## Supplementary material to the article "Sensitivity analysis and choosing between alternative polytomous IRT models using Bayesian model comparison criteria" published by "Communications in Statistics - Simulation and Computation"

Marcelo A. da Silva<sup>1,2</sup>, Jorge L. Bazán<sup>1</sup> and Anne Corinne Huggins-Manley<sup>3</sup>

<sup>1</sup> Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo. Av. Trabalhador São Carlense, 400. 13566-590 - São Carlos, SP, Brasil

<sup>2</sup> Departamento de Estatística, Universidade Federal de São Carlos, Rod. Washington Luiz, km 235. 13565-905 - São Carlos - SP, Brasil

<sup>3</sup> College of Education, University of Florida, USA

## **1 DINA model estimation methods**

Following the results shown in Section 5 in the paper, we use the GR model in this application. The fit of this model showed that for items (2), (4), (6), (9), (10), (11), (16), (19), (20) and (21), at least one of the response categories has no probability greater than the others of being selected for the whole intensity of depressive symptoms, as can be seen the item characteristic curves for item (9) (concerning suicidal thoughts) in the second graph in Figure 1. In this case, for all values of depressive symptoms, curve (3) of the item characteristic curves for the item characteristic curves for item (7) (concerning self-hate), shown in the first graph in Figure 1. The curves relate to the item response categories.

The use of GR model indicated through item information curves that the items (1) to (9) and items (12), (13), (15) and (17) provide a amount of relevant information to measure the latent trait of the respondents.

We present in Table 1 estimates of the items' parameters from the GR model after performing the grouping of the categories described above. We can observe that some items have the slope parameter below 1 (items (11), (14), (16), (18), (19), (20) and (21)), and that the item (19) has the lowest value, which is 0.452. Regarding the difficulty parameters, item (19) also draws our attention, because it has high values ( $b_1 = 5.193, b_2 = 9.581$  and  $b_3 = 11.604$ ).

rabera 1. Estimates of the terms parameters of the GK model to the BDI data.					
Items		a	$b_1$	$b_2$	$b_3$
1	Sadness	1.592	-0.546	2.032	3.097
2	Pessimism	1.581	0.429	2.301	2.932
3	Feelings of failure	1.990	0.630	2.056	2.972
4	Dissatisfaction	1.687	-0.084	1.715	2.249
5	Guilt	1.728	0.636	2.262	3.350
6	Punishment	1.235	0.656	1.713	1.954
7	Self-hate	1.767	-0.200	1.199	2.657
8	Self-accusation	1.350	-0.588	2.396	3.554
9	Suicidal thoughts	1.715	1.500	3.121	3.590
10	Weeping	1.186	0.982	2.236	2.555
11	Irritability	0.915	-0.036	2.245	3.053
12	Social withdrawal	1.294	0.527	2.279	3.930
13	Indecision	1.484	0.386	1.355	2.735
14	Change in self-awareness	0.922	0.773	1.809	3.359
15	Work difficulties	1.696	0.075	1.586	3.054
16	Insomnia	0.860	0.000	2.911	4.361
17	Tiredness	1.246	-0.476	1.814	3.373
18	Loss of appetite	0.895	1.182	3.377	5.156
19	Loss of weight	0.452	5.193	9.581	11.604
20	Somatic worries	0.805	1.007	4.164	6.358
21	Loss of libido	0.706	1.836	3.941	6.292

Tabela 1: Estimates of the items' parameters of the GR model to the BDI data.

\_



Figura 1: Item characteristic curves for items (7) and (9), concerning self-hate and suicidal thoughts, respectively. Curve x corresponds to category x.