TECHNICAL
REPORT

06109

AN INTEGRATED SENSITIVITY ANALYSIS OF THE SLOVENIAN PUBLIC OPINION SURVEY DATA

BEUNCKENS, C., SOTTO, C., MOLENBERGHS, G. and G. VERBEKE

An Integrated Sensitivity Analysis of the Slovenian Public Opinion Survey Data

Caroline Beunckens[1], **Cristina Sotto**[1, 2], **Geert Molenberghs**[1], **and Geert Verbeke**[3]

[1] Center for Statistics, Hasselt University,
Agoralaan, Building D, B3590 Diepenbeek, Belgium.

[2] School of Statistics, University of the Philippines,
Diliman, Quezon City, Philippines.

[3] Biostatistical Centre, Katholieke Universiteit Leuven, Belgium.

Abstract

Many models to analyze incomplete data have been developed, which allow the missingness to be non-random. Since such models rely on unverifiable modelling assumptions, research nowadays is devoted to assess the sensitivity of resulting inferences. A popular sensitivity tool is based on local influence (Cook, 1986), a technique that studies the effect of small perturbations around a given null model to detect subjects which have an unduly influence on the analysis. Jansen *et al* (2003) developed a local-influence approach for binary data, subject to non-monotone missingness, focusing on the model proposed by Baker, Rosenberger and DerSimonian (1992). These authors focus on perturbations of a given BRD model in the direction of an alternative model with one additional parameter. Additionally, in this paper, we will consider perturbations in the observed cell probabilities, rather than the parameters of the model. Further, we show that not only the model parameters, but also the cell counts can have an important influence on the conclusions. To this end, we derived influence measures for functions of the model parameters in general, following the reasoning of Cook (1986). Both local influence approaches are applied to the Slovenian Public Opinion Survey data, and juxtaposed with the interval-of-ignorance based sensitivity analysis of Molenberghs, Kenward and Goetghebeur (2001). 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 rather complete sensitivity analysis picture.

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

Table 1: *Results of the Slovenian Public Opinion Survey. The* Don't Know *category is indicated by* $*$.

		Independence		
Secession	Attendance	Yes	No	$*$
Yes	Yes	1191	8	21
	No	8	0	4
	$*$	107	3	9
No	Yes	158	68	29
	No	7	14	3
	$*$	18	43	31
$*$	Yes	90	2	109
	No	1	2	25
	$*$	19	8	96

1 Introduction

In 1991 Slovenians voted for independence from former Yugoslavia in a plebiscite. To prepare for this result, the Slovenian government collected data in the so-called Slovenian Public Opinion Survey (SPO), a month prior to the plebiscite. Rubin, Stern and Verhovar (1995) studied the three fundamental questions added to the SPO and, in comparing it to the plebiscite's outcome, drew conclusions about the missing data process.

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 since independence would have been possible in confederal form as well and therefore the secession question is added. Question (3) is highly relevant since the political decision was taken that not attending was treated as an effective NO to question (1). Thus, the primary estimand is the proportion θ of people that will be considered as voting YES, which is the fraction of people answering yes to both the attendance and independence question. The raw data are presented in Table 1.

Clearly, the data are incomplete, hampering the straightforward estimation of θ. Missingness indeed adds a source of uncertainty and it is useful to distinguish between two types of *statistical uncertainty*. The first one, *statistical imprecision*, is due to finite sampling. The Slovenian Public

Table 2: *Theoretical distribution of the probability mass over complete and observed cells, respectively, for a bivariate binary outcome with non-monotone missingness.*

(a) Complete cells

$\pi_{11,11}$	$\pi_{11,12}$
$\pi_{11,21}$	$\pi_{11,22}$

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

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

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

(b) Observed cells

$\pi_{11,11}$	$\pi_{11,12}$
$\pi_{11,21}$	$\pi_{11,22}$

$\pi_{10,1+}$
$\pi_{10,2+}$

$\pi_{01,+1}$	$\pi_{01,+2}$

$\pi_{00,++}$

Table 3: *Observed cells for the Slovenian Public Opinion Survey, collapsed over the secession question. A simplified cell indexing system has been used.*

$m_1 : 1439$	$m_2 : 78$
$m_3 : 16$	$m_4 : 16$

$m_5 : 159$
$m_6 : 32$

$m_7 : 144$	$m_8 : 54$

$m_9 : 136$

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 complete and observed data. Let us focus on two binary questions, such as the independence and attendance questions in the Slovenian Public Opinion Survey. The 16 theoretical complete cell probabilities are as in Table 2, thus producing 15 complete data degrees of freedom. The generic expression for the cell probabilities is $\pi_{r_1 r_2, j_1 j_2}$, where $r_\ell = 0$ (1) if the answer to question ℓ is missing (observed) and $j_\ell = 1$ (2) if the answer to question ℓ is yes (no).

Similarly, the 9 observed cells can be represented as in Table 2. The latter table is directly comparable to the observed data structure. In the SPO case, for example, these 9 counts are obtained from collapsing Table 1 over the secession question, hence producing Table 3.

In the ensuing discussion, particular cells in Table 3 shall be referred to in terms of cell number references. The labels used are illustrated in Table 3. Rather than probabilities, the data refer to cell counts. We write m_ℓ for the cell counts ($\ell = 1, \ldots, 9$), adopting a simplified indexing system.

In what follows, we will refer to the mechanism governing missingness using the taxonomy introduced by 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). Further concepts will be introduced in the rest of the paper as the need arises.

In Section 2, a tour is made of a number of simple analyses. The family of 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 bodyguard 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.

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* (2006) 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

Table 4: *The Slovenian Public Opinion Survey. Some estimates of the proportion θ attending the plebiscite and voting for independence, as presented in Rubin, Stern and Verhovar (1995) and Molenberghs, Kenward and Goetghebeur (2001).*

Estimation method	Voting in favour of independence: $\widehat{\theta}$
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
Plebiscite	0.885

contingency tables, subject to non-monotone missingness. Rubin, Stern and Verhovar (1995) conducted several analyses of the data. Their main emphasis was on determining the proportion θ of the population that would attend the plebiscite and vote for independence. Their estimates are reproduced in Table 4.

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 θ 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, underscoring the growing conviction that they should routinely be disregarded and a move towards, at least, MAR is in place (Molenberghs and Kenward, 2007). Rubin, Stern and Verhovar (1995) considered two MAR models, also reported in Table 4, 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 Verhovar (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 to be 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: $\nu_{10,j_1j_2} = \nu_{11,j_1j_2}\beta_{j_1j_2}$, $\nu_{01,j_1j_2} = \nu_{11,j_1j_2}\alpha_{j_1j_2}$, and $\nu_{00,j_1j_2} = \nu_{11,j_1j_2}\alpha_{j_1j_2}\beta_{j_1j_2}\gamma$, with

$$\alpha_{j_1j_2} = \frac{\phi_{01|j_1j_2}}{\phi_{11|j_1j_2}}, \qquad \beta_{j_1j_2} = \frac{\phi_{10|j_1j_2}}{\phi_{11|j_1j_2}}, \qquad \gamma = \frac{\phi_{11|j_1j_2}\phi_{00|j_1j_2}}{\phi_{10|j_1j_2}\phi_{01|j_1j_2}}.$$

Furthermore, $\nu_{r_1r_2,j_1j_2}$ is the model for the four cells, indexed by j_1 and j_2, in pattern (r_1, r_2), where $(r_1, r_2) = (1, 1)$ corresponds to completers, etc.

The α (β) parameters describe missingness in the independence (attendance) question, and γ captures the interaction between both. The subscripts are missing from γ since Baker, Rosenberger and DerSimonian (1992) have shown that this quantity is independent of j and k in every identifiable model. These authors considered nine models, based on setting $\alpha_{j_1j_2}$ and $\beta_{j_1j_2}$ constant in one or more indices, and enumerated using the 'BRD' abbreviation:

$$\begin{array}{llllll}
\text{BRD1}: & (\alpha_{..}, \beta_{..}) & \text{BRD4}: & (\alpha_{..}, \beta_{.j_2}) & \text{BRD7}: & (\alpha_{.j_2}, \beta_{.j_2}) \\
\text{BRD2}: & (\alpha_{..}, \beta_{j_1.}) & \text{BRD5}: & (\alpha_{j_1.}, \beta_{..}) & \text{BRD8}: & (\alpha_{j_1.}, \beta_{.j_2}) \\
\text{BRD3}: & (\alpha_{.j_2}, \beta_{..}) & \text{BRD6}: & (\alpha_{j_1.}, \beta_{j_1.}) & \text{BRD9}: & (\alpha_{.j_2}, \beta_{j_1.}).
\end{array}$$

Interpretation is straightforward, for example, BRD1 is MCAR, and in BRD4 missingness in the first variable is constant, while missingness in the second variable depends on its value. BRD6–BRD9 saturate the observed data degrees of freedom, while the lower numbered ones leave room for a non-trivial model fit to the observed data.

Molenberghs, Kenward and Goetghebeur (2001) and Molenberghs *et al* (2006) fitted the BRD models and Table 5 summarizes the results. BRD1 produces $\widehat{\theta} = 0.892$, exactly the same as the first MAR estimate obtained by Rubin, Stern and Verhovar (1995). This does not come as a surprise, since both models assume MAR and use information from the two main questions. A graphical representation of the original analyses and the BRD models combined is given in Figure 2.

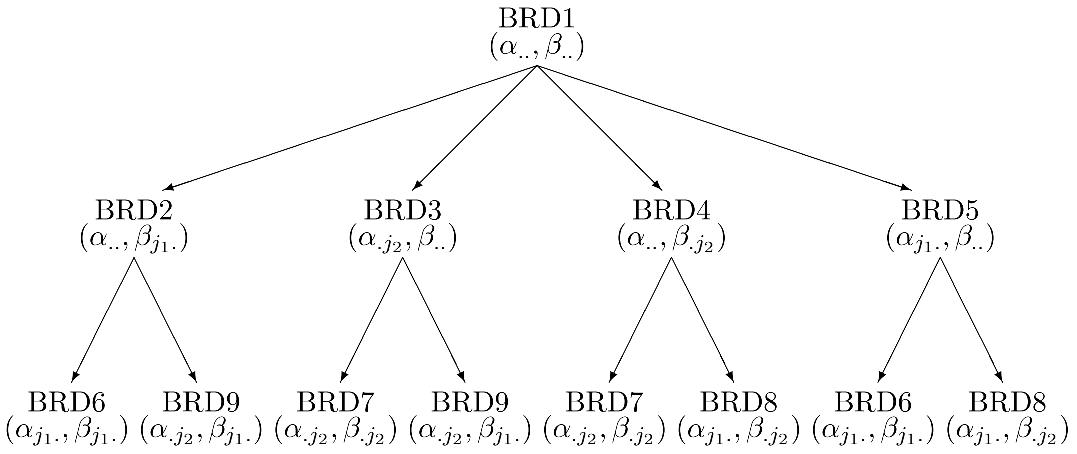


Figure 1: *Graphical representation of the BRD model nesting structure.*

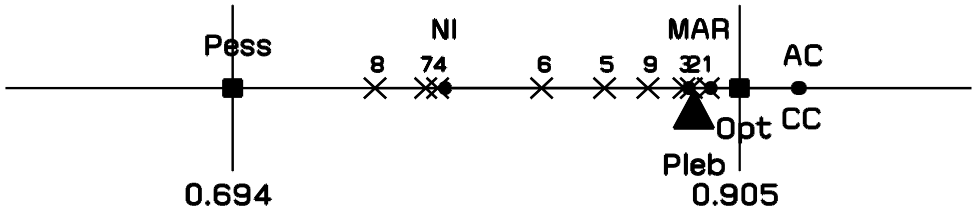


Figure 2: *The Slovenian Public Opinion Survey. Relative position for the estimates of "proportion of YES votes", based on the models considered in Rubin, Stern and Verhovar (1995) and on the Baker, Rosenberger and DerSimonian (1992) models. The vertical lines indicate the non-parametric pessimistic–optimistic 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. Numbers refer to the BRD models.)*

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 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 belief that can be put into them. Variations to a basic model can be constructed in different ways. The most obvious strategy is to consider various dependencies of the missing data process on the outcomes and/or on covariates. Alternatively, the distributional

Table 5: *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 '$\widehat{\theta}_{MAR}$' 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	$\widehat{\theta}$	C.I.	$\widehat{\theta}_{\text{MAR}}$
BRD1	(α, β)	6	-2495.29	0.892	[0.878;0.906]	0.8920
BRD2	(α, β_{j_1})	7	-2467.43	0.884	[0.869;0.900]	0.8915
BRD3	(α_{j_2}, β)	7	-2463.10	0.881	[0.866;0.897]	0.8915
BRD4	(α, β_{j_2})	7	-2467.43	0.765	[0.674;0.856]	0.8915
BRD5	(α_{j_1}, β)	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

assumptions of the models can be changed. This route will be followed in Section 4.1.

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 (Section 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 2 produces empirical proportions representing the π's with error. This imprecision disappears as the sample size tends to infinity. What remains is ignorance regarding the redistribution of all but the first four π's over the missing outcomes value. 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, θ say, a region of possible values which is consistent with Table 2 is called a region of ignorance. Analogously an observed incomplete table leaves ignorance regarding the would-be observed complete table, which in turn leaves imprecision regarding the true complete probabilities. The region of estimators for θ consistent with the observed data provides an estimated region of ignorance. The $(1-\alpha)100\%$ *region of uncertainty* is a larger region in the spirit of a confidence region, designed to capture the combined effects of imprecision and ignorance. Various ways for constructing regions of ignorance and regions of uncertainty are conceivable. For a single parameter, the regions obviously become intervals. These ideas have been developed in Molenberghs, Kenward and Goetghebeur (2001) and Kenward, Goetghebeur and Molenberghs (2001) and formalized by Vansteelandt *et al* (2006).

The estimated intervals of ignorance and intervals of uncertainty are shown in Table 6, while a graphical representation of the YES votes is given in Figure 3. Model 10 is defined as $(\alpha_{j_2}, \beta_{j_1 j_2})$ with

$$\beta_{j_1 j_2} = \beta_0 + \beta_{j_1} + \beta_{j_2}, \tag{1}$$

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

$$\alpha_{j_1 j_2} = \alpha_0 + \alpha_{j_1} + \alpha_{j_2}, \tag{2}$$

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, it should be noted that 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

Table 6: *The Slovenian Public Opinion Survey. Intervals of ignorance and intervals of uncertainty for the proportion θ (confidence interval) attending the plebiscite following from fitting.*

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

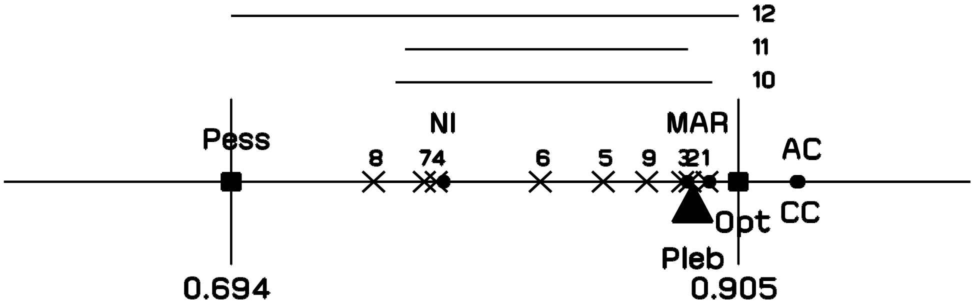


Figure 3: *The Slovenian Public Opinion Survey. Relative position for the estimates of "proportion of YES votes", based on the models considered in Rubin, Stern and Verhovar (1995) and on the BRD Models. The vertical lines indicate the nonparametric pessimistic-optimistic Bounds. (Pess: pessimistic boundary; 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. Numbers refer to the BRD models. Intervals of ignorance (Models 10–12) are represented by horizontal bars.)*

would have to be included. Thus, it appears that a relatively simple sensitivity analysis is sufficient to increase the insight in the information provided by the incomplete data about the proportion of valid YES votes.

4.2 An MAR Bodyguard for an MNAR Model

Molenberghs *et al* (2007) showed that, strictly speaking, the correctness of the alternative 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, the plausibility of which cannot be verified in empirical terms alone. This implies that an overall (omnibus) assessment of MAR *versus* MNAR is not possible, since 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 produces exactly 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. While this so-called MAR bodyguard generally does not belong to a conventional parametric family, its existence has important ramifications. Let us illustrate the use of the MAR bodyguard by means of 4 models from the BRD family, fitted to the independence and attendance outcomes, i.e., collapsing Table 1. We select models BRD1, BRD2, BRD7, and BRD9. Model BRD1 assumes missingness to be MCAR. All others are of the MNAR type. 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 7. Each of the four models is doubled up with its MAR counterpart.

Table 7 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 three models. The reason is clear: since model BRD1 belongs to the MAR family from the start, its counterpart BRD1(MAR) will not produce any difference, but merely copies the fit of BRD1 to the unobserved data, given the observed ones. 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 BRD7(MAR)≡BRD9(MAR) all saturate the observed data degrees of freedom, their complete-data fits are dramatically different.

Let us consider the results for the primary estimand θ obtained from fitting each of the nine BRD models (Table 5). BRD1 produces $\widehat{\theta} = 0.892$, exactly the same estimate as the first MAR estimate obtained by Rubin, Stern and Verhovar (1995). This should not come as a surprise, since both BRD1 and Rubin's model assume MAR and use information from the two main questions. 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 modes 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 the models BRD1(MAR)–BRD9(MAR), as displayed in the last column of Table 5, it is clear that this value is remarkably stable and hence a value of $\widehat{\theta} = 0.892$, based on the four bodyguards BRD6(MAR)–BRD9(MAR), is a sensible choice after all. Thus, a main contribution resulting from considering the bodyguards in this particular example, is the provision of a solid basis for the MAR-based estimate. Obviously, since Models BRD6(MAR)–BRD9(MAR) are exactly the same and exhibit a perfect fit, the corresponding probabilities $\widehat{\theta}_{\text{MAR}}$ are exactly 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, but only slightly so. For example, the probability of being in favour of independence combined with not attending ranges over 0.066–0.0685 across these four models.

We have made the following 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 into their MAR counterpart, from which inferences are drawn. As such, the MAR counterpart usefully supplements the original models BRD6–BRD9 and provide one further, important scenario to model the incomplete data. In principle, the same exercise can be conducted when the additional secession variable would be used.

4.3 Global Influence

One of the tools to perform a sensitivity analysis with an eye on individual observations is global influence, starting from case deletion. The methodology is based on the difference in log-likelihood between the model fitted to the entire data set on the one hand, and the data set minus one subject on the other hand. One might also consider, as we do here, the reverse operation of adding single case. Denoting by $\ell_i(\boldsymbol{\phi})$ the contribution of the i^{th} individual to the log-likelihood, where $\boldsymbol{\phi}$ is the s-dimensional vector of unknown parameters of the particular BRD model, the complete log-likelihood is

$$\ell(\boldsymbol{\phi}) = \sum_{i=1}^{N} \ell_i(\boldsymbol{\phi}). \tag{3}$$

Further, denote by

$$\ell_{(\pm i)}(\boldsymbol{\phi}) \tag{4}$$

the log-likelihood function, where the contribution of the i^{th} subject has been removed $(-i)$ or added $(+i)$. Cook's distances (CD) are based on measuring the discrepancy between either the maximized log-likelihoods (3) and (4) or (subsets of) the estimated parameter vectors $\widehat{\boldsymbol{\phi}}$ and $\widehat{\boldsymbol{\phi}}_{(\pm i)}$, with obvious notation. Precisely, we can consider

$$CD_{1i}(\boldsymbol{\phi}) = 2\left[\widehat{\ell}(\boldsymbol{\phi}) - \widehat{\ell}_{(\pm i)}(\boldsymbol{\phi})\right], \tag{5}$$

however, the focus in this paper is on

$$CD_{2i}(\boldsymbol{\phi}) = 2(\widehat{\boldsymbol{\phi}} - \widehat{\boldsymbol{\phi}}_{(\pm i)})' \, \ddot{L}^{-1} \, (\widehat{\boldsymbol{\phi}} - \widehat{\boldsymbol{\phi}}_{(\pm i)}), \tag{6}$$

with $\ddot{L}$ the matrix of second-order derivatives of $\ell(\boldsymbol{\phi})$, with respect to $\boldsymbol{\phi}$, evaluated at $\widehat{\boldsymbol{\phi}}$. Performing a global influence analysis on 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 3. Thus, instead of removing subjects on a one by one basis, we only need to remove one subject per cell and per covariate level, in case covariates are considered too.

Figure 4 shows a selection of the results for the global influence analysis on the SPO survey data. Only results for BRD4, 7, and 8 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 YES on the attendance question but a missing response on the independence question (cell #5), also influences BRD4's model parameters, though to a lesser extent.

For models BRD7–8, 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 NO for attendance and a missing independence response, or excluding a respondent with NO on both questions, yields changes in the model parameters of BRD7–8. These finding hint on the influential nature of subjects with a NO on the attendance question, which is likely related with this group's sparseness.

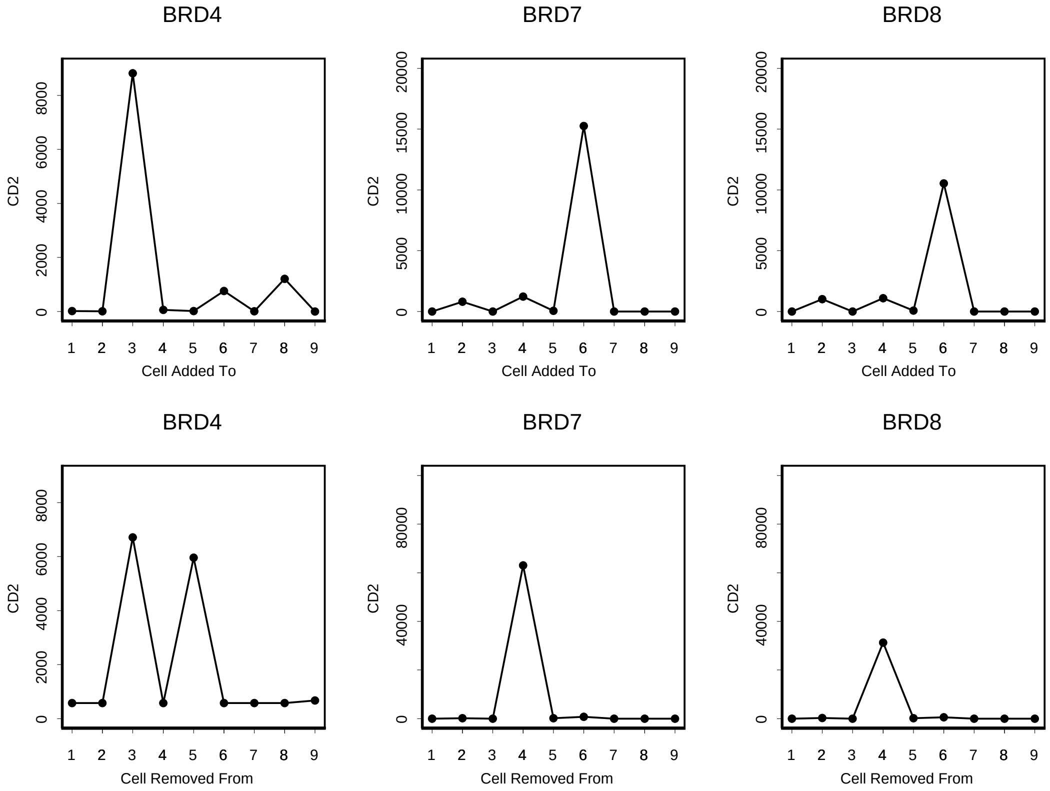


Figure 4: *The Slovenian Public Opinion Survey. Global influence analysis for BRD4, BRD7 and BRD8. Cook's distance measure, CD_{2i}, is evaluated when an observation is added to a specific cell (first row) and when an observation is deleted from a specific cell (second row).*

For all other models, 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.

4.4 Local Influence

A drawback of global influence is that the specific cause of the influence is hard to retrieve since, by deleting or adding a subject, all types of influence stemming from it are lumped together. Local influence, studying the effect of infinitesimally small model perturbations around a given null model, is more suitable for this purpose.

Let us now introduce the key concepts of local influence (Cook, 1986; Verbeke *et al*, 2001). Since the resulting influence diagnostics can in many cases be expressed analytically, they often allow for a decomposition into interpretable components, thus yielding additional insight. We denote the log-likelihood corresponding to a particular BRD model by $\ell(\boldsymbol{\phi}|\boldsymbol{\omega}) = \sum_{i=1}^{N} \ell_i(\boldsymbol{\phi}|\omega_i)$, in which $\ell_i(\boldsymbol{\phi}|\omega_i)$ is the contribution of the i^{th} individual, and where $\boldsymbol{\phi}=(\boldsymbol{\theta},\boldsymbol{\psi})$ is the s-dimensional vector, grouping, respectively, the parameters of the measurement and dropout models, but not including the $N \times 1$ vector $\boldsymbol{\omega} = (\omega_1, \omega_2, \ldots, \omega_N)'$ of weights defining the perturbation. Assume that $\boldsymbol{\omega}$ belongs to an open subset Ω of $\mathbb{R}^N$. For $\boldsymbol{\omega}$ equal to $\boldsymbol{\omega}_o = (0, 0, \ldots, 0)'$, $\ell(\boldsymbol{\phi}|\boldsymbol{\omega}_o)$ is the log-likelihood corresponding to the simpler of the two models.

Let $\widehat{\boldsymbol{\phi}}$ be the maximum likelihood estimator for $\boldsymbol{\phi}$, obtained by maximizing $\ell(\boldsymbol{\phi}|\boldsymbol{\omega}_o)$, and let $\widehat{\boldsymbol{\phi}}_{\boldsymbol{\omega}}$ denote the maximum likelihood estimator for $\boldsymbol{\phi}$ under $\ell(\boldsymbol{\phi}|\boldsymbol{\omega})$. The local influence approach compares $\widehat{\boldsymbol{\phi}}_{\boldsymbol{\omega}}$ with $\widehat{\boldsymbol{\phi}}$. Similar values indicate that the parameter estimates are robust with respect to perturbations in the direction of the extended model. Cook (1986) proposed to measure the distance between $\widehat{\boldsymbol{\phi}}_{\boldsymbol{\omega}}$ and $\widehat{\boldsymbol{\phi}}$ by the likelihood displacement, defined as $LD(\boldsymbol{\omega}) = 2[\ell(\widehat{\boldsymbol{\phi}}|\boldsymbol{\omega}_o) - \ell(\widehat{\boldsymbol{\phi}}_{\boldsymbol{\omega}}|\boldsymbol{\omega}_o)]$. This takes into account the variability of $\widehat{\boldsymbol{\phi}}$. Indeed, $LD(\boldsymbol{\omega})$ will be large if $\ell(\boldsymbol{\phi}|\boldsymbol{\omega}_o)$ is strongly curved at $\widehat{\boldsymbol{\phi}}$, which means that $\boldsymbol{\phi}$ is estimated with high precision, and small otherwise. Therefore, a graph of $LD(\boldsymbol{\omega})$ versus $\boldsymbol{\omega}$ contains essential information on the influence perturbations. It is useful to view this graphs as the geometric surface formed by values of the $N+1$ dimensional vector $\zeta(\boldsymbol{\omega}) = (\boldsymbol{\omega}', LD(\boldsymbol{\omega}))'$ as $\boldsymbol{\omega}$ varies throughout Ω. Since 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., at the normal curvatures $C_{\boldsymbol{h}}$ of $\zeta(\boldsymbol{\omega})$ in $\boldsymbol{\omega}_o$, in the direction of some N-dimensional vector $\boldsymbol{h}$ of unit length. Let $\boldsymbol{\Delta}_i$ be the s-dimensional vector defined by:

$$\boldsymbol{\Delta}_i = \left.\frac{\partial^2 \ell_i(\boldsymbol{\phi}|\omega_i)}{\partial \omega_i \partial \boldsymbol{\phi}}\right|_{\boldsymbol{\phi}=\widehat{\boldsymbol{\phi}}, \omega_i=0}, \tag{7}$$

and define $\boldsymbol{\Delta}$ as the $(s \times N)$ matrix with $\boldsymbol{\Delta}_i$ as its i^{th} column. Let $\ddot{L}$ denote the $(s \times s)$ matrix of second order derivatives of $\ell(\boldsymbol{\phi}|\boldsymbol{\omega}_o)$ with respect to $\boldsymbol{\phi}$, also evaluated at $\boldsymbol{\phi} = \widehat{\boldsymbol{\phi}}$. Cook (1986) has then shown that $C_{\boldsymbol{h}}$ can be written as

$$C_{\boldsymbol{h}} = 2\left|\boldsymbol{h}'\,\boldsymbol{\Delta}'\,(\ddot{L})^{-1}\,\boldsymbol{\Delta}\,\boldsymbol{h}\right|, \tag{8}$$

for any direction $\boldsymbol{h}$. One choice is the vector $\boldsymbol{h}_i$ containing one in the i^{th} position and zero elsewhere, corresponding to the perturbation of the i^{th} subject only, reflecting the influence of allowing the i^{th} subject to drop out in a more general fashion than the others. The corresponding local influence measure, denoted by C_i, then becomes $C_i = 2\left|\boldsymbol{\Delta}_i'\,(\ddot{L})^{-1}\,\boldsymbol{\Delta}_i\right|$. Another important direction is the direction $\boldsymbol{h}_{\text{max}}$ of maximal normal curvature C_{max}. It shows how to perturb the model to obtain the largest local changes in the likelihood displacement. It is readily seen that C_{max} is the largest eigenvalue of $-2\,\boldsymbol{\Delta}'\,(\ddot{L})^{-1}\,\boldsymbol{\Delta}$, with $\boldsymbol{h}_{\mathsf{max}}$ the corresponding eigenvector. Calculation of local influence measures reduces to evaluation of $\boldsymbol{\Delta}$ and $\ddot{L}$ and a convenient computational scheme can be used whenever a program is available to fit the full alternative model, i.e., the model at the end of an edge in Figure 1 since it then suffices to compute the second derivative at $(\widehat{\boldsymbol{\phi}}, \omega_i = 0)$, for each observation separately, from which the $\boldsymbol{\Delta}_i = (\boldsymbol{\phi}, \omega_i)$ subvector is selected.

It should be noted that $C_{\boldsymbol{h}}$ is a measure of the local influence on the log-likelihood function, i.e., quantifying the effect of perturbations in terms of the displacement in the log-likelihood. At times, however, it might be more meaningful to assess the influence that infinitesimal changes may have on a particular function of the parameters, rather than on the log-likelihood itself. In the case of contingency tables, for instance, one might be more interested in the impact of perturbations on the predicted cell counts, $Z_{r_1 r_2, j_1 j_2}$, which are functions of the parameter vector $\boldsymbol{\phi}$. In such cases, when $Z(\boldsymbol{\phi})$ denotes a particular function of the model parameters, the expression for $C_{\boldsymbol{h}}$ can be further generalized as:

$$C_{\boldsymbol{h}} = 2\left| h'\,\boldsymbol{\Delta}'\,(\ddot{\boldsymbol{L}})^{-1}\,\ddot{\boldsymbol{Z}}\,(\ddot{\boldsymbol{L}})^{-1}\,\boldsymbol{\Delta}\,h \right|, \tag{9}$$

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

$$\boldsymbol{\Delta}_{ij} = \left.\frac{\partial^2 \ell(\boldsymbol{\phi}|\boldsymbol{\omega})}{\partial\phi_i \partial\omega_j}\right|_{\boldsymbol{\phi}=\widehat{\boldsymbol{\phi}}, \boldsymbol{\omega}=\boldsymbol{\omega}_o}, \tag{10}$$

$$\ddot{\boldsymbol{L}}_{i\ell} = \left.\frac{\partial^2 \ell(\boldsymbol{\phi}|\boldsymbol{\omega}_o)}{\partial\phi_i \partial\phi_\ell}\right|_{\boldsymbol{\phi}=\widehat{\boldsymbol{\phi}}} \tag{11}$$

$$\ddot{\boldsymbol{Z}}_{il} = \left.\frac{\partial^2 Z(\boldsymbol{\phi})}{\partial\phi_i \partial\phi_\ell}\right|_{\boldsymbol{\phi}=\widehat{\boldsymbol{\phi}}}, \tag{12}$$

with $i, \ell = 1, \ldots, p$ and $j = 1, \ldots, q$. It is easy to see that (9) reduces to (8) when the function of interest, $Z(\boldsymbol{\phi})$, is the log-likelihood $\ell(\boldsymbol{\phi}|\boldsymbol{\omega})$ itself. Whereas (8) quantifies influence in terms of the displacement in the log-likelihood function, (9) describes influence through the displacement in the particular function of interest.

4.4.1 Perturbation in Parameters: One BRD Model vs. Another

We first consider perturbations of a given BRD model in the direction of another BRD model with one or more parameters in which the first model is nested, implying that perturbations lie along the edges of Figure 1. For such a nested pair, the simpler of the two models equates two parameters from the more complex one. For example, BRD4 includes the parameter $\beta_{.j_2}, (j_2 = 1, 2)$, whereas for BRD1, only $\beta_{..}$ is included. For the influence analysis, ω_i is then included as a contrast between two such parameters; for the perturbation of BRD1 in the direction of BRD4, one considers $\beta_{..}$ and $\beta_{..} + \omega_i$. The vector of all ω_i's defines the direction in which such a perturbation is considered.

To illustrate this approach, we begin by first defining the log-likelihood for the BRD family of models. We have

$$\begin{aligned} \ell(\boldsymbol{\phi}|\boldsymbol{\omega}) \quad &= \quad \sum_{j_1,j_2} Z_{11,j_1j_2} \ \ln \pi_{11,j_1j_2} \quad + \quad \sum_{j_1} Z_{10,j_1+} \ \ln \pi_{10,j_1+} \\ &+ \quad \sum_{j_2} Z_{01,+j_2} \ \ln \pi_{01,+j_2} \quad + \quad Z_{00,++} \ \ln \pi_{00,++}, \end{aligned} \tag{13}$$

where $\pi_{r_1r_2,j_1j_2} = p_{j_1j_2}\, q_{r_1r_2|j_1j_2}$, with $p_{11} = p_1, p_{12} = p_2, p_{21} = p_3, p_{22} = 1 - p_1 - p_2 - p_3$, and

$$q_{r_1r_2|j_1j_2} = \frac{\exp\left\{\alpha_{j_1j_2}(1-r_1) + \beta_{j_1j_2}(1-r_2) + \gamma(1-r_1)(1-r_2)\right\}}{1 + \exp\left(\alpha_{j_1j_2}\right) + \exp\left(\beta_{j_1j_2}\right) + \exp\left(\alpha_{j_1j_2} + \beta_{j_1j_2} + \gamma\right)}. \tag{14}$$

Distinction among the 9 BRD models occurs in the latter expression. For instance, for BRD4 with $(\alpha_{..}, \beta_{.j_2})$, expression (14) yields:

$$\begin{aligned} q_{r_1r_2|j_11} \quad &= \quad \frac{\exp\left\{\alpha_{..}(1-r_1) + \beta_{..}(1-r_2) + \gamma(1-r_1)(1-r_2)\right\}}{1 + \exp\left(\alpha_{..}\right) + \exp\left(\beta_{..}\right) + \exp\left(\alpha_{..} + \beta_{..} + \gamma\right)}, \\ q_{r_1r_2|j2} \quad &= \quad \frac{\exp\left\{\alpha_{..}(1-r_1) + \left(\beta_{..} + \omega_i\right)(1-r_2) + \gamma(1-r_1)(1-r_2)\right\}}{1 + \exp\left(\alpha_{..}\right) + \exp\left(\beta_{..} + \omega_i\right) + \exp\left(\alpha_{..} + \beta_{..} + \omega_i + \gamma\right)}. \end{aligned}$$

Note that, for $\omega_i = 0$, the two previous expressions are equivalent and BRD4 reduces to the simpler BRD1. For this pair of nested models, BRD4 contains one parameter more than BRD1, this extra parameter being the distinguishing feature between both models. That is, under the more complicated model BRD4, the extra parameter ω_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. Similar motivations hold for the other pairs of nested BRD models. Given now the fully-defined log-likelihood, one can proceed with deriving local influence measures (8) and (9).

Note that the influence analysis focuses on the missingness model, rather than on the measurement model parameters. This may be seen as slightly odd, since often scientific interest focuses on the measurement model parameters. However, it has been documented (Rubin, 1994; Kenward, 1998; Verbeke *et al*, 2001) that the missingness model parameters are often the most sensitive ones to take up all kinds of misspecification and influential features. These may then, in turn, impact conclusions coming from the measurement model parameters, such as time evolution, or combinations thereof, such as covariate effects for certain groups of responders.

For the SPO data, we consider local influence measures on both the likelihood displacement (8) and the predicted cell counts (9) for different model pairs. Although 12 model nestings are possible(Figure 1), we focus on the model pairs BRD1 *vs.* BRD4, BRD3 *vs.* BRD7, and BRD4 *vs.* BRD7. The rationale for these choices, in addition to conciseness, is that in these 3 model pairs substantial influence was seen when considering local influence on the likelihood displacement. In addition, for the local influence on the predicted cell counts, discussed in the next section, these three model pairs are indicative for the various features that were seen across all 12 comparisons.

Figure 5 shows, for the 3 comparisons considered, the influence measures C_i, plotted against the i^{th} observed cell, as well as against each subject within that cell, and $\boldsymbol{h}_{\text{max}}$ against the i^{th} observed cell. A peak in the graph at a particular cell indicates that the corresponding cell drives the data towards the more complex, rather than the simpler model. For the comparison of BRD1 *vs.* BRD4, a peak is observed at cell #6, for both C_i and $\boldsymbol{h}_{\text{max}}$, 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 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 cell #9, subjects with a missing response on both questions, implying that such subjects drive the data in the direction of BRD7 $(\alpha_{.j_2}, \beta_{.j_2})$ rather than BRD3 $(\alpha_{.j_2}, \beta_{..})$. That is, missingness in the independence question is driven to depend on the corresponding unobserved answer by subjects with missing responses on both questions, and, also 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 $(\alpha_{.j_2}, \beta_{.j_2})$ from BRD4 $(\alpha_{..}, \beta_{.j_2})$. 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. Graphs of the local influence measure (9) on the predicted cell counts are presented in Figure 6, with the graphs for the 16 predicted cell counts arranged in their respective positions as in Table 2. From the first panel, for model pair BRD1 *vs.* BRD4, we observe that the influence graphs show similar shapes, albeit with differing magnitudes, for a particular cell (j_1, j_2), across the four missingness patterns. For instance, the influence curves for $Z_{r_1 r_2,11}$ (upper left corners) for $(r_1, r_2) = (1,1), (1,0), (0,1), (0,0)$ have more or less identical shapes. Occurrences of peaks at particular cells are thus common across the missingness patterns, yielding more or less a clear result for each cell (j_1, j_2). 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 therefore 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. Cells #2 and #5 denote, respectively, respondents having YES on attendance/NO on independence, and YES on attendance/missing value on independence, and these respondents make the predicted cell count $Z_{r_1 r_2,11}$ seem to have come more from BRD4 rather than from 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 6b) 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. For $(j_1, j_2) = (1,1)$ and $(j_1, j_2) = (1,2)$, i.e., top row of the 4 sets of tables, although relative peaks are observed at the same positions across the 4 sets of tables, the degree of the peak varies across the missingness patterns, causing some to appear more like a peak and some less so. 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. We proceed to look at the results for $(j_1, j_2) = (1,1)$ and $(j_1, j_2) = (1,2)$. Across the 4 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 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 $Z_{r_1 r_2,22}$, since we observe similarly shaped influence curves across the missingness patterns for cell position $(j_1, j_2) = (2, 2)$, with peaks either at cell #9 or cell #6.

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 what was observed for this model pair when considering influence on the likelihood displacement. 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.

4.4.2 Perturbation in Cell Probabilities

Another route to studying local influence is to add an infinitesimally small value to the cell probabilities. Such an approach leads to the following expression for the log-likelihood:

$$\begin{aligned}\ell(\boldsymbol{\phi}|\boldsymbol{\omega}) &= \sum_{j_1,j_2} (Z_{11,j_1j_2} + N\omega_{11,j_1j_2}) \ln \pi_{11,j_1j_2} \\ &\quad + \sum_{j_1} (Z_{10,j_1+} + N\omega_{10,j_1+}) \ln \pi_{10,j_1+} \\ &\quad + \sum_{j_2} (Z_{01,+j_2} + N\omega_{01,+j_2}) \ln \pi_{01,+j_2} \\ &\quad + (Z_{00,++} + N\omega_{00,++}) \ln \pi_{00,++},\end{aligned}$$

where $\pi_{r_1r_2,j_1j_2} = p_{j_1j_2}\, q_{r_1r_2|j_1j_2}$, with $p_{11} = p_1, p_{12} = p_2, p_{21} = p_3, p_{22} = 1 - p_1 - p_2 - p_3$, and

$$q_{r_1r_2|j_1j_2} = \frac{\exp\{\alpha_{j_1j_2}(1-r_1) + \beta_{j_1j_2}(1-r_2) + \gamma(1-r_1)(1-r_2)\}}{1 + \exp(\alpha_{j_1j_2}) + \exp(\beta_{j_1j_2}) + \exp(\alpha_{j_1j_2} + \beta_{j_1j_2} + \gamma)}.$$

Under the null model, $\boldsymbol{\omega} = \boldsymbol{\omega}_o = \mathbf{0}$, and the above log-likelihood reduces to:

$$\begin{aligned} \ell(\boldsymbol{\phi}|\boldsymbol{\omega}_o) &= \sum_{j_1,j_2} Z_{11,j_1j_2} \ln \pi_{11,j_1j_2} + \sum_{j_1} Z_{10,j_1+} \ln \pi_{10,j_1+} \\ &\quad + \sum_{j_2} Z_{01,+j_2} \ln \pi_{01,+j_2} + Z_{00,++} \ln \pi_{00,++}. \end{aligned}$$

Computation of local influence measures (8) and (9) is straightforward once the log-likelihood, $\ell(\boldsymbol{\phi}|\boldsymbol{\omega})$, is clearly defined.

It is important to note that the approach to local influence described in the previous section differs from the approach proposed here, since now the perturbation is done directly in the observed cell probabilities, rather than the parameters of the model. This implies that we are perturbing the cells one at a time and observing which one brings about the largest changes, in likelihood or in the predicted cell counts, within a given BRD model. Consequently, although influence measures are computed in the same fashion, a difference in interpretation is warranted. A peak in the influence curve now represents the particular observed cell at which a probability perturbation causes substantial displacement in either the log-likelihood or in the predicted cell counts.

Let us consider the results of this approach to local influence for the SPO survey. We first derive influence measures on the likelihood displacement; these are graphed in Figure 7. For most BRD models, it seems small perturbations in the probabilities of cells #3 and/or #4 has a large influence. That is, if we slightly alter the probabilities with which the "NO-on-attendance/YES-on- independence" or the NO/NO respondents occur, we can expect substantial likelihood displacement. Also notable is the influence of changes in cell #6 for BRD8, implying that under this model, changing the probability of the NO/missingness category slightly causes displacement in the likelihood. These observations suggest that the most influential cells for virtually all BRD models are the completers answering NO on attendance, likely attributable to the small counts in these cells, while for BRD8, it is those subjects answering NO on attendance and unobserved response on independence that are influential.

Figure 8 provides a summary of the results of the local influence analysis on the predicted cell counts when perturbing a particular cell probability. No particular influence can be seen for any BRD model when perturbing probabilities of cells #1 and #2, as might be expected since the observed cell counts in these cells are large, and thus infinitesimal changes in their respective cell probabilities may not have a large impact. We can also see that small perturbations in cell probability 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 result that can be observed from Figure 8 is that for perturbations in cell probability 4 (NO/NO respondents), which yields influence on *all* 16 predicted cell counts in most of the higher-numbered BRD models 4 to 9. Also of particular interest are the results for perturbations in cell #6, indicating that it is primarily under BRD8 where a large influence is observed in the most of the predicted cell counts. Finally, we note that changes in the probability of the doubly missing category (cell #9) affects only the predicted cell counts of this missingness pattern and only under BRD's 1,2 and/or 3.

5 Concluding Remarks

In this paper, we have presented a variety of sensitivity analyses for the incomplete set of data arising from the well-known Slovenian Public Opinion Survey. A first family of sensitivity analyses is based on considering a variety of models. First, simple analyses have been augmented with a non-parametric interval, providing absolute bounds for the proportion of people in favor of independence, which is expressed through at the same time possessing and expressing an opinion in favor of independence. A nonparametric interval was supplemented with a finite collection of identifiable models from the BRD family, as well as infinite collections, resulting from overspecified models. In addition, for each of these models, the corresponding MAR bodyguards were calculated; these are models with the same fit to the observed data than their corresponding original models, but with missingness of the MAR type. Whereas the nonparametric range is [0.694,0.905], with the parametric ranges subsets thereof, the MAR models center around 0.89, close to the actual plebiscite values.

A second family of sensitivity analyses studies influence of observations on the model's conclusions, expressed through either parameters or cell counts. Both global influence, i.e., case deletion, and local influence, based on infinitesimal perturbation, were considered. It was found that perturbing some, but not all, small counts can have an extremely large effect on the conclusions, often through partially observed or unobserved cells. Such influences can strongly affect conclusions about estimands such as the one considered here. Indeed, the proportion of people attending the plebiscite and at the same time being in favor of independence is made up of adding up the

(1,1) cell across all missingness patterns, and hence depends on how a model distributes partially observed counts over the cells.

Therefore, we believe it is important to conduct a sensitivity analysis and in particular that such an analysis that combines insight from considering families of models on the one hand and from studying influence is able to paint a relatively complete picture. This allows one to put a perspective on the conclusions that can confidently be reached about an estimand based on an incomplete set of data, something that considering a single model arguably never can.

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.

Cook, R.D. (1986). Influential observations in linear regression. *J. Am. Statist. Ass.* **74**, 169–174.

Cook, R.D. (1986). Assessment of local influence. *J. Roy. Statist. Soc. Ser. 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., 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 bodyguard. *Submitted for publication*.

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., Thijs, H., Kenward, M.G. and Verbeke, G. (2003). Sensitivity analysis of continuous incomplete longitudinal outcomes. *Statistica Neerlandica* **57**, 112–135.

Molenberghs, G., Beunckens, C., Sotto, C. and Kenward, M.G. (2006). Every missing not at random model has got a missing at random bodyguard. *Submitted for publication*.

Molenberghs, G. and Kenward, M.G. (2007). *Handling Incomplete Data From Clinical Studies.* New York: Wiley.

Rubin, D.B., Stern, H.S. and Verhovar, V. (1995). Handling "Don't Know" survey responses: the case of the Slovenian plebiscite. *J. Am. Statist. Ass.* **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.

Thijs, H., Molenberghs, G. and Verbeke, G. (2000). The milk protein trial: influence analysis of the dropout process. *Biometrical J.* **42**, 617–646.

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., 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.

Table 7: *Analysis of the Slovenian Public Opinion Survey, 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

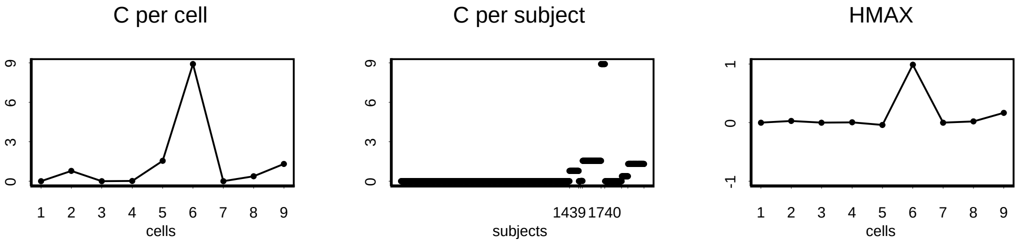
1439	78	3.2	155.8	142.4	44.8	0.4	112.5
16	16	0.0	32.0	1.6	9.2	0.0	23.1

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

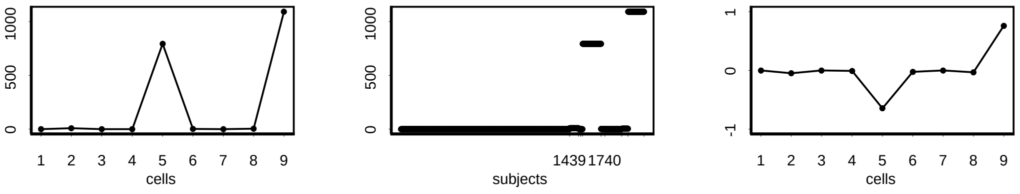
Fit of BRD7(MAR) and BRD9(MAR) to complete data

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



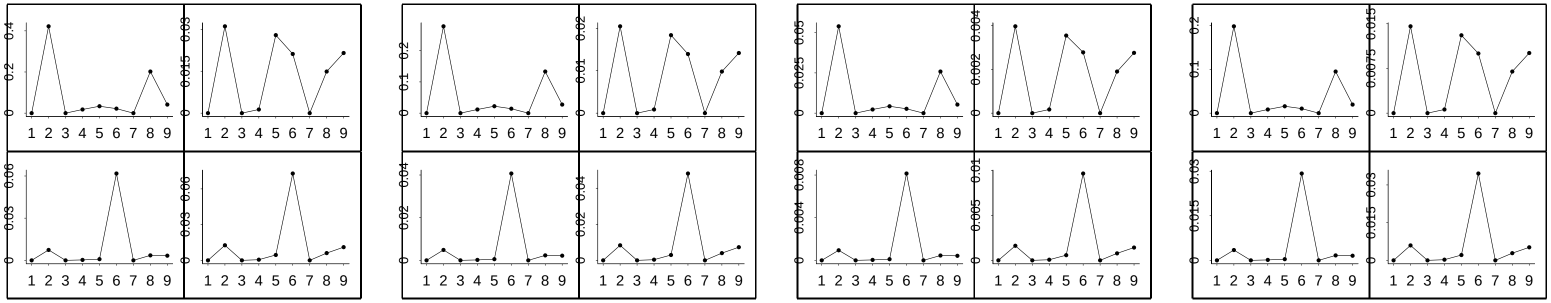
(a) *BRD1 vs BRD4*

(b) *BRD3 vs. BRD7*

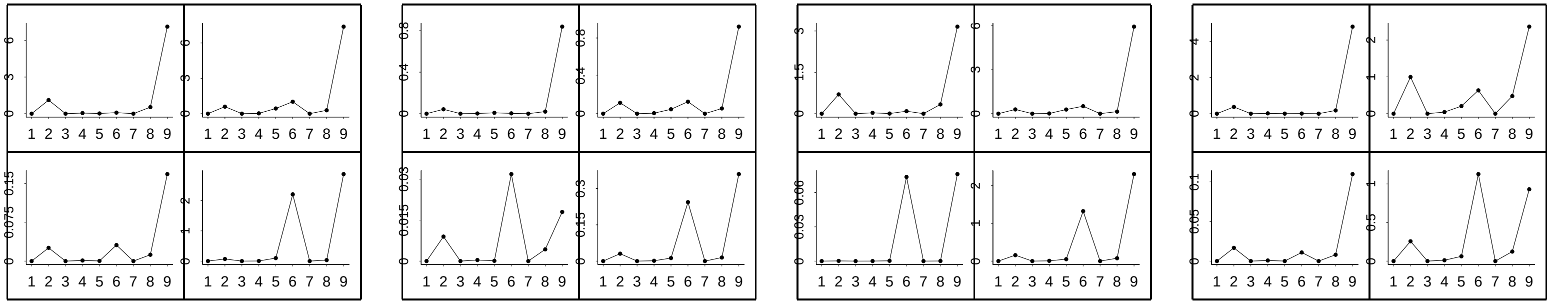


(c) *BRD4 vs. BRD7*

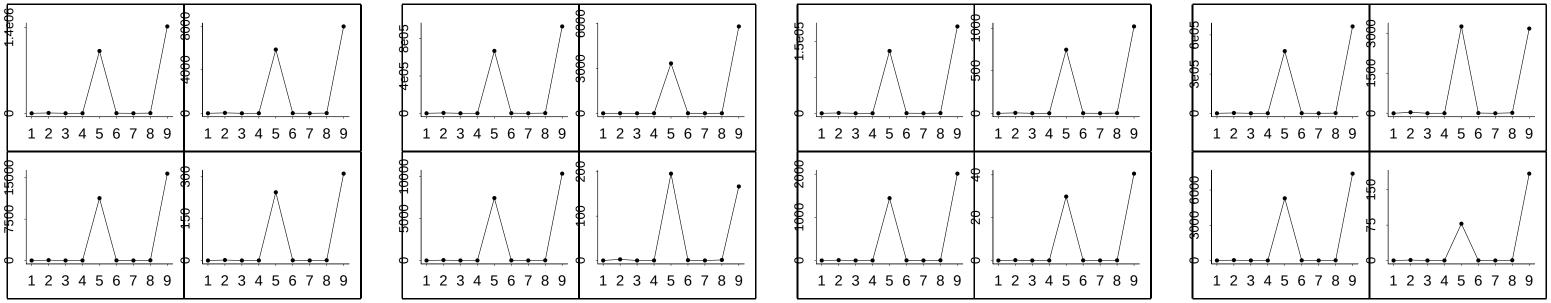
Figure 5: *The Slovenian Public Opinion Survey. Local influence analysis on parameters for model pairs (a) BRD1 vs. BRD4, (b) BRD3 vs. BRD7, and (c) BRD4 vs. BRD7. First column shows the local influence measure C_i at the i^{th} observed cell; the second column shows the same measure but plotted for each of the subjects within the i^{th} observed cell; and, the third column shows $\boldsymbol{h}_{max}$ for the i^{th} observed cell.*



(a) *BRD1 vs. BRD4*



(b) *BRD3 vs. BRD7*



(c) *BRD4 vs. BRD7*

Figure 6: *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 2) against the 9 observed cells (as labelled in Table 3).*

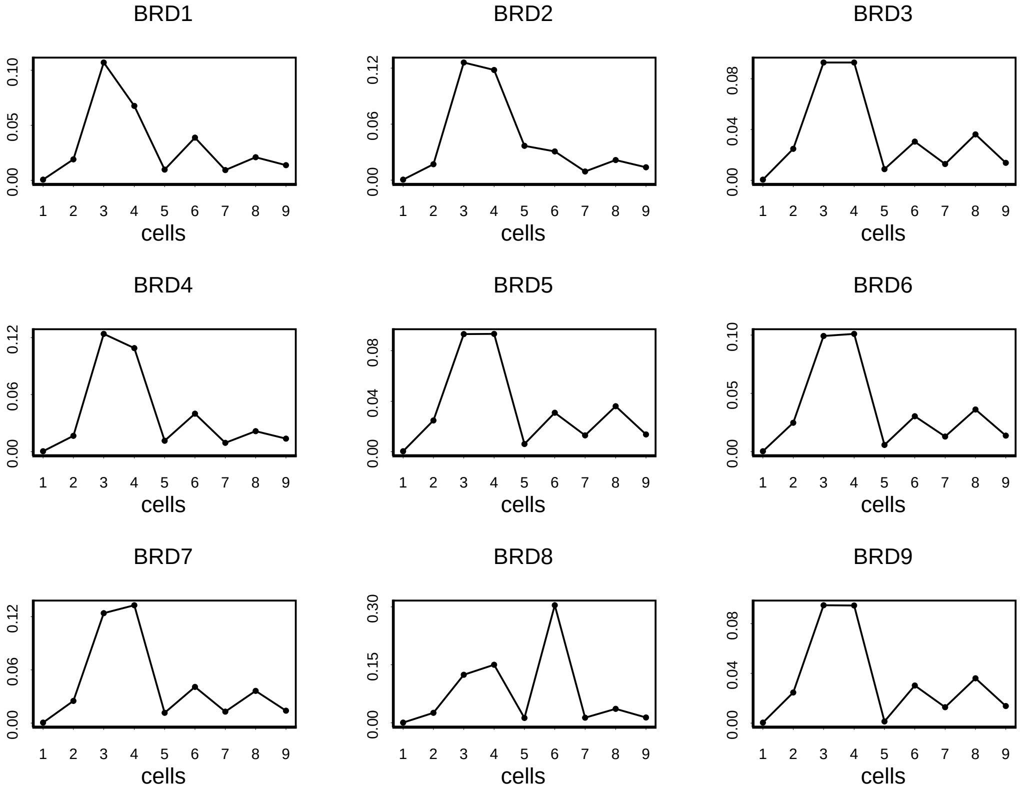
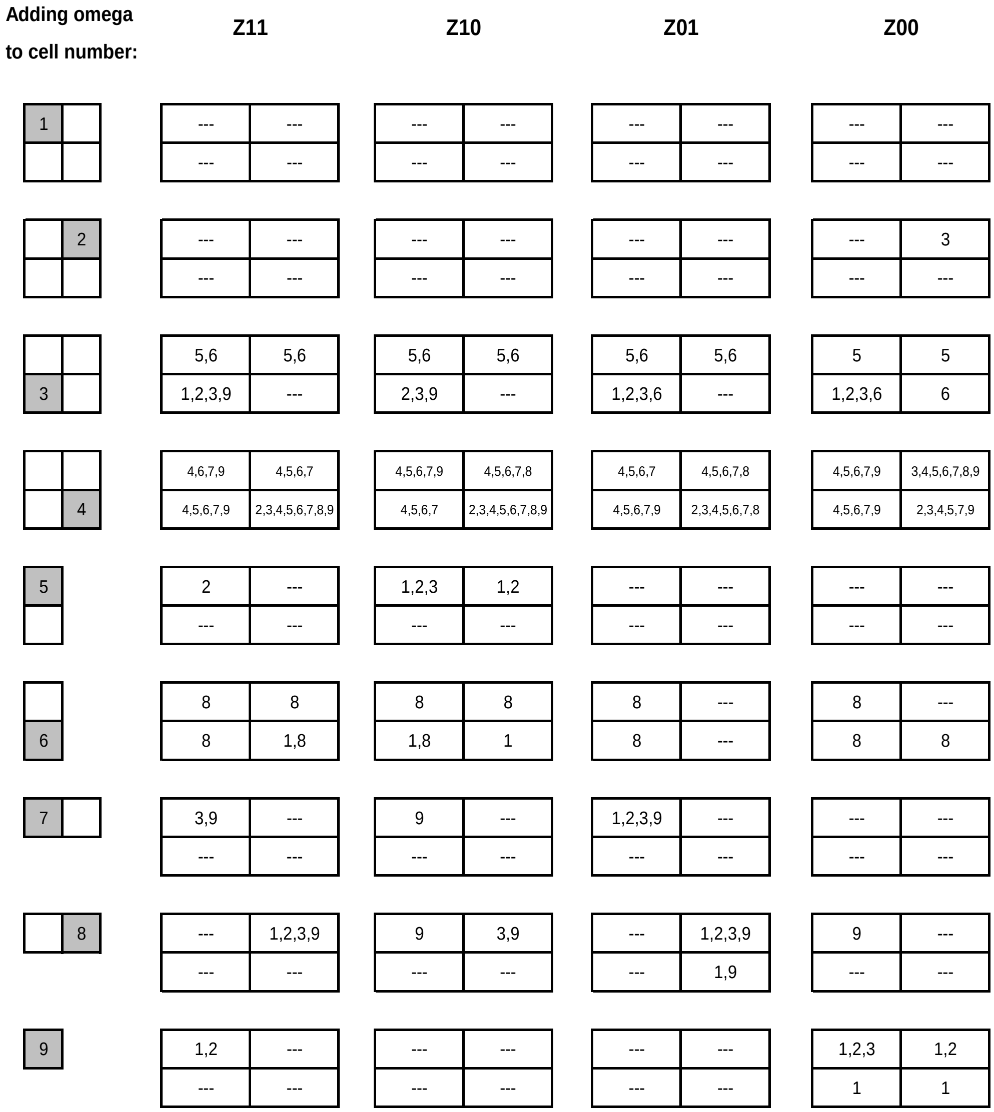


Figure 7: *The Slovenian Public Opinion Survey. Local influence analysis on the log-likelihood for the 9 BRD models. Plots show C_i values against the 9 observed cells for each BRD model.*



Adding omega to cell number:	Z11		Z10		Z01		Z00	
1	---	---	---	---	---	---	---	---
	---	---	---	---	---	---	---	---
2	---	---	---	---	---	---	---	3
	---	---	---	---	---	---	---	---
3	5,6	5,6	5,6	5,6	5,6	5,6	5	5
	1,2,3,9	---	2,3,9	---	1,2,3,6	---	1,2,3,6	6
4	4,6,7,9	4,5,6,7	4,5,6,7,9	4,5,6,7,8	4,5,6,7	4,5,6,7,8	4,5,6,7,9	3,4,5,6,7,8,9
	4,5,6,7,9	2,3,4,5,6,7,8,9	4,5,6,7	2,3,4,5,6,7,8,9	4,5,6,7,9	2,3,4,5,6,7,8	4,5,6,7,9	2,3,4,5,7,9
5	2	---	1,2,3	1,2	---	---	---	---
	---	---	---	---	---	---	---	---
6	8	8	8	8	8	---	8	---
	8	1,8	1,8	1	8	---	8	8
7	3,9	---	9	---	1,2,3,9	---	---	---
	---	---	---	---	---	---	---	---
8	---	1,2,3,9	9	3,9	---	1,2,3,9	9	---
	---	---	---	---	---	1,9	---	---
9	1,2	---	---	---	---	---	1,2,3	1,2
	---	---	---	---	---	---	1	1

Figure 8: *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.*