

A multifaceted sensitivity analysis of the Slovenian public opinion survey data

Peer-reviewed author version

BEUNCKENS, Caroline; SOTTO, Cristina; MOLENBERGHS, Geert & VERBEKE, Geert (2009) A multifaceted sensitivity analysis of the Slovenian public opinion survey data. In: JOURNAL OF THE ROYAL STATISTICAL SOCIETY SERIES C-APPLIED STATISTICS, 58. p. 171-196.

DOI: 10.1111/j.1467-9876.2009.00647.x

Handle: <http://hdl.handle.net/1942/9516>



A multifaceted sensitivity analysis of the Slovenian public opinion survey data [Link](#)

**Peer-reviewed author version**

Made available by Hasselt University Library in [Document Server@UHasselt](#)

**Reference** (Published version):

Beunckens, Caroline; Sotto, Cristina Remedio; Molenberghs, Geert & Verbeke, Geert(2009) A multifaceted sensitivity analysis of the Slovenian public opinion survey data. In: JOURNAL OF THE ROYAL STATISTICAL SOCIETY SERIES C-APPLIED STATISTICS, 58. p. 171-196

DOI: 10.1111/j.1467-9876.2009.00647.x

Handle: <http://hdl.handle.net/1942/9516>

# A Multi-faceted Sensitivity Analysis of the Slovenian Public Opinion Survey Data

Caroline Beunckens<sup>1</sup>, Cristina Sotto<sup>1,2</sup>, Geert Molenberghs<sup>1,3</sup>,  
and Geert Verbeke<sup>3,1</sup>

<sup>1</sup> Center for Statistics, Hasselt University,  
Agoralaan, Building D, B3590 Diepenbeek, Belgium.

<sup>2</sup> School of Statistics, University of the Philippines,  
Diliman, Quezon City, Philippines.

<sup>3</sup> Biostatistical Centre, Katholieke Universiteit Leuven, Belgium.

## Abstract

Many models to analyze incomplete data have been developed that allow the missing data to be missing not at random. Awareness has grown that such models are based on unverifiable assumptions, in the sense that they rest on the (incomplete) data only in part, but that inferences nevertheless depend on what the model predicts about the unobserved data, given the observed data. This explains why, nowadays, considerable work is being devoted to assess how sensitive models for incomplete data are to: (1) the particular model chosen; (2) a family of models chosen; (3) the impact of (a group of) influential subjects. For each of these categories, a number of proposals has been formulated, studied theoretically and/or by way of simulations, and applied to sets of data. It is, however, uncommon to explore various sensitivity analysis avenues simultaneously. In this paper, we apply a collection of such tools, some after extension, to incomplete counts arising from cross-classified binary data from the so-called Slovenian Public Opinion Survey. Thus for the first time bringing together a variety of sensitivity analysis tools on the same set of data, we are in a position to sketch a comprehensive sensitivity analysis picture. We will show that MAR estimates of the proportion voting in favour of independence are insensitive to the precise choice of MAR model and close to the actual plebiscite results, whereas the MNAR models that are furthest from the plebiscite results are vulnerable to the impact of outlying cases. Our approach helps to illustrate the value of comprehensive sensitivity analysis. Ideas are formulated as to the methodology's use beyond the data analysis considered here.

*Keywords:* categorical data; interval of ignorance; local influence; missing at random; missing not at random

# 1 Introduction

In 1991, Slovenians voted for independence from Yugoslavia in a plebiscite. In anticipation of this result, the Slovenian government collected data on its possible outcome by inserting questions in the so-called Slovenian Public Opinion (SPO) Survey. The survey, administered to 2074 voting-age Slovenians, was conducted a month prior to the plebiscite using a three-stage sampling design (Barnett, 2002). Rubin, Stern and Vehovar (1995) studied the three fundamental questions added to the SPO and, in comparison to the plebiscite's outcome, drew conclusions about the missing data process. We will analyze the data as an ordinary contingency table with incomplete margins, ignoring sample survey features, but in line with analyses conducted by Rubin, Stern and Vehovar (1995).

The three questions added were: (1) Are you in favour of Slovenian independence? (2) Are you in favour of Slovenia's secession from Yugoslavia? (3) Will you attend the plebiscite? In spite of their apparent equivalence, questions (1) and (2) are different because independence had been possible in a transition from the existing federal structure to a looser, confederal, form as well, and therefore the secession question is added. Question (3) is highly relevant because the political decision was taken that not attending was treated as an effective NO to question (1). Thus, the primary estimand is the proportion  $\theta$  of people that will be considered as voting YES, and in the context of the three questions, this can be defined as the fraction of people answering YES on *both* the independence and attendance questions (1) and (3), respectively, regardless of their response to question (2). A cross-classification of the attendance versus independence questions is presented in Table 2.

Clearly, the data are incomplete, hampering the straightforward estimation of  $\theta$ . Missingness adds a source of uncertainty and it is useful to distinguish between two types of *statistical uncertainty*. The first one, *statistical imprecision*, stems from finite sampling. The Slovenian Public Opinion Survey included not all Slovenians but only 2074 respondents. However, even if all would have been included, there would have been residual uncertainty because some fail to report at least one answer. This second source of uncertainty, stemming from incompleteness, will be called *statistical ignorance*. Statistical imprecision is classically quantified by means of estimators (standard error and variance, confidence region,...) and properties of estimators (consistency, asymptotic distribution, efficiency,...). In order to quantify statistical ignorance, it is useful to distinguish between full and observed data.

**Table 1:** Theoretical distribution of the probability mass over full and observed cells, respectively, for a bivariate binary outcome with missingness in none, one, or both responses.

(a) Complete cells

$\pi_{11,11}$	$\pi_{11,12}$	$\pi_{10,11}$	$\pi_{10,12}$	$\pi_{01,11}$	$\pi_{01,12}$	$\pi_{00,11}$	$\pi_{00,12}$
$\pi_{11,21}$	$\pi_{11,22}$	$\pi_{10,21}$	$\pi_{10,22}$	$\pi_{01,21}$	$\pi_{01,22}$	$\pi_{00,21}$	$\pi_{00,22}$

(b) Observed cells

$\pi_{11,11}$	$\pi_{11,12}$	$\pi_{10,1+}$	$\pi_{01,+1}$	$\pi_{01,+2}$	$\pi_{00,++}$
$\pi_{11,21}$	$\pi_{11,22}$	$\pi_{10,2+}$			

We focus on two binary questions, such as the independence and attendance questions in the Slovenian Public Opinion Survey. The 16 theoretical full cell probabilities are as in Table 1a, producing 15 full data degrees of freedom. The generic expression for the cell probabilities is  $\pi_{r_1 r_2, j_1 j_2} = P(R_1 = r_1, R_2 = r_2, J_1 = j_1, J_2 = j_2)$ , where  $R_1$  and  $R_2$  denote the missingness indicators for the independence and attendance questions, respectively, while  $J_1$  and  $J_2$  denote the actual responses for the independence and attendance questions, respectively. Further,  $r_1 = 0$  (1) if the answer to the independence question is missing (observed) and  $j_1 = 1$  (2) if the answer to the independence question is yes (no). The indices  $r_2$  and  $j_2$  are similarly defined for the attendance question.

Similarly, the 9 observed cells are shown in Table 1b. Presenting cell counts instead of probabilities gives Table 2, which also shows the cell indexing system used in the rest of the paper.

The full cell probabilities in Table 1a are four-way joint probabilities of the two missingness indicators and the two binary responses. Moreover, the parameter of interest,  $\theta$ , i.e., the proportion of people answering YES to *both* the independence and attendance questions, is the marginal probability that the two responses are both equal to 1 (marginalized over the missingness indicators), and can be expressed as:

$$\theta = \sum_{r_1=0}^1 \sum_{r_2=0}^1 \pi_{r_1 r_2, 11} = \sum_{r_1=0}^1 \sum_{r_2=0}^1 P(R_1 = r_1, R_2 = r_2, J_1 = 1, J_2 = 1).$$

It is clear from the above expression that when the data are complete, evaluation of  $\theta$  would simply entail summing the upper-left cells from each the four  $(2 \times 2)$  tables in Table 1a. However, when

**Table 2:** Observed cells for the Slovenian Public Opinion Survey, collapsed over the secession question. A simplified cell indexing system is shown in relation to the original cell indexing system. The columns refer to ‘independence’ while the rows refer to ‘attendance.’

Cell 1 $Z_{11,11} = 1439$	Cell 2 $Z_{11,12} = 78$	Cell 5 $Z_{10,1+} = 159$	Cell 7 $Z_{01,+1} = 144$	Cell 8 $Z_{01,+2} = 54$	Cell 9 $Z_{00,++} = 136$
Cell 3 $Z_{11,21} = 16$	Cell 4 $Z_{11,22} = 16$	Cell 6 $Z_{10,2+} = 32$			

the data are incomplete, as in Table 1b, the relation is not straightforward as the collapsed cells have to be split in order to obtain values for the cells pertinent to the estimation of  $\theta$ .

We refer to the mechanism governing missingness using the taxonomy of Rubin (1976). A mechanism is *missing completely at random* (MCAR) if the processes governing missingness and outcomes are independent, perhaps conditional on covariates. *Missing at random* (MAR) is the situation where missingness may depend on observed outcomes and covariates but, given these, not further on unobserved outcomes. When, in addition to such dependencies, the unobserved data provide further information about the missing data mechanism, then we name the mechanism *missing not at random* (MNAR). Note that MAR relaxes MCAR, and that MNAR relaxes MAR. Additional concepts will be introduced in the rest of the paper as the need arises.

It is our goal to conduct a number of sensitivity analyses on this set of data. While this has been done before for this and many other examples, it is less customary to consider several sensitivity analyses simultaneously. Some tools will be used in their currently available form, while others will be extended and adapted for our purposes.

The aims of our analyses are: (1) sensitivity analysis for the Slovenian Public Opinion Survey Data, with emphasis on  $\theta$ , the proportion of people voting in favour of independence, and (2) what can be learned from this about how sensitivity analyses could be carried out in other settings. Using these complementary approaches, we are able to show that MAR estimates of  $\theta$  are insensitive to the precise choice of MAR model and close to the actual plebiscite results, whereas the MNAR models that are furthest from the plebiscite results, the so-called pessimistic ones, are markedly less robust. Influence-analysis techniques will underscore that the seeming credibility of the more pessimistic scenarios owes to a few influential subjects only. We thus argue that our approach illustrates the value of comprehensive sensitivity analysis.

In Section 2, a tour is made of a number of simple analyses. The family of models proposed by Baker, Rosenberger and DerSimonian (1992), to be used in the remainder of the paper, is introduced in Section 3. Three main strands of sensitivity analysis are then presented in Section 4, i.e., the interval of ignorance (Section 4.1), global influence (Section 4.3), local influence (Section 4.4), and the computation of a so-called MNAR counterpart to the model considered (Section 4.2). We thereby bring together and contrast existing sensitivity assessments with new, local influence based analyses that have never been applied to the SPO data. Moreover, the local influence technology based on cell counts rather than parameters is new, as well as the approach of perturbing the cell probabilities rather than the model parameters. The lessons to be learned from undertaking a combined sensitivity analysis, both from a practitioner's as well as from a modeller's point of view, are touched upon in Section 5.

## 2 Review of Simple Analyses

The data were used by Molenberghs, Kenward and Goetghebeur (2001) to illustrate their proposed sensitivity analysis tool, the interval of ignorance. Molenberghs *et al* (2007) used the data to exemplify results about the relationship between MAR and MNAR models. An overview of various analyses can be found in Molenberghs and Kenward (2007). These authors used the models proposed by Baker, Rosenberger and DerSimonian (1992) for the setting of two-way contingency tables, subject to missingness in either none, one, or both responses. Rubin, Stern and Vehovar (1995) conducted several analyses of the data. Their main emphasis was on determining the proportion  $\theta$  of the population that would attend the plebiscite and vote for independence. Their estimates are reproduced in Table 3.

The pessimistic (optimistic) bounds, or non-parametric bounds, are obtained by setting all incomplete data that can be considered a yes (no), as yes (no). The complete case estimate for  $\theta$  is based on the subjects answering all three questions and the available case estimate is based on the subjects answering the two questions of interest here. It is noteworthy that both of these estimates are out of bounds. This is not a mistake: it should be recalled that the political decision was made to treat a NO on the attendance question as an effective NO vote in the plebiscite. Disregarding incomplete cases ignores this aspect and thus discards available information, thereby causing the estimate to exceed the bounds. Note that the bounds apply to only those estimators making use of all available data, not to estimators based on subsets.

**Table 3:** *The Slovenian Public Opinion Survey. Some estimates of the proportion  $\theta$  attending the plebiscite and voting for independence, as presented in Rubin, Stern and Vehovar (1995) and Molenberghs, Kenward and Goetghebeur (2001), as well as sensitivity-analysis results.*

Estimation method	Voting in favour of independence: $\hat{\theta}$
Standard analyses	
Non-parametric bounds	[0.694;0.905]
Complete cases	0.928
Available cases	0.929
MAR (2 questions)	0.892
MAR (3 questions)	0.883
MNAR	0.782
Sensitivity analyses	
BRD1–BRD9	[0.741;0.892]
Well-fitting BRD6–BRD9	[0.741;0.867]
BRD1(MAR)–BRD9(MAR)	[0.892;0.892]
II: Model 10	[0.762;0.893]
II: Model 11	[0.766;0.883]
II: Model 12	[0.694;0.905]
Plebiscite	0.885

This underscores the growing conviction that such estimates should be disregarded in favour of more routine use of MAR (Molenberghs and Kenward, 2007). Nevertheless, it may sometimes be worthwhile to consider the complete case (CC) analysis as simply one valid under MCAR. Ideally, it should be set against the background of bounds, as is done here, or at least a number of competing models. Rubin, Stern and Vehovar (1995) considered two MAR models, also reported in Table 3, the first one solely based on the two questions of direct interest, the second one using all three. Finally, they considered a single MNAR model, based on the assumption that missingness on a question depends on the answer to that question but not on the other questions. Rubin, Stern and Vehovar (1995) concluded, owing to the proximity of the MAR analysis to the plebiscite value, that MAR in this and similar cases may be considered a plausible assumption. As argued before (Kenward, Goetghebeur and Molenberghs, 2001), one has to be careful with this conclusion, however. Arguments to support this position will be provided in Section 4.1, based on the BRD family introduced next.



### 3 The BRD Family of Models

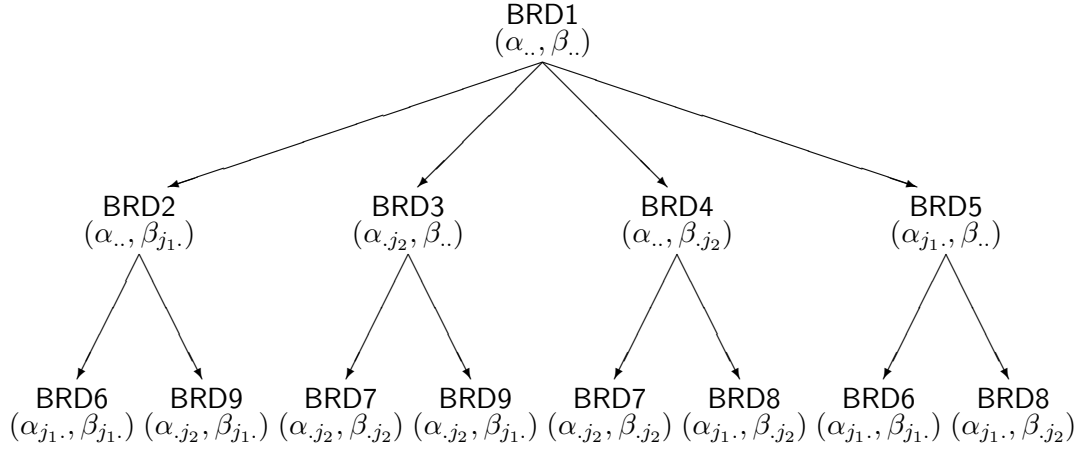
Baker, Rosenberger and DerSimonian (1992) proposed a log-linear based family of models for the four-way classification of both outcomes, together with their missingness indicators. In line with the notation in Table 1a, we denote the cell counts by  $Z_{r_1 r_2, j_1 j_2}$ , where  $r_1, r_2 = 0, 1$  indicates, respectively, whether the measurement is missing or observed at occasions 1, 2, and  $j_1, j_2 = 1, 2$  indicates the response categories for both outcomes. The models are written as:

$$\begin{aligned} E(Z_{11, j_1 j_2}) &= Z_{++, ++} \pi_{11, j_1 j_2} = m_{j_1 j_2} & E(Z_{01, j_1 j_2}) &= m_{j_1 j_2} \alpha_{j_1 j_2} \\ E(Z_{10, j_1 j_2}) &= m_{j_1 j_2} \beta_{j_1 j_2} & E(Z_{00, j_1 j_2}) &= m_{j_1 j_2} \alpha_{j_1 j_2} \beta_{j_1 j_2} \gamma. \end{aligned}$$

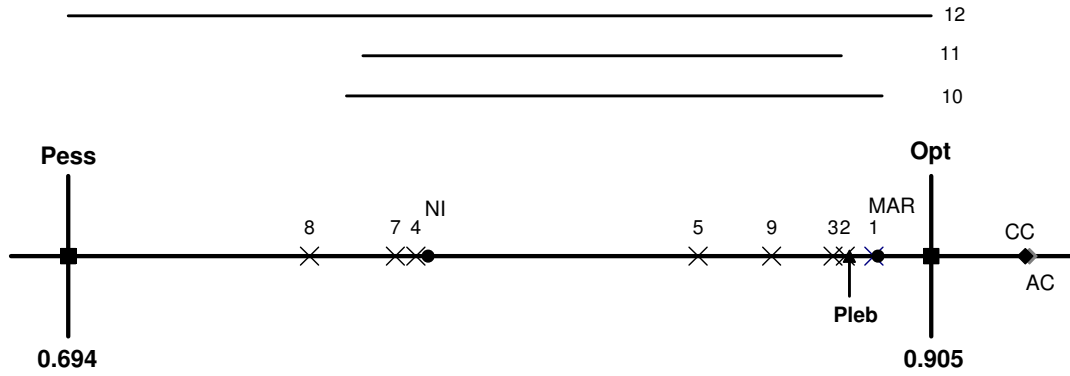
The  $\alpha$  ( $\beta$ ) parameters describe missingness in the independence (attendance) question as the proportion of subjects with a missing response on the independence (attendance) question relative to the proportion of subjects with both responses present, given a particular response combination ( $j_1, j_2$ ). The  $\gamma$  parameter, on the other hand, captures the interaction between the two missingness indicators via the (conditional) odds ratio for a given response combination. The subscripts are missing from  $\gamma$  since Baker, Rosenberger and DerSimonian (1992) have shown that this quantity needs to be independent of  $j_1$  and  $j_2$  in any identifiable model.

Baker, Rosenberger and DerSimonian (1992) considered nine models, based on setting  $\alpha_{j_1 j_2}$  and  $\beta_{j_1 j_2}$  constant in one or more indices, and can be enumerated using the ‘BRD’ abbreviation as displayed in Figure 1, together with their nesting structure. Interpretation is straightforward; for example, in BRD1, both missingness indicators do not depend on the responses, thereby characterizing an MCAR mechanism. In view, however, of the nesting structure among the different types of mechanisms described in Section 1, although BRD1 is more precisely classified as MCAR, it can also be considered a special case of MAR, or even MNAR. For instance, in BRD4, missingness in the first variable is constant, while missingness in the second variable depends on its value. Moreover, BRD6–BRD9 saturate the observed data degrees of freedom, while the lower numbered ones leave room for evaluating the goodness-of-fit of the model to the observed data.

Molenberghs, Kenward and Goetghebeur (2001) and Molenberghs *et al* (2007) fitted the BRD models; Table 4 summarizes the results. BRD1 produces  $\hat{\theta} = 0.892$ , exactly the same as the first MAR estimate obtained by Rubin, Stern and Vehovar (1995). This does not come as a surprise, because BRD1, though MCAR, belongs to the MAR family, as does Rubin, Stern and Vehovar (1995)’s model; both use information from the two main questions. A graphical representation



**Figure 1:** Graphical representation of the BRD model nesting structure.



**Figure 2:** The Slovenian Public Opinion Survey. Relative position for the estimates of “proportion of YES votes”, based on the models considered in Rubin, Stern and Vehovar (1995) and on the BRD Models. The vertical lines indicate the nonparametric pessimistic-optimistic Bounds. (Pess(■): pessimistic boundary; Opt(■): optimistic boundary; MAR(●): Rubin et al’s MAR model; NI(●): Rubin et al ’s MNAR model; AC(♦): available cases; CC(♦): complete cases; Pleb(▲): plebiscite outcome. Crosses (×) and the numbers above them refer to the BRD models. Intervals of ignorance (Models 10–12) are represented by horizontal bars.)

of the original analyses and the BRD models combined is given in Figure 2.

## 4 Sensitivity Analysis

We will use the working definition that a sensitivity analysis is one in which several statistical models are considered simultaneously and/or where a statistical model is further scrutinized

**Table 4:** *The Slovenian Public Opinion Survey. Analysis restricted to the independence and attendance questions. Summaries on each of the Models BRD1–BRD9 are presented, with obvious column labels. The column labelled ' $\hat{\theta}_{MAR}$ ' refers to the model corresponding to the given one, with the same fit to the observed data, but with missing data mechanism of the MAR type.*

Model	Structure	d.f.	loglik	$\hat{\theta}$	C.I.	$\hat{\theta}_{MAR}$
BRD1	$(\alpha, \beta)$	6	-2495.29	0.892	[0.878;0.906]	0.8920
BRD2	$(\alpha, \beta_{j_1})$	7	-2467.43	0.884	[0.869;0.900]	0.8915
BRD3	$(\alpha_{j_2}, \beta)$	7	-2463.10	0.881	[0.866;0.897]	0.8915
BRD4	$(\alpha, \beta_{j_2})$	7	-2467.43	0.765	[0.674;0.856]	0.8915
BRD5	$(\alpha_{j_1}, \beta)$	7	-2463.10	0.844	[0.806;0.882]	0.8915
BRD6	$(\alpha_{j_1}, \beta_{j_1})$	8	-2431.06	0.819	[0.788;0.849]	0.8919
BRD7	$(\alpha_{j_2}, \beta_{j_2})$	8	-2431.06	0.764	[0.697;0.832]	0.8919
BRD8	$(\alpha_{j_1}, \beta_{j_2})$	8	-2431.06	0.741	[0.657;0.826]	0.8919
BRD9	$(\alpha_{j_2}, \beta_{j_1})$	8	-2431.06	0.867	[0.851;0.884]	0.8919

using specialized tools, such as diagnostic measures. This informal definition encompasses a wide variety of useful approaches. The simplest procedure is to fit a selected number of (non-random) models which are all deemed plausible or in which a preferred (primary) analysis is supplemented with a number of variations. The extent to which conclusions (inferences) are stable across such ranges provides an indication about the degree to which they are robust to inherently untestable assumptions about the missingness mechanism. Variations to a basic model can be constructed in different ways. The most obvious strategy, cast within the selection model paradigm, is to consider various dependencies of the missing data process on the outcomes and/or on covariates. Alternatively, the distributional assumptions of the models can be changed. This route will be followed in Section 4.1. Thijs *et al* (2000) consider sensitivity analysis within the context of pattern-mixture models.

Related to this, we can assess how an MNAR model, or a collection of MNAR models, differ from the set of models with equal fit to the observed data but that are of a MAR nature. This path is followed in Section 4.2.

Additionally, a sensitivity analysis can also be performed on the level of individual observations instead of on the level of the models. In that case, interest is directed towards finding those individuals who drive the conclusions towards one or more MNAR models. Therefore, the influence of every individual separately will be explored. Two techniques exist, i.e., global influence (Sec-

tion 4.3) and local influence (Section 4.4, Cook 1986). The global influence methodology, also known as the case-deletion method (Cook and Weisberg, 1986), is introduced by Cook (1979, 1986) in linear regression, and by Molenberghs *et al* (2003) and Thijs, Molenberghs and Verbeke (2000) in linear mixed models. Verbeke *et al* (2001); Thijs, Molenberghs and Verbeke (2000) already used local influence on the Diggle and Kenward (1994) model, which is based on a selection model, integrating a linear mixed model for continuous outcomes with logistic regression for dropout. Later, Van Steen *et al* (2001) adapted these ideas to the model of Molenberghs, Kenward and Lesaffre (1997), for monotone repeated ordinal data.

## 4.1 Interval of Ignorance

A sample from Table 1 produces empirical proportions representing the  $\pi$ 's with error. This imprecision disappears asymptotically. What remains is ignorance regarding the redistribution of all but the first four  $\pi$ 's over the missing values. This leaves ignorance regarding any probability in which at least one of the first or second indices is equal to 0, and hence regarding any derived parameter of scientific interest. For such a parameter,  $\theta$  say, a region of possible values which is consistent with Table 1 is called a region of ignorance. Evidently, such a region will depend, not only on the data and the way it is incomplete, but also on the model for which it is constructed. Analogously, an observed incomplete table leaves ignorance regarding the would-be observed full table, which leaves imprecision regarding the true full probabilities. The region of estimators for  $\theta$  consistent with the observed data provides an estimated region of ignorance. For a single parameter, the region becomes the *interval of ignorance*. Various ways of constructing regions of ignorance are conceivable. Practically, one selects the largest possible set of identifiable parameters. The remaining ones are then termed sensitivity parameters. For every value chosen for the latter, the former can be estimated by means of, for example, maximum likelihood. Repeating this for all values of the sensitivity parameters or, practically speaking, a sufficiently refined grid, one effectively obtains a region or, in the univariate case, an interval of estimates. The  $(1 - \alpha)100\%$  *region of uncertainty* is a larger region, encompassing the region of ignorance, in the spirit of a confidence region, designed to capture the combined effects of imprecision and ignorance. Practically, for every point in the region of ignorance, a confidence region is constructed, the union of which then produces the interval of uncertainty. Details regarding construction and asymptotic properties can be found in Molenberghs, Kenward and Goetghebeur (2001), Kenward, Goetghebeur and Molenberghs (2001), and Vansteelandt *et al* (2006).

**Table 5:** *The Slovenian Public Opinion Survey. Intervals of ignorance and intervals of uncertainty for the proportion  $\theta$  (confidence interval) attending the plebiscite, following from fitting overspecified Models 10, 11, and 12.*

Model	d.f.	loglik	$\hat{\theta}$	
			II	IU
10	9	-2431.06	[0.762;0.893]	[0.744;0.907]
11	9	-2431.06	[0.766;0.883]	[0.715;0.920]
12	10	-2431.06	[0.694;0.905]	

The estimated intervals of ignorance and intervals of uncertainty are shown in Table 5, while a graphical representation of the YES votes is given in Figure 2. Model 10 is defined as  $(\alpha_{j_2}, \beta_{j_1 j_2})$  with

$$\beta_{j_1 j_2} = \beta_0 + \beta_{j_1} + \beta_{j_2}, \quad (1)$$

while Model 11 assumes  $(\alpha_{j_1 j_2}, \beta_{j_1})$  and uses

$$\alpha_{j_1 j_2} = \alpha_0 + \alpha_{j_1} + \alpha_{j_2}, \quad (2)$$

Finally, Model 12 is defined as  $(\alpha_{j_1 j_2}, \beta_{j_1 j_2})$ , a combination of both (1) and (2). Model 10 shows an interval of ignorance which is very close to [0.741, 0.892], the range produced by the models BRD1–BRD9, while Model 11 is somewhat sharper and just fails to cover the plebiscite value. However, the corresponding intervals of uncertainty contain the true value.

Interestingly, Model 12 virtually coincides with the non-parametric range even though it does not saturate the complete data degrees of freedom. To do so, not 2 but in fact 7 sensitivity parameters would have to be included. Thus, it appears that a relatively simple sensitivity analysis is adequate to increase insight in the information provided by the incomplete data about the proportion of valid YES votes, in the study under consideration here.

## 4.2 An MAR Counterpart for an MNAR Model

Molenberghs *et al* (2007) showed that, strictly speaking, the correctness of an alternative, MNAR, model can only be verified in as far as it fits the *observed* data. Thus, evidence for or against MNAR can only be provided within a particular, predefined parametric family, whose plausibility

cannot be verified in empirical terms alone. This implies that an overall (omnibus) assessment of MAR *versus* MNAR is not possible, because every MNAR model can be doubled up with a uniquely defined MAR counterpart, producing exactly the same fit as the original MNAR model, in the sense that it leads to entirely the same predictions to the observed data (e.g., fitted counts in an incomplete contingency table) as the original MNAR model, and depending on exactly the same parameter vector. ‘Unique’ means that, for a given MNAR model, there is one and only one MAR model with the specified features corresponding to it. However, an entire class of MNAR models will share such an MAR counterpart, we have a many-to-one map (cf. related ideas in Tsiatis (1975)). We construct the companion as follows: (1) an MNAR model is fitted to the data; (2) the fitted model is reformulated in a pattern-mixture model form; (3) the density or distribution of the unobserved measurements given the observed ones, and given a particular response pattern, is replaced by its MAR counterpart; and (4) it is established that such an MAR counterpart uniquely exists.

First, an MNAR model is fitted to the observed set of data. Second, the likelihood contribution for subject  $i$ , expressed in pattern-mixture form, is:

$$f(\mathbf{y}_i^o | \mathbf{r}_i, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\psi}}) f(\mathbf{r}_i | \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\psi}}) f(\mathbf{y}_i^m | \mathbf{y}_i^o, \mathbf{r}_i, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\psi}}). \quad (3)$$

Third, based on Molenberghs *et al* (1998), we use the fact that MAR within the PMM framework means:  $f(\mathbf{y}_i^m | \mathbf{y}_i^o, \mathbf{r}_i, \boldsymbol{\theta}) = f(\mathbf{y}_i^m | \mathbf{y}_i^o, \boldsymbol{\theta})$ . Hence, clearly  $f(\mathbf{y}_i^m | \mathbf{y}_i^o, \mathbf{r}_i, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\psi}})$  needs to be replaced by

$$h(\mathbf{y}_i^m | \mathbf{y}_i^o, \mathbf{r}_i) = h(\mathbf{y}_i^m | \mathbf{y}_i^o) = f(\mathbf{y}_i^m | \mathbf{y}_i^o, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\psi}}), \quad (4)$$

where the  $h(\cdot)$  notation is used for brevity. Thus, we have a unique way of extending the model fit to the observed data, within the MAR family.

The key computational consequence is the need to obtain  $h(\mathbf{y}_i^m | \mathbf{y}_i^o)$  in (4). This means, for each pattern, the conditional density of the unobserved measurements given the observed ones needs to be extracted from the marginal distribution of the complete set of measurements. Suggestions for implementation can be found in Molenberghs *et al* (2007). Again, the main consequence is that one cannot test MNAR against MAR, without making unverifiable assumptions about the alternative model.

We illustrate the use of the MAR counterpart by means of 4 models from the BRD family, fitted to the independence and attendance outcomes. We select models BRD1, BRD2, BRD7, and BRD9.

Although BRD1 assumes missingness to be MCAR, it can, in fact, be viewed as belonging to the MAR, and more generally, to the MNAR family. It is thus possible although trivial to obtain its MAR counterpart. All other BRD models are MNAR. Model BRD2 has 7 free parameters, and hence does not saturate the observed data degrees of freedom, while models BRD7 and BRD9 saturate the 8 data degrees of freedom. The collapsed data, together with the model fits, are displayed in Table 6. Each of the four models is doubled up with its MAR counterpart.

Table 6 presents, apart from the raw data, for each of the models and its MAR counterpart, the fit to the observed and the hypothetical complete data. The fits of models BRD7, BRD9, and their MAR counterparts to the observed data, coincide with the observed data. As the theory states, every MNAR model and its MAR counterpart produce exactly the same fit to the observed data, which is therefore also seen for BRD1 and BRD2. However, while Models BRD1 and BRD1(MAR) coincide in their fit to the hypothetical complete data, this is not the case for the other models. The reason is clear: since BRD1, though MCAR, belongs to the MAR family from the start, its counterpart BRD1(MAR) does not produce any difference, but merely copies the fit of BRD1 to the unobserved data, given the observed ones. BRD1 and BRD1(MAR) are therefore identical. This is a feature shared with all MAR models, and hence, in particular, with all MCAR models. Finally, while BRD7 and BRD9 produce a different fit to the complete data, BRD7(MAR) and BRD9(MAR) coincide. This is because the fits of BRD7 and BRD9 coincide with respect to their fit to the observed data, and indeed, due to their saturation, coincide with the observed data as such. This fit is the sole basis for the models' MAR extensions. It is noteworthy that, while BRD7, BRD9, and  $\text{BRD7(MAR)} \equiv \text{BRD9(MAR)}$  all saturate the observed data degrees of freedom, their complete-data fits are dramatically different.

Before continuing with the models' interpretation, it is necessary to assess their fit. Conducting likelihood ratio tests for BRD1 versus the ones with 7 parameters, BRD2–BRD5, and then in turn for BRD2–BRD5 versus the saturated models BRD6–BRD9, suggests the lower numbered models do not fit well, leaving us with BRD6–BRD9. The impression might be generated that the poor model fit of BRD1 might be seen as evidence for discarding the MAR-based value 0.892. However, studying the MAR values from each of BRD1(MAR)–BRD9(MAR), as displayed in the last column of Table 4, this value is remarkably stable and hence a value of  $\hat{\theta} = 0.892$ , based on the four counterparts BRD6(MAR)–BRD9(MAR), is a sensible choice after all. Thus, a main contribution resulting from considering the counterparts in this particular example, is the provision

**Table 6:** *The Slovenian Public Opinion Survey. Analysis restricted to the independence and attendance questions. The observed data are shown, as well as the fit of models BRD1, BRD2, BRD7, and BRD9, and their MAR counterparts, to the observed data and to the hypothetical complete data. The contingency tables' rows (columns) correspond to YES vs. NO on the independence (attendance) question.*

Observed data & fit of BRD7, BRD7(MAR), BRD9, and BRD9(MAR) to incomplete data									
1439	78	159		144	54	136			
16	16	32							
Fit of BRD1 and BRD1(MAR) to incomplete data									
1381.6	101.7	182.9		179.7	18.3	136.0			
24.2	41.4	8.1							
Fit of BRD2 and BRD2(MAR) to incomplete data									
1402.2	108.9	159.0		181.2	16.8	136.0			
15.6	22.3	32.0							
Fit of BRD1 and BRD1(MAR) to complete data									
1381.6	101.7	170.4	12.5	176.6	13.0	121.3	9.0		
24.2	41.4	3.0	5.1	3.1	5.3	2.1	3.6		
Fit of BRD2 to complete data									
1402.2	108.9	147.5	11.5	179.2	13.9	105.0	8.2		
15.6	22.3	13.2	18.8	2.0	2.9	9.4	13.4		
Fit of BRD2(MAR) to complete data									
1402.2	108.9	147.7	11.3	177.9	12.5	121.2	9.3		
15.6	22.3	13.3	18.7	3.3	4.3	2.3	3.2		
Fit of BRD7 to complete data									
1439	78	3.2	155.8	142.4	44.8	0.4	112.5		
16	16	0.0	32.0	1.6	9.2	0.0	23.1		
Fit of BRD9 to complete data									
1439	78	150.8	8.2	142.4	44.8	66.8	21.0		
16	16	16.0	16.0	1.6	9.2	7.1	41.1		
Fit of BRD7(MAR) and BRD9(MAR) to complete data									
1439	78	148.1	10.9	141.5	38.4	121.3	9.0		
16	18	11.8	20.2	2.5	15.6	2.1	3.6		



of a solid basis for the MAR-based estimate. There is every reason to reject the MAR/MCAR estimate from  $BRD1 \equiv BRD1(MAR)$ , since the model simply does not fit the observed data well enough. At first sight, this is strange since the estimate happens to be close to the plebiscite value. Now, essentially the same estimate results from *all* MAR counterparts, including  $BRD6(MAR)$ – $BRD9(MAR)$ , which saturate the observed data. Hence, the MAR estimate can be deemed at the same time consistent with the observed data *and* close to the plebiscite value. Obviously, because Models  $BRD6(MAR)$ – $BRD9(MAR)$  are exactly the same and of perfect fit, the corresponding probabilities  $\hat{\theta}_{MAR}$  are equal, too. In this particular case, even though  $BRD2(MAR)$ – $BRD5(MAR)$  differ among each other, the probability of being in favor of independence and attending the plebiscite is constant across these four models. This is a mere coincidence, since all three other cell probabilities are different, although only slightly so. For example, the probability of being in favour of independence while not attending ranges over 0.066–0.0685 across these four models.

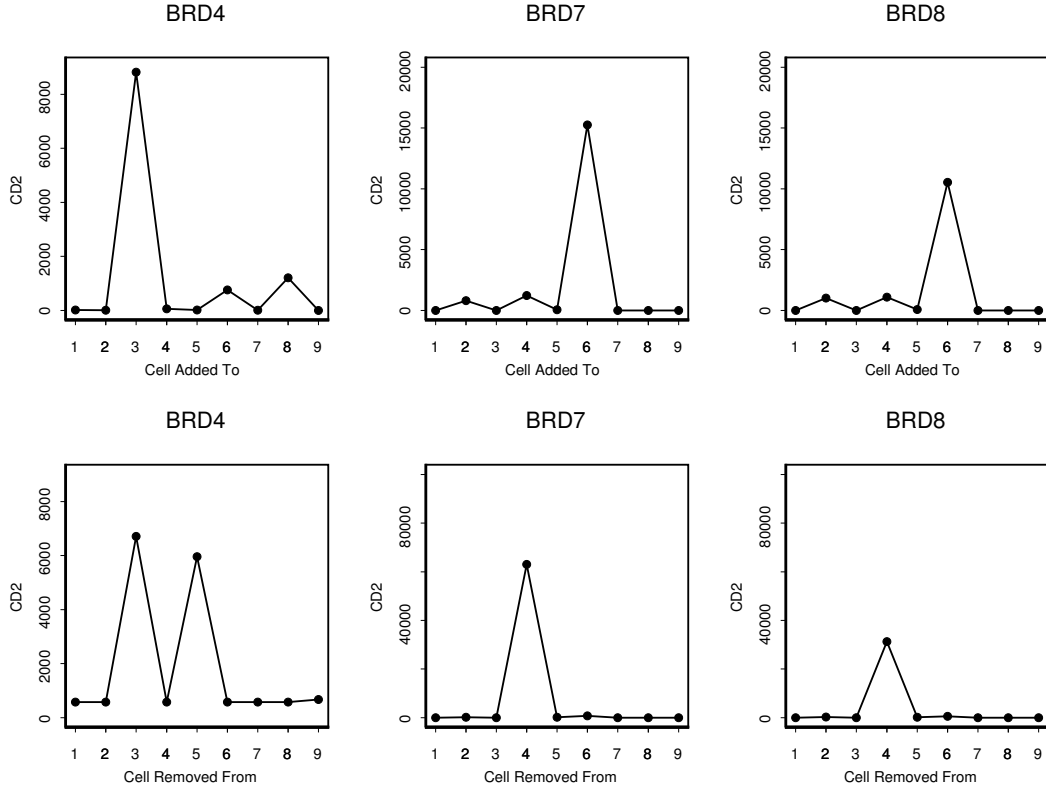
We have made ‘two-stage’ use of Models  $BRD6(MAR)$ – $BRD9(MAR)$ . At the first stage, in a conventional way, the fully saturated model is selected as the only adequate description of the observed data. At the second stage, these models are transformed to their MAR counterpart, from which inferences are drawn. The MAR counterpart thus usefully supplements the original models  $BRD6$ – $9$  and provide a further, important scenario to model the incomplete data.

### 4.3 Global Influence

One sensitivity-analysis tool is global influence, which starts from case deletion and is based on the difference in log-likelihood between the model fitted to the entire data set  $\ell(\phi) = \sum_{i=1}^N \ell_i(\phi)$ , with  $\ell_i(\phi)$  the contribution of the  $i^{\text{th}}$  individual, on the one hand, and the data set minus one subject or a subject doubled,  $\ell_{(\pm i)}(\phi)$ , on the other hand. Here,  $\phi$  parameterizes the particular BRD model. Cook’s distances (CD) are based on measuring the discrepancy, induced by deletion or doubling, in either the likelihood or the parameter vector:

$$CD_{1i} = 2 \left[ \hat{\ell}(\phi) - \hat{\ell}_{(\pm i)}(\phi) \right] \quad \text{or} \quad CD_{2i} = 2(\hat{\phi} - \hat{\phi}_{(\pm i)})' \ddot{L}^{-1} (\hat{\phi} - \hat{\phi}_{(\pm i)}), \quad (5)$$

with  $\ddot{L}$  the matrix of second derivatives of  $\ell(\phi)$ , with respect to  $\phi$ , evaluated at  $\hat{\phi}$ . Our focus is on the latter version, because interest is on changes in the parameter estimates rather than on the likelihood. Both measures can be constructed for the entire parameter vector or for sub-vectors thereof; this includes, of course, a single parameter. Performing a global influence analysis on



**Figure 3:** *The Slovenian Public Opinion Survey. Global influence analysis for BRD4, BRD7 and BRD8. Cook's distance measure,  $CD_{2i}$ , is evaluated when an observation is added to a specific cell (first row) and when an observation is deleted from a specific cell (second row).*

data with categorical outcomes is less time consuming than on data with continuous outcomes, since the data can then be organized into cells, as in Table 2.

Figure 3 shows a selection of the results for the global influence analysis on the SPO survey data. Inasmuch as Cook's distance measure  $CD_{2i}$  was approximately zero for all cells, indicating no substantial influence when adding or removing a single case from a particular cell, for all other models, only the results for BRD4, BRD7, and BRD8 are presented. Observe that, for BRD4, adding a single observation to cell 3 has a large influence on the parameters, as well as deletion from either cells 3 or 5. Cell 3 represents subjects with a NO on the attendance question and a YES on the independence question. An addition or removal of one such respondent can largely affect the parameters of BRD4. Similarly, exclusion of a single respondent with a NO on the independence question but a missing response on the attendance question (cell 5), also influences BRD4's model parameters, though to a lesser extent. For models BRD7 and BRD8, an additional observation in cell 6 or a deletion from cell 4 leads to significant influence on

these models' parameters. Thus, adding a subject with a YES for independence and a missing attendance response, or excluding a respondent with YES on both questions, yields changes in the model parameters of BRD7 and BRD8. These findings hint on the influential nature of subjects with a YES on the independence question, which is likely related with this group's sparseness.

#### 4.4 Local Influence

A drawback of global influence is that the specific cause of the influence is hard to retrieve because, by deleting or adding a subject, all types of influence stemming from it are lumped together. Local influence (Cook, 1986; Verbeke *et al*, 2001) allegedly is more suitable for this purpose, in particular also because the method often leads to closed forms.

Denote the log-likelihood corresponding to a particular BRD model by  $\ell(\phi|\omega) = \sum_{i=1}^N \ell_i(\phi|\omega_i)$ , where  $\phi = (\theta, \psi)$  is the  $s$ -dimensional vector, grouping, respectively, the parameters of the measurement and dropout models. Further,  $\omega = (\omega_1, \omega_2, \dots, \omega_N)'$ , belonging to an open subset  $\Omega$  of  $\mathbb{R}^N$ , is a vector defining infinitesimal perturbations around the model studied. Obviously,  $\omega_o = (0, 0, \dots, 0)'$  corresponds to the original model.

Let  $\hat{\phi}$  be the maximum likelihood estimator for  $\phi$ , obtained by maximizing  $\ell(\phi|\omega_o)$ , and let  $\hat{\phi}_\omega$  denote the maximum likelihood estimator for  $\phi$  under  $\ell(\phi|\omega)$ . The relative change in likelihood is a measure for influence; Cook (1986) captured this through the likelihood displacement:  $LD(\omega) = 2[\ell(\hat{\phi}|\omega_o) - \ell(\hat{\phi}_\omega|\omega_o)]$ . A graph of  $LD(\omega)$  versus  $\omega$ , i.e., the geometric surface formed by values of the  $N + 1$  dimensional vector  $\zeta(\omega) = (\omega', LD(\omega))'$ , depicts the influence of perturbations. Because this so-called *influence graph* (Lesaffre and Verbeke, 1998) can only be depicted when  $N = 2$ , Cook (1986) proposed to consider local influence, i.e., the normal curvatures  $C_h$  of  $\zeta(\omega)$  in  $\omega_o$ , in the direction of some  $N$ -dimensional vector  $h$  of unit length. A general expression is

$$C_h = 2 \left| h' \Delta' (\ddot{L})^{-1} \Delta h \right|, \quad (6)$$

with

$$\Delta_i = \frac{\partial^2 \ell_i(\phi|\omega_i)}{\partial \omega_i \partial \phi} \bigg|_{\phi = \hat{\phi}_{\omega_i=0}},$$

$\Delta$  the  $(s \times N)$  matrix with  $\Delta_i$  as its  $i^{\text{th}}$  column, and  $\ddot{L}$  the  $(s \times s)$  matrix of second order derivatives of  $\ell(\phi|\omega_o)$  with respect to  $\phi$ , evaluated at  $\phi = \hat{\phi}$ . A sensible choice for  $h_i$  is the vector with a one in the  $i^{\text{th}}$  position and zero elsewhere, corresponding to the perturbation of the

$i^{\text{th}}$  subject only. Another important direction is the direction  $\mathbf{h}_{\max}$  of maximal normal curvature  $C_{\max}$ . It shows how to perturb the model to obtain the largest local changes in the likelihood displacement. Details can be found in Verbeke and Molenberghs (2000).

The above development is geared towards studying the influence on the likelihood function. Other choices are possible, too. In our contingency-table setting, it is instructive to study influence in predicted cell counts,  $Z_{r_1 r_2, j_1 j_2}$ . In such cases, when  $Z(\phi)$  denotes a particular function of the model parameters, the expression for  $C_{\mathbf{h}}$  can be further generalized as:

$$C_{\mathbf{h}} = 2 \left| \mathbf{h}' \Delta' (\ddot{\mathbf{L}})^{-1} \ddot{\mathbf{Z}} (\ddot{\mathbf{L}})^{-1} \Delta \mathbf{h} \right|, \quad (7)$$

with  $\|\mathbf{h}\| = 1$  and  $\Delta$ ,  $\ddot{\mathbf{L}}$ , and  $\ddot{\mathbf{Z}}$  defined as:

$$\Delta_{ij} = \frac{\partial^2 \ell(\phi | \omega)}{\partial \phi_i \partial \omega_j} \Big|_{\phi = \hat{\phi}, \omega = \omega_o}, \quad \ddot{\mathbf{L}}_{i\ell} = \frac{\partial^2 \ell(\phi | \omega_o)}{\partial \phi_i \partial \phi_\ell} \Big|_{\phi = \hat{\phi}}, \quad \ddot{\mathbf{Z}}_{i\ell} = \frac{\partial^2 Z(\phi)}{\partial \phi_i \partial \phi_\ell} \Big|_{\phi = \hat{\phi}},$$

with  $i, \ell = 1, \dots, p$  and  $j = 1, \dots, q$ .

Here, the local influence measure is not calibrated, because (6) takes the form of a squared second derivative, owing to the double occurrence of  $\Delta$ , 'divided by' another second derivative,  $\ddot{\mathbf{L}}$ . Changing units, therefore, changes scale, unlike in the mixed-models application of Lesaffre and Verbeke (1998) and Verbeke and Molenberghs (2000), where the influence measures approximately sum to twice the sample size. As a consequence, when applying local influence as presented here, interpretation ought to be relative rather than absolute. The important issue remains then as to 'how large is large?' It is extremely hard to provide firm guidelines, but it may be wise, as one possible rule of thumb, to scrutinize the subjects with the largest 5% of influence values. The issue has been studied in detail in Jansen *et al* (2004).

Local influence is useful when assessing which (groups of) observations are most influential in driving the conclusions about the nature of the missing data mechanism in the direction of the more elaborate MNAR model. As Jansen *et al* (2004) indicate, such a phenomenon should not be seen as evidence, let alone proof, that some observations are genuinely influenced by a complex MNAR mechanism rather than, for example, by a simpler MAR mechanism. Indeed, this would conflict with the MAR-counterpart results. Rather, such influence graphs are instructive when assessing which observations have the power to drive the conclusions towards a more complex mechanism. Often, the issue is that other outlying features, such as unusual values, unusual slopes in longitudinal observations, etc. are responsible for the apparent conclusion about

the missing data mechanism. Like the MAR-counterpart result, this type of sensitivity analysis issues a cautionary warning against excessive confidence regarding the nature of the missing data mechanism. A limiting feature, unsurprising in view of the foregoing discussion, is the absence of a ‘yardstick,’ or a threshold, demarcating influential subjects; arguably, influence graphs will blow a whistle over subjects that need further scrutiny.

We first consider perturbations of a given BRD model in the direction of a more elaborate one. For example, BRD4 includes the parameter  $\beta_{.j_2}$ , ( $j_2 = 1, 2$ ), whereas BRD1 only includes  $\beta_{..}$ . For the influence analysis,  $\omega_i$  is not a parameter but an infinitesimal perturbation of the simpler model towards the more complex one, confined to a single subject. For example, for the perturbation of BRD1 in the direction of BRD4, one considers  $\beta_{..}$  and  $\beta_{..} + \omega_i$ . The vector of all  $\omega_i$ ’s defines the direction in which such a perturbation is considered. The BRD log-likelihood is:

$$\ell(\phi|\omega) = \sum_{j_1, j_2} Z_{11, j_1 j_2} \ln \pi_{11, j_1 j_2} + \sum_{j_1} Z_{10, j_1 +} \ln \pi_{10, j_1 +} + \sum_{j_2} Z_{01, + j_2} \ln \pi_{01, + j_2} + Z_{00, ++} \ln \pi_{00, ++}, \quad (8)$$

where  $\pi_{r_1 r_2, j_1 j_2} = p_{j_1 j_2} q_{r_1 r_2 | j_1 j_2}$ , with  $p_{11} = p_1, p_{12} = p_2, p_{21} = p_3, p_{22} = 1 - p_1 - p_2 - p_3$ , and

$$q_{r_1 r_2 | j_1 j_2} = \frac{\exp \{ \alpha_{j_1 j_2} (1 - r_1) + \beta_{j_1 j_2} (1 - r_2) + \gamma (1 - r_1) (1 - r_2) \}}{1 + \exp(\alpha_{j_1 j_2}) + \exp(\beta_{j_1 j_2}) + \exp(\alpha_{j_1 j_2} + \beta_{j_1 j_2} + \gamma)}. \quad (9)$$

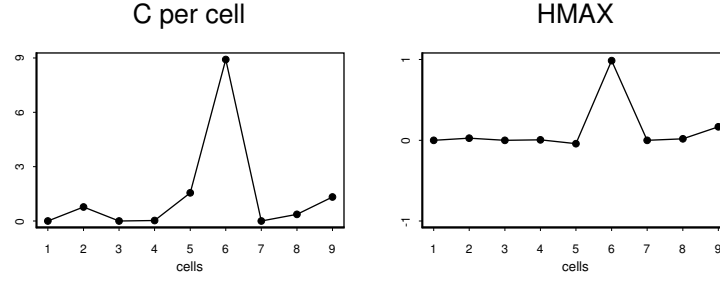
For BRD4, with  $(\alpha_{..}, \beta_{.j_2})$ , expression (9) yields:

$$\begin{aligned} q_{r_1 r_2 | j_1 1} &= \frac{\exp \{ \alpha_{..} (1 - r_1) + \beta_{..} (1 - r_2) + \gamma (1 - r_1) (1 - r_2) \}}{1 + \exp(\alpha_{..}) + \exp(\beta_{..}) + \exp(\alpha_{..} + \beta_{..} + \gamma)}, \\ q_{r_1 r_2 | j_2} &= \frac{\exp \{ \alpha_{..} (1 - r_1) + (\beta_{..} + \omega_i) (1 - r_2) + \gamma (1 - r_1) (1 - r_2) \}}{1 + \exp(\alpha_{..}) + \exp(\beta_{..} + \omega_i) + \exp(\alpha_{..} + \beta_{..} + \omega_i + \gamma)}. \end{aligned}$$

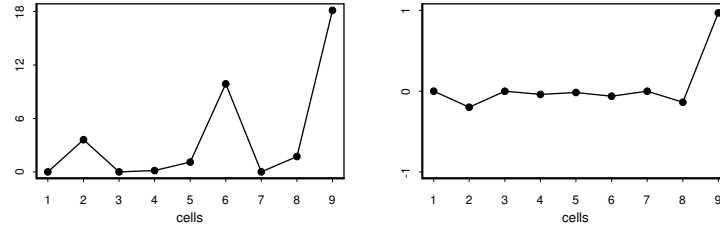
The perturbation  $\omega_i$  defines a difference between the dropout probabilities above, while under the simpler (null) model BRD1, the two expressions reduce to a single dropout probability. Local influence measures now follow from the general logic described above.

For the SPO data, we focus on the model pairs BRD1 vs. BRD4, BRD3 vs. BRD7, and BRD4 vs. BRD7. These pairs are precisely the ones with high influence on the likelihood displacement.

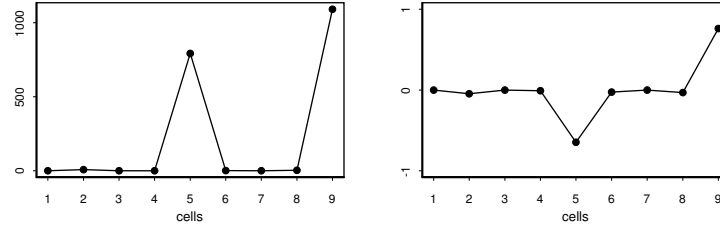
Figure 4 shows the influence measures  $C_i$ , plotted against the  $i^{\text{th}}$  observed cell and  $h_{\max}$  against the  $i^{\text{th}}$  observed cell. For the comparison of BRD1 vs. BRD4, a peak is observed at cell 6, for both  $C_i$  and  $h_{\max}$ , implying that respondents in this cell drive the data more towards BRD4  $(\alpha_{..}, \beta_{.j_2})$  rather than BRD1  $(\alpha_{..}, \beta_{..})$ . That is, subjects with a NO on the attendance question



(a) *BRD1 vs BRD4*



(b) *BRD3 vs. BRD7*



(c) *BRD4 vs. BRD7*

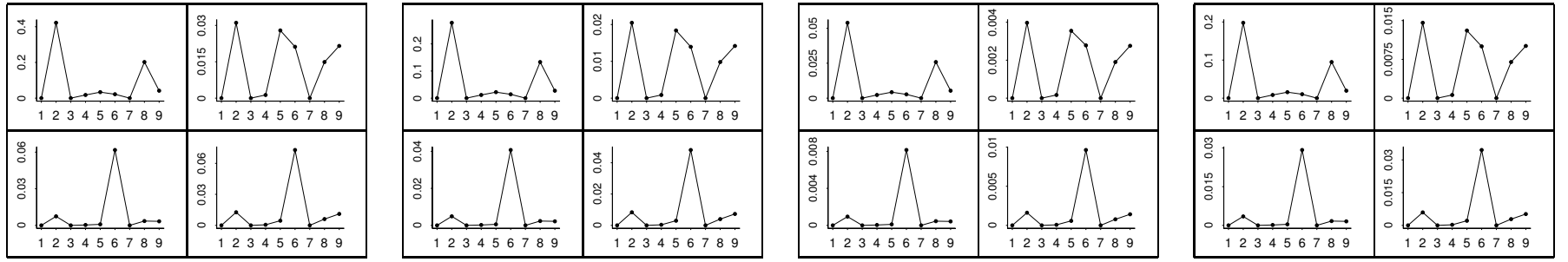
**Figure 4:** *The Slovenian Public Opinion Survey. Local influence analysis on parameters for model pairs (a) BRD1 vs. BRD4, (b) BRD3 vs. BRD7, and (c) BRD4 vs. BRD7. The first column shows the local influence measure  $C_i$  at the  $i^{th}$  observed cell; the second column shows  $h_{max}$  for the  $i^{th}$  observed cell.*

and a missing value on the independence question are influential when perturbing the model such that missingness in the independence question depends on the corresponding unobserved answer (BRD4) rather than being constant (BRD1). For BRD3 vs. BRD7, a peak is observed at ‘fully missing’ cell 9. This means that missingness in the independence question is driven to depend on the corresponding unobserved answer by subjects with missing responses on both questions, and slightly by those with a NO on attendance and a missing value on independence (cell 6). Finally, it is primarily subjects with missing responses on both questions (cell 9) that seem to push the

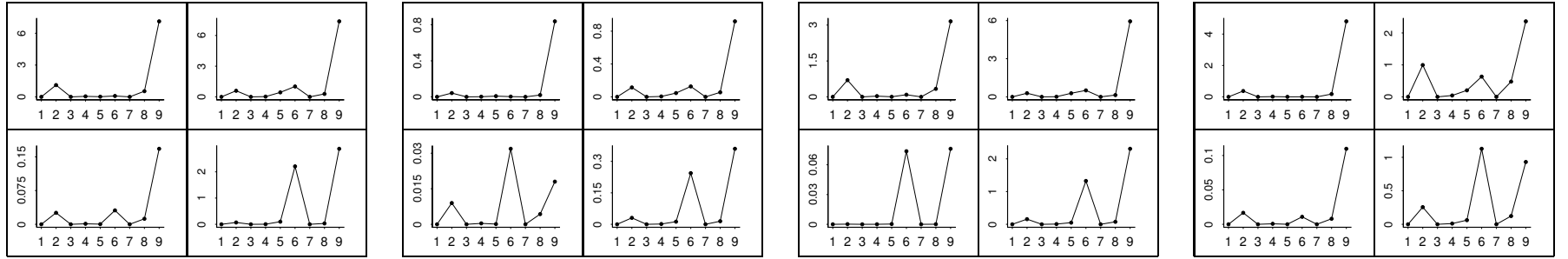
data towards BRD7 ( $\alpha_{.j_2}, \beta_{.j_2}$ ). These subjects, along with those that have a YES on attendance and a missing value on independence (cell 5), make the missingness in the attendance question depend on the response of the independence question.

We now turn to the results of the local influence analysis on the fitted cell counts (Figure 5). The first panel, for model BRD1 vs. BRD4, shows similar shapes for the influence graphs, albeit with differing magnitudes, for a particular cell  $(j_1, j_2)$ , across the four missingness patterns. For  $(j_1, j_2) = (1, 1)$ , it is cell 2 that shows influence, and also slightly cell 8. Respondents with either a YES or a missing value on attendance and a NO on independence thus drive the predicted cell count  $Z_{r_1 r_2, 11}$  towards a model in which the missingness in the independence question depends on its value (BRD4). For  $(j_1, j_2) = (1, 2)$ , cells 2 and 5, as well as 6 and 9, stand out. These respondents make the predicted cell count  $Z_{r_1 r_2, 11}$  seemingly come from BRD4 rather than BRD1. For  $(j_1, j_2) = (2, 1)$  and  $(j_1, j_2) = (2, 2)$ , similar curves are obtained across the four missingness patterns, with a clear peak at cell 6, implying that the “NO-on-attendance/missingness-on-independence” responses perturb predicted cell counts  $Z_{r_1 r_2, 21}$  and  $Z_{r_1 r_2, 22}$  in the direction of a model in which the missingness in the independence question is dependent on its value, rather than on one in which missingness in the independence question is constant.

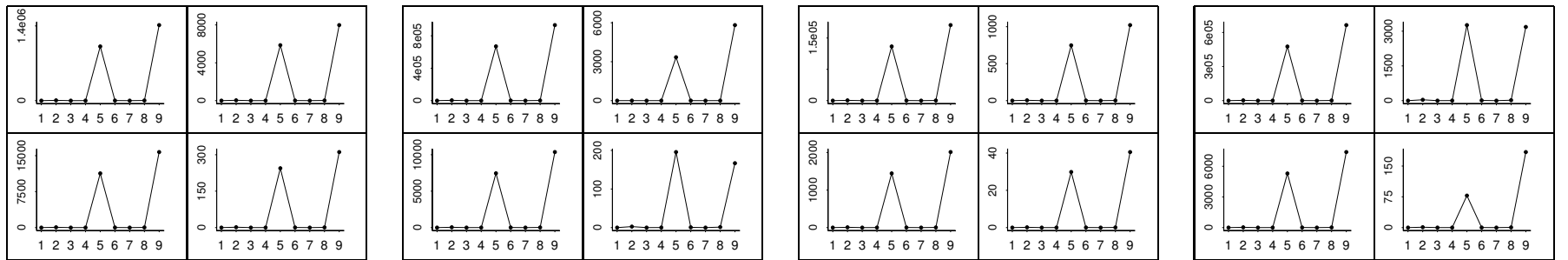
The resulting patterns for the comparison of BRD3 against BRD7 (Figure 5b) differs from what was observed for BRD1 vs. BRD4. Whereas for the latter, influence curves for a particular cell  $(j_1, j_2)$  remained the same across the missingness patterns, for BRD3 vs. BRD7, variations now arise across these missingness patterns, leading to a less clear-cut overall picture. This is further complicated by what can be observed for  $(j_1, j_2) = (2, 1)$  and  $(j_1, j_2) = (2, 2)$ , i.e., the bottom row of the tables, for which curve shapes vary across the missingness patterns. Consider  $(j_1, j_2) = (1, 1)$  and  $(j_1, j_2) = (1, 2)$ . Across missingness patterns, the predicted cell counts  $Z_{r_1 r_2, 11}$  and  $Z_{r_1 r_2, 12}$  are primarily influenced by subjects with both responses missing, and slightly by those having a YES on attendance/NO on independence. For cell  $(j_1, j_2) = (2, 1)$ , similar graphs are obtained for  $(r_1, r_2) = (1, 1)$  and  $(r_1, r_2) = (0, 0)$ , i.e., the completers and double non-responders, respectively, with a peak at cell 9. It is therefore subjects with both responses missing that influence cell counts  $Z_{11, 21}$  and  $Z_{00, 21}$ , in the direction of a model in which missingness in the independence question depends on its value. For the other two missingness patterns,  $(r_1, r_2) = (1, 0)$  and  $(r_1, r_2) = (0, 1)$ , referring to subjects with a single nonresponse, peaks occur at cells 6 and 9. Thus, subjects with a NO on attendance/missingness of independence



(a)  $BRD1$  vs.  $BRD4$



(b)  $BRD3$  vs.  $BRD7$



(c)  $BRD4$  vs.  $BRD7$

**Figure 5:** The Slovenian Public Opinion Survey. Local influence analysis on the predicted cell counts for model pairs (a)  $BRD1$  vs.  $BRD4$ , (b)  $BRD3$  vs.  $BRD7$ , and (c)  $BRD4$  vs.  $BRD7$ . Plots show  $C_i$  values for each of the 16 predicted cell counts (in their respective positions as in Table 1) against the 9 observed cells.



and those with both responses missing have an influence on predicted cell counts  $Z_{10,21}$  and  $Z_{01,21}$ . These same subjects also influence the predicted cell counts  $Z_{r_1 r_2, 22}$ .

Whereas the comparison of BRD3 vs. BRD7 presents the most variable influence graphs, BRD4 vs. BRD7 shows the most consistent ones. All 16 influence curves exhibit a single shape, although of varying magnitudes, implying that influence on any predicted cell count is coming from a common source, regardless of the missingness pattern. Here, we see a clear peak at cells 9 and 5, similar to the likelihood-displacement results. Subjects with missing responses on both questions and those with YES on attendance/missingness on independence, have an influence that drives any predicted cell count towards a model where the missingness in the attendance question depends on the response of the independence question.

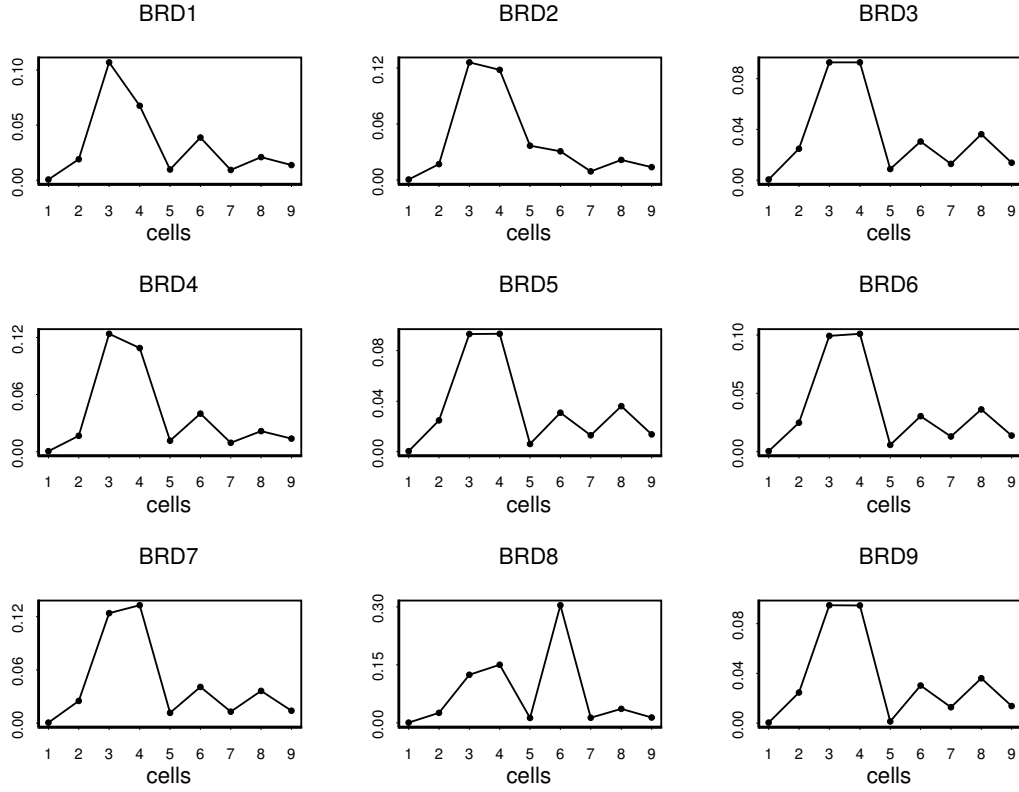
A final local influence analysis introduces infinitesimal perturbations in cell probabilities. The perturbed likelihood then is

$$\begin{aligned} \ell(\phi|\omega) = & \sum_{j_1, j_2} (Z_{11, j_1 j_2} + N\omega_{11, j_1 j_2}) \ln \pi_{11, j_1 j_2} + \sum_{j_1} (Z_{10, j_1 +} + N\omega_{10, j_1 +}) \ln \pi_{10, j_1 +} \\ & + \sum_{j_2} (Z_{01, +j_2} + N\omega_{01, +j_2}) \ln \pi_{01, +j_2} + (Z_{00, ++} + N\omega_{00, ++}) \ln \pi_{00, ++}, \end{aligned}$$

with all notation as before. Under the null model,  $\omega = \omega_o = 0$ , and the above log-likelihood reduces to the standard multinomial one. Because we now focus on cell probabilities rather than observations, the interpretation will be different, in spite of similarities in computation.

For the SPO data, consider first the influence on the likelihood displacement (Figure 6). For most models, the probabilities of cells 3 and/or 4 are influential. Also notable is the influence of changes in cell 6. Thus, the most influential cells for almost all models are the completers answering NO on attendance, likely attributable to the small counts in these cells, while for BRD8, those answering NO on attendance and unobserved response on independence are influential.

Figure 7 summarizes the results. Cells 1 and 2 are generally without influence, owing to large counts. Note that small perturbations in cell 3 seem to affect only the predicted cell counts in the top row ( $j = 1$ , YES on attendance) under BRD5 and/or BRD6, while such changes impact the cell position  $(j_1, j_2) = (1, 2)$  (NO on attendance/YES on independence) under BRD's 1, 2, 3, 6, and/or 9. Perhaps the most striking in Figure 7 is that for perturbations in cell probability 4 (NO/NO respondents) yields influence on *all* 16 predicted cell counts in most of the higher-numbered BRD4–9. Also, the results for perturbations in cell 6, indicate that it is primarily under



**Figure 6:** *The Slovenian Public Opinion Survey. Local influence analysis on the log-likelihood for the 9 BRD models. Plots show  $C_i$  values against the 9 observed cells for each BRD model.*

BRD8 where a large influence is observed in most of the predicted cell counts. Finally, changes in the probability of the doubly missing cell 9 affect only the predicted cell counts of this missingness pattern and only under BRD's 1, 2, and/or 3.

## 5 How Sensitive is the Proportion of 'Yes' Voters?

The inferential target of the Slovenian Public Opinion Survey analyses is estimating the proportion of people voting in favour of independence, a goal hampered by incompleteness. It has been argued that putting blind belief in a single model may be too strong; it is not possible, from a purely statistical point of view, to unambiguously validate a single model, motivating the consideration of sensitivity analyses. We contemplated a variety of these, going beyond conventional sensitivity analysis applications, which are often confined to a single sensitivity assessment tool. Table 3 summarizes various methods. The simplest analysis considers the non-parametric bounds and deduces that even the least supportive scenario for independence would still produce a, roughly,

Adding omega to cell number:	Z11	Z10	Z01	Z00																				
<table><tr><td>1</td><td></td></tr><tr><td></td><td></td></tr></table>	1				<table><tr><td>---</td><td>---</td></tr><tr><td>---</td><td>---</td></tr></table>	---	---	---	---	<table><tr><td>---</td><td>---</td></tr><tr><td>---</td><td>---</td></tr></table>	---	---	---	---	<table><tr><td>---</td><td>---</td></tr><tr><td>---</td><td>---</td></tr></table>	---	---	---	---	<table><tr><td>---</td><td>---</td></tr><tr><td>---</td><td>---</td></tr></table>	---	---	---	---
1																								
---	---																							
---	---																							
---	---																							
---	---																							
---	---																							
---	---																							
---	---																							
---	---																							
<table><tr><td></td><td>2</td></tr><tr><td></td><td></td></tr></table>		2			<table><tr><td>---</td><td>---</td></tr><tr><td>---</td><td>---</td></tr></table>	---	---	---	---	<table><tr><td>---</td><td>---</td></tr><tr><td>---</td><td>---</td></tr></table>	---	---	---	---	<table><tr><td>---</td><td>---</td></tr><tr><td>---</td><td>---</td></tr></table>	---	---	---	---	<table><tr><td>---</td><td>3</td></tr><tr><td>---</td><td>---</td></tr></table>	---	3	---	---
	2																							
---	---																							
---	---																							
---	---																							
---	---																							
---	---																							
---	---																							
---	3																							
---	---																							
<table><tr><td></td><td></td></tr><tr><td>3</td><td></td></tr></table>			3		<table><tr><td>5,6</td><td>5,6</td></tr><tr><td>1,2,3,9</td><td>---</td></tr></table>	5,6	5,6	1,2,3,9	---	<table><tr><td>5,6</td><td>5,6</td></tr><tr><td>2,3,9</td><td>---</td></tr></table>	5,6	5,6	2,3,9	---	<table><tr><td>5,6</td><td>5,6</td></tr><tr><td>1,2,3,6</td><td>---</td></tr></table>	5,6	5,6	1,2,3,6	---	<table><tr><td>5</td><td>5</td></tr><tr><td>1,2,3,6</td><td>6</td></tr></table>	5	5	1,2,3,6	6
3																								
5,6	5,6																							
1,2,3,9	---																							
5,6	5,6																							
2,3,9	---																							
5,6	5,6																							
1,2,3,6	---																							
5	5																							
1,2,3,6	6																							
<table><tr><td></td><td></td></tr><tr><td></td><td>4</td></tr></table>				4	<table><tr><td>4,6,7,9</td><td>4,5,6,7</td></tr><tr><td>4,5,6,7,9</td><td>2,3,4,5,6,7,8,9</td></tr></table>	4,6,7,9	4,5,6,7	4,5,6,7,9	2,3,4,5,6,7,8,9	<table><tr><td>4,5,6,7,9</td><td>4,5,6,7,8</td></tr><tr><td>4,5,6,7</td><td>2,3,4,5,6,7,8,9</td></tr></table>	4,5,6,7,9	4,5,6,7,8	4,5,6,7	2,3,4,5,6,7,8,9	<table><tr><td>4,5,6,7</td><td>4,5,6,7,8</td></tr><tr><td>4,5,6,7,9</td><td>2,3,4,5,6,7,8</td></tr></table>	4,5,6,7	4,5,6,7,8	4,5,6,7,9	2,3,4,5,6,7,8	<table><tr><td>4,5,6,7,9</td><td>3,4,5,6,7,8,9</td></tr><tr><td>4,5,6,7,9</td><td>2,3,4,5,7,9</td></tr></table>	4,5,6,7,9	3,4,5,6,7,8,9	4,5,6,7,9	2,3,4,5,7,9
	4																							
4,6,7,9	4,5,6,7																							
4,5,6,7,9	2,3,4,5,6,7,8,9																							
4,5,6,7,9	4,5,6,7,8																							
4,5,6,7	2,3,4,5,6,7,8,9																							
4,5,6,7	4,5,6,7,8																							
4,5,6,7,9	2,3,4,5,6,7,8																							
4,5,6,7,9	3,4,5,6,7,8,9																							
4,5,6,7,9	2,3,4,5,7,9																							
<table><tr><td>5</td><td></td></tr><tr><td></td><td></td></tr></table>	5				<table><tr><td>2</td><td>---</td></tr><tr><td>---</td><td>---</td></tr></table>	2	---	---	---	<table><tr><td>1,2,3</td><td>1,2</td></tr><tr><td>---</td><td>---</td></tr></table>	1,2,3	1,2	---	---	<table><tr><td>---</td><td>---</td></tr><tr><td>---</td><td>---</td></tr></table>	---	---	---	---	<table><tr><td>---</td><td>---</td></tr><tr><td>---</td><td>---</td></tr></table>	---	---	---	---
5																								
2	---																							
---	---																							
1,2,3	1,2																							
---	---																							
---	---																							
---	---																							
---	---																							
---	---																							
<table><tr><td></td><td></td></tr><tr><td>6</td><td></td></tr></table>			6		<table><tr><td>8</td><td>8</td></tr><tr><td>8</td><td>1,8</td></tr></table>	8	8	8	1,8	<table><tr><td>8</td><td>8</td></tr><tr><td>1,8</td><td>1</td></tr></table>	8	8	1,8	1	<table><tr><td>8</td><td>---</td></tr><tr><td>8</td><td>---</td></tr></table>	8	---	8	---	<table><tr><td>8</td><td>---</td></tr><tr><td>8</td><td>8</td></tr></table>	8	---	8	8
6																								
8	8																							
8	1,8																							
8	8																							
1,8	1																							
8	---																							
8	---																							
8	---																							
8	8																							
<table><tr><td>7</td><td></td></tr><tr><td></td><td></td></tr></table>	7				<table><tr><td>3,9</td><td>---</td></tr><tr><td>---</td><td>---</td></tr></table>	3,9	---	---	---	<table><tr><td>9</td><td>---</td></tr><tr><td>---</td><td>---</td></tr></table>	9	---	---	---	<table><tr><td>1,2,3,9</td><td>---</td></tr><tr><td>---</td><td>---</td></tr></table>	1,2,3,9	---	---	---	<table><tr><td>---</td><td>---</td></tr><tr><td>---</td><td>---</td></tr></table>	---	---	---	---
7																								
3,9	---																							
---	---																							
9	---																							
---	---																							
1,2,3,9	---																							
---	---																							
---	---																							
---	---																							
<table><tr><td></td><td>8</td></tr><tr><td></td><td></td></tr></table>		8			<table><tr><td>---</td><td>1,2,3,9</td></tr><tr><td>---</td><td>---</td></tr></table>	---	1,2,3,9	---	---	<table><tr><td>9</td><td>3,9</td></tr><tr><td>---</td><td>---</td></tr></table>	9	3,9	---	---	<table><tr><td>---</td><td>1,2,3,9</td></tr><tr><td>---</td><td>1,9</td></tr></table>	---	1,2,3,9	---	1,9	<table><tr><td>9</td><td>---</td></tr><tr><td>---</td><td>---</td></tr></table>	9	---	---	---
	8																							
---	1,2,3,9																							
---	---																							
9	3,9																							
---	---																							
---	1,2,3,9																							
---	1,9																							
9	---																							
---	---																							
<table><tr><td>9</td><td></td></tr><tr><td></td><td></td></tr></table>	9				<table><tr><td>1,2</td><td>---</td></tr><tr><td>---</td><td>---</td></tr></table>	1,2	---	---	---	<table><tr><td>---</td><td>---</td></tr><tr><td>---</td><td>---</td></tr></table>	---	---	---	---	<table><tr><td>---</td><td>---</td></tr><tr><td>---</td><td>---</td></tr></table>	---	---	---	---	<table><tr><td>1,2,3</td><td>1,2</td></tr><tr><td>1</td><td>1</td></tr></table>	1,2,3	1,2	1	1
9																								
1,2	---																							
---	---																							
---	---																							
---	---																							
---	---																							
---	---																							
1,2,3	1,2																							
1	1																							

**Figure 7:** *The Slovenian Public Opinion Survey. Local influence analysis on cell counts when perturbing each of the 9 observed cell probabilities. Entries in boxes denote the BRD model number for which influence is largest when the particular cell probability is perturbed.*

70% majority. Alternatively, one can fit a discrete class of models, such as the nine BRD models, or merely the well-fitting models BRD6–9, and then construct the resulting interval; this narrows the non-parametric interval and even excludes the plebiscite value. It is here that the MAR counterparts come in quite useful, given that they essentially produce a single point estimate of 89.2%, very close to the true value. The concept of interval naturally leads to considering of intervals of ignorance, nicely interpolating between the non-parametric interval and a single,

identified model. Recall that Model 12 reproduces the non-parametric interval, while Models 10 and 11 further narrow it.

Turning to substantive considerations, the estimates for  $\theta$  (Table 4), can be split into two:

**Optimistic:** MAR bodyguards, BRD1, BRD3, (BRD5), (BRD6), and BRD9.

**Pessimistic:** BRD4, BRD7, and BRD8.

The parenthetical ones are slightly less pronounced than the others. The assumptions regarding the missingness mechanism, underpinning all twelve models, can be read from the second column in Table 4; they are also spelled out in Table 7. It is striking that the three fully identified, pessimistic estimates allow missingness in the independence question to depend on independence, whereas the other six do not: these three models support the thesis that there is a large group of people, in favor of independence, that would not partake in the plebiscite, as can be seen from Table 6. We will study these in more detail, using BRD7, BRD9, and their MAR counterparts. The pessimistic fit of BRD7 asserts that the proportion of people in favour of independence, yet that would not attend the plebiscite, amounts to  $(78 + 155.8 + 44.8 + 112.5)/2074$ , i.e., 18.9%. The optimistic BRD9 predicts this fraction to be as low as 7.3%, while for the counterparts it goes down further to 6.6%. We infer with reasonable confidence that, in the actual plebiscite, people expressed their opinion, regardless of real or perceived pressure, and that an overwhelming majority favoured independence, supporting the optimistic scenario.

There are more nuances in the overspecified Models 10–12. While Models 10 and 12 also include the “pessimistic relationship” of missingness in independence on the independence answer, the interval is not pessimistic. Rather, the three intervals encompass both very pessimistic and very optimistic scenarios. This is because, unlike the fully identified models, there is no need to sacrifice one type of dependence to maintain identifiability. Precisely, while none of BRD4, BRD7, and BRD8 allow missingness in the independence question to depend on the attendance response, the three IIs do allow for such a dependence, a feature shared with BRD2, BRD6, and BRD9. It ought not to go unnoticed that, among the optimistic ones, BRD5 and BRD6 are somewhat less pronounced; these models allow missingness in the attendance question to depend on the respondent’s attendance position. The effect of this is similar, but less sharp, than in the ‘independence on independence’ scenario sketched above. It is therefore no longer a surprise that BRD8, where these two effects play together, produces the most pessimistic estimate.

The question remains why the impact of changing the assumptions is rather spectacular, in the

**Table 7:** Slovenian Public Opinion Survey. Meaning of the missingness mechanism in the nine identified Models BRD1–BRD9 and its three overspecified extensions Models 10–12.

Missingness in Depends on	→	attendance		independence	
		attendance	independence	attendance	independence
BRD1		—	—	—	—
BRD2		—	—	✓	—
BRD3		—	✓	—	—
BRD4		—	—	—	✓
BRD5		✓	—	—	—
BRD6		✓	—	✓	—
BRD7		—	✓	—	✓
BRD8		✓	—	—	✓
BRD9		—	✓	✓	—
Model 10		—	✓	✓	✓
Model 11		✓	✓	✓	—
Model 12		✓	✓	✓	✓

sense that the most pessimistic scenarios are very close to the pessimistic bound, whereas the most optimistic scenarios are virtually at the optimistic bound. It is instructive to return to the global and local influence analyses. First and foremost, we learn that there is relatively little influence all together, *except in BRD4, BRD7, and BRD8*, i.e., the entire pessimistic group. This is clear from the discussion in Sections 4.3 and 4.4, and from Figures 3, 4, and 6. From global influence we learn that there is strong impact of the (relatively small) no-on-independence cells, in the pessimistic models: changes in these counts can dramatically alter the way in which the incomplete cells are split over the hypothetical complete cells, rendering  $\hat{\theta}$  unstable. Local influence focuses on a different aspect: which observations/cells drive the conclusions away from a given null model? For BRD1 versus BRD4, this is Cell 6, the not-in-favour-of-independence respondents without declared attendance status. Thus, a small but influential count is simultaneously responsible for a move towards a pessimistic scenario and the extent of pessimism.

Sensitivity analyses combined and substantive considerations demonstrate that the optimistic scenarios, whether from the MAR counterparts or the optimistic group of BRD models, are plausible descriptions of the mechanisms operating during the plebiscite exercise. The influence analyses show that the pessimistic ones are rather different from their optimistic counterparts and constitute plausible scenarios only owing to the presence of one or a few influential cells.

## 6 Concluding Remarks

We have presented a variety of existing and expanded sensitivity analyses, to gain insight in the impact of the missing data generating mechanism governing the Slovenian Public Opinion Survey data. A first family of sensitivity analyses is based on considering a collection of models, (1) through non-parametric bounds, (2) a family of identified models, or (3) intervals resulting from over-specified models. Next to these, we conducted observation-varying analyses: (4) global influence, (5) local influence in terms of infinitesimally varying particular parameters, and (6) local influence through perturbing cell counts. We combined conclusions of these sensitivity analyses with substantive considerations, arriving at a coherent picture that points at increased confidence in the more optimistic scenarios, and provides a plausible explanation for these.

How to proceed in similar, related, or different data-analysis problems with incomplete data? First, many of the methods have been developed for contingency tables and beyond. This is true for local and global influence technology that also exists for continuous outcomes, possibly with covariates, and intervals of ignorance that have been developed for logistic regression. For a review, see Molenberghs and Kenward (2007). Other methods, not considered for the purpose of our analysis, can be considered as well. We briefly touch upon some.

First, pattern-mixture models (Thijs *et al*, 2000) are worthy of attention, either for their own sake because they might appropriate to address a particular scientific question, or as a useful contrast to selection models either: (1) to answer the same scientific question, based on these different modeling strategies; or (2) to gain additional insight by supplementing the selection model results with those from a pattern-mixture approach. Pattern-mixture models also have a special role in some multiple-imputation-based sensitivity analyses (Molenberghs and Kenward, 2007). Second, one could turn towards the shared-parameter framework, one of its main advantages being that it can easily handle non-monotone missingness. Nevertheless, these models are based on very strong parametric assumptions, such as normality for the shared random effect(s). Sensitivities abound in the selection and pattern-mixture frameworks as well, but the assumption of unobserved, random, or latent effects, further complicate the issue (Tsonaka, Verbeke, and Lesaffre, 2007; Rizopoulos, Verbeke, and Molenberghs, 2007). Third, Beunckens *et al* (2007) proposed a so-called *latent-class mixture model*, bringing together features of the selection, pattern-mixture, and shared-parameter frameworks. Fourth, one can take a semi-parametric standpoint, where weighted generalized estimating equations (WGEE), proposed by Robins, Rotnitzky, and Zhao

(1994) play a central role. Rather than jointly modelling the outcome and missingness processes, the centerpiece is inverse probability weighting of a subject's contribution, where the weights are specified in terms of factors influencing missingness, such as covariates and observed outcomes (Robins, Rotnitzky, and Scharfstein, 1998; Scharfstein *et al*, 1999). Robins, Rotnitzky, and Scharfstein (2000) and Robins *et al* (2001) use this framework to conduct sensitivity analysis.

The field of sensitivity analysis towards the impact of incomplete data is vibrantly active and methodology as well as insight surrounding it is bound to emerge in time to come. The modeler and practitioner have got an ever expanding toolkit of sensitivity analysis machinery at their disposal. Carefully selected sensitivity analysis equipment, supplemented with substantive arguments, can produce valuable insight and perhaps even confidence, above and beyond what is obtainable from a single analysis, be it of the MAR or MNAR type.

## Acknowledgment

The authors gratefully acknowledge financial support from the Interuniversity Attraction Pole Research Network P6/03 of the Belgian Government (Belgian Science Policy).

## References

- Baker, S.G., Rosenberger, W.F. and DerSimonian, R. (1992). Closed-form estimates for missing counts in two-way contingency tables. *Statistics in Medicine* **11**, 643-657.
- Barnett, V. (2002). *Sample Survey: Principles and Methods* (3rd ed.). London: Arnold.
- Beunckens, C., Molenberghs, G., Verbeke, G., and Mallinckrodt, C. (2007). A latent-class mixture model for incomplete longitudinal Gaussian data. *Biometrics*, **00**, 000–000.
- Cook, R.D. (1979). Influential observations in linear regression. *Journal of the American Statistical Association* **74**, 169–174.
- Cook, R.D. (1986). Assessment of local influence. *Journal of the Royal Statistical Society Series B* **2**, 133–169.
- Cook, R.D. and Weisberg, S. (1982). *Residuals and Influence in Regression*. London: Chapman and Hall.

- Diggle, P.J. and Kenward, M.G. (1994). Informative drop-out in longitudinal data analysis (with discussion). *Applied Statistics* **43**, 49-93.
- Jansen, I., Hens, N., Molenberghs, G., Aerts, M., Verbeke, G., and Kenward, M.G. (2006). The nature of sensitivity in missing not at random models. *Computational Statistics and Data Analysis*, **50**, 830-858.
- Jansen, I., Molenberghs, G., Aerts, M., Thijs, H. and Van Steen, K. (2003). A local influence approach applied to binary data from a psychiatric study. *Biometrics* **59**, 410-419.
- Kenward, M.G. (1998). Selection models for repeated measurements with nonrandom dropout: An illustration of sensitivity analysis. *Statistics in Medicine* **17**, 2723-2732.
- Kenward, M.G., Goetghebeur, E. and Molenberghs, G. (2001). Sensitivity analysis of incomplete categorical data. *Statistical Modelling* **1**, 31-48.
- Lesaffre, E. and Verbeke, G. (1998). Local influence in linear mixed models. *Biometrics* **54**, 570-582.
- Molenberghs, G., Beunckens, C., Sotto, C. and Kenward, M.G. (2007). Every missing not at random model has got a missing at random counterpart with equal fit. *Journal of the Royal Statistical Society Series B*. In press.
- Molenberghs, G. and Kenward, M.G. (2007). *Handling Incomplete Data From Clinical Studies*. New York: Wiley.
- Molenberghs, G., Kenward, M.G. and Goetghebeur, E. (2001). Sensitivity analysis for incomplete contingency tables: the Slovenian plebiscite case. *Applied Statistics* **50**, 15-29.
- Molenberghs, G., Kenward, M.G. and Lesaffre, E. (1997). The analysis of longitudinal ordinal data with non-random dropout. *Biometrika* **84**, 33-44.
- Molenberghs, G., Michiels, B., Kenward, M.G., and Diggle, P.J. (1998). Missing data mechanisms and pattern-mixture models. *Statistica Neerlandica* **52**, 153-161.
- Molenberghs, G., Thijs, H., Kenward, M.G. and Verbeke, G. (2003). Sensitivity analysis of continuous incomplete longitudinal outcomes. *Statistica Neerlandica*, **57**, 112-135.



- Rizopoulos D., Verbeke G., and Molenberghs G. (2007). Shared parameter models under random-effects misspecification. *Biometrika*, **94**, 000–000.
- Robins, J. M., Rotnitzky, A., and Scharfstein, D.O. (1998). Semiparametric regression for repeated outcomes with non-ignorable non-response. *Journal of the American Statistical Association*, **93**, 1321–1339.
- Robins, J. M., Rotnitzky, A., and Scharfstein, D.O. (2000). Sensitivity analysis for selection bias and unmeasured confounding in missing data and causal inference models. In *Statistical Models in Epidemiology, the Environment, and Clinical Trials*, M. E. Halloran and D. A. Berry (eds), 1–94. New York: Springer.
- Robins, J. M., Rotnitzky, A., and Zhao, L. P. (1994). Estimation of regression coefficients when some regressors are not always observed. *Journal of the American Statistical Association*, **89**, 846–866.
- Rotnitzky, A., Scharfstein, D., Su, T.L., and Robins, J.M. (2001). Methods for conducting sensitivity analysis of trials with potentially nonignorable competing causes of censoring. *Biometrics*, **57**, 103–113.
- Rubin, D.B., Stern, H.S. and Vehovar, V. (1995). Handling “Don’t Know” survey responses: the case of the Slovenian plebiscite. *Journal of the American Statistical Association* **90**, 822–828.
- Rubin, D.B., 1976. Inference and missing data. *Biometrika* **63**, 581–592.
- Rubin, D.B. (1994). Discussion to Diggle, P.J. and Kenward, M.G.: Informative dropout in longitudinal data analysis. *Applied Statistics*, **43**, 80–82.
- Scharfstein, D.O., Rotnitzky, A., and Robins, J.M. (1999). Adjusting for nonignorable drop-out using semiparametric nonresponse models (with discussion). *Journal of the American Statistical Association*, **94**, 1096–1146.
- Thijs, H., Molenberghs, G., Michiels, B., Verbeke, G., and Curran, D. (2002). Strategies to fit pattern-mixture models. *Biostatistics*, **3**, 245–265.
- Thijs, H., Molenberghs, G. and Verbeke, G. (2000). The milk protein trial: influence analysis of the dropout process. *Biometrical Journal*, **42**, 617–646.

- Tsiatis, A.A. (1975). A nonidentifiability aspect of the problem of competing risks. *Proceedings of the National Academy of Science*, **72**, 20–22.
- Tsonaka R., Verbeke G., and Lesaffre E. (2007). A semi-parametric shared parameter model to handle non-monotone non-ignorable missingness. Submitted for publication.
- Vansteelandt, S., Goetghebeur, E., Kenward, M.G., and Molenberghs, G. (2006). Ignorance and uncertainty regions as inferential tools in a sensitivity analysis. *Statistica Sinica* **16**, 953–979.
- Van Steen, K., Molenberghs, G., Verbeke, G. and Thijs, H. (2001). A local influence approach to sensitivity analysis of incomplete longitudinal ordinal data. *Statistical Modelling: An International Journal* **1**, 125–142.
- Verbeke, G. and Molenberghs, G. (2000). *Linear Mixed Models for Longitudinal Data*. New York: Springer.
- Verbeke, G., Molenberghs, G., Thijs, H., Lesaffre, E. and Kenward, M.G. (2001). Sensitivity analysis for non-random dropout: a local influence approach. *Biometrics* **57**, 7–14.