On-line Appendix for Gingerich, DW, V Oliveros, A Corbacho, and M Ruiz-Vega, "When to Protect? Using the Crosswise Model to Integrate Protected and Direct Responses in Surveys of Sensitive Behavior"

EM Algorithm for SST Questioning Only

Here we present the EM algorithm employed in the Monte Carlo analyses to estimate the parameters of the explanatory model when only SST questioning is utilized. Appendix Table 6 presents the relevant probability table for the complete data under the assumption of honesty given protection.

Utilizing the notation employed previously in the text, the expected value of the log-likelihood of the complete data (up to a constant) is equal to

$$\mathbb{E}[\ln L_c(\boldsymbol{\xi}|Y,\mathbf{X},Z)] = \sum_i^n \mathbb{P}_Z(\{2,4\})_i \ln f(\mathbf{X}_i;\boldsymbol{\beta}) + \sum_i^n (1-\mathbb{P}_Z(\{2,4\})_i)\ln(1-f(\mathbf{X}_i;\boldsymbol{\beta})), \tag{A1}$$

where $\mathbb{P}_Z(\{2,4\})_i$ is the conditional probability of bearing the sensitive trait given a respondent's observed response and background characteristics.

On the jth iteration of the E-M algorithm, this quantity is equal to

$$\mathbb{P}_Z(\{2,4\})_i^{(j)} = \begin{cases} \frac{(1-p)f(\mathbf{X}_i;\boldsymbol{\beta}^{(j)})}{p(1-f(\mathbf{X}_i;\boldsymbol{\beta}^{(j)}))+(1-p)f(\mathbf{X}_i;\boldsymbol{\beta}^{(j)})} & \text{if} \quad Y_i = 0 \\ \frac{pf(\mathbf{X}_i;\boldsymbol{\beta}^{(j)})}{pf(\mathbf{X}_i;\boldsymbol{\beta}^{(j)})+(1-p)(1-f(\mathbf{X}_i;\boldsymbol{\beta}^{(j)}))} & \text{if} \quad Y_i = 1 \end{cases} \tag{A2}$$

Based on the above, the current conditional parameter vector $\boldsymbol{\beta}^{(j+1)}$ is identical to that shown in (13), save for the fact that $\mathbb{P}_Z(\{2,4\})_i^{(j)}$ is inserted into the equation as the relevant response variable. The EM algorithm proceeds by iterating through $\boldsymbol{\beta}^{(j)}$ and $\mathbb{P}_Z(\{2,4\})_i^{(j+1)}$ until convergence is achieved.

Appendix Table 6. Probability table for complete data for SST questioning only

Z	Y	outcome	probability
1	0	$(y^A=0,\theta=0)$	$p(1-\pi)$
2	0	$(y^A=0,\theta=1)$	$(1-p)\pi$
3	1	$(y^A=1,\theta=0)$	$(1-p)(1-\pi)$
4	1	$(y^A=1,\theta=1)$	$p\pi$