Supplementary material

for

# The number of response categories in ordered response models

Maria Iannario[1], Anna Clara Monti[2], Pietro Scalera[1],

[1] Department of Political Sciences, University of Naples Federico II, Italy

[2] Department of Law, Economics, Management and Quantitative Methods, University of Sannio, Italy

This supplementary material supplies figures and tables which integrate the analyses on the choice of the thresholds (in Section S.1), on merging categories (in Section S.2) and on the choice of $m$ for a given sample size $n$ (in Section S.3).

## S.1 Thresholds

This section describes the construction of the four sets of thresholds and develops further analyses on the impact of the thresholds with respect to the efficiency of the estimators, the power of the test, the efficiency loss produced by merging categories, and the choice of $m$ for a given sample size $n$. Tables are supplied which extend the analyses carried out in Sections 5, 6 and 7 of the main paper with the $\alpha^{ED}$ thresholds to the other sets of thresholds.

### S.1.1 Construction of the thresholds

Let $R = Y^*_{0.995} - Y^*_{0.005}$ where $Y^*_{0.005}$ and $Y^*_{0.995}$ are the 0.5% and 99.5% percentiles of the distribution of $Y^*$ simulated from 5 millions of observations. The equidistant thresholds satisfy the constraint $\alpha^{ED}_j - \alpha^{ED}_{j-1} = h$, where $h = R/m$. The first threshold is $\alpha^{ED}_1 = Y^*_{0.005} + h$ while the following thresholds are obtained as $\alpha^{ED}_j = \alpha^{ED}_1 + (j-1)h$.

The thresholds $\alpha_j^{ID}$ consider a central class centered at the median of $Y^*$ in case of odd $m$ and two central classes adjacent to the median in case of even $m$ of minimum length $\delta$. The length of the classes is progressively increased by a factor of $(1 + 1/m)$ when moving towards the tails. Hence the classes adjacent to the central one(s) have length $\delta(1 + 1/m)$, the following classes have length $\delta(1 + 1/m)^2$, then the length increases to $\delta(1 + 1/m)^3$, and so forth. The minimum length $\delta$ is determined as follows

$$\begin{aligned} \delta \left\{ 1 + 2 \sum_{r=1}^{(m-1)/2} \left(1 + \frac{1}{m}\right)^r \right\} &= R, \qquad \textit{for odd } m \ , \\ 2\delta \left\{ 1 + \sum_{r=1}^{m/2-1} \left(1 + \frac{1}{m}\right)^r \right\} &= R, \qquad \textit{for even } m \ . \end{aligned} \tag{S.1.1}$$

The threshold $\alpha_j^{EP}$, which generate classes with equal probability, are given by the quantiles of $Y^*$ corresponding to the probability $1/m$, $2/m$, ..., $(m-1)/m$ obtained from the simulated distribution.

The thresholds $\alpha_j^{DP}$ yield classes with larger probability at the center and decreasing probability, by a factor of $(1 - 1/m)$, when moving from the center towards the extremes. Let $\pi_C$ be the probability of the central class in case of odd $m$ and of the two central classes in case of even $m$. The probability of the adjacent classes is $\pi_C(1 - 1/m)$, the probability of the following classes is $\pi_C(1 - 1/m)^2$, then the probability decreases to $\pi_C(1 - 1/m)^3$, and so forth. The probability $\pi_C$ of the central classes is determined as follows

$$\begin{aligned} \pi_C \left\{ 1 + 2 \sum_{r=1}^{(m-1)/2} \left(1 - \frac{1}{m}\right)^r \right\} &= 1, \qquad \textit{for odd } m \ , \\ 2\pi_C \left\{ 1 + \sum_{r=1}^{m/2-1} \left(1 - \frac{1}{m}\right)^r \right\} &= 1, \qquad \textit{for even } m \ . \end{aligned} \tag{S.1.2}$$

The thresholds are given by the corresponding quantiles from the simulated distribution of $Y^*$.

## S.1.2 Asymptotic Efficiency

Figures S.1, S.2 and S.3 compare the efficiency of the estimators of the regression coefficients obtained with the four sets of thresholds $\alpha_j^{ED}$, $\alpha_j^{ID}$, $\alpha_j^{EP}$ and $\alpha_j^{DP}$.

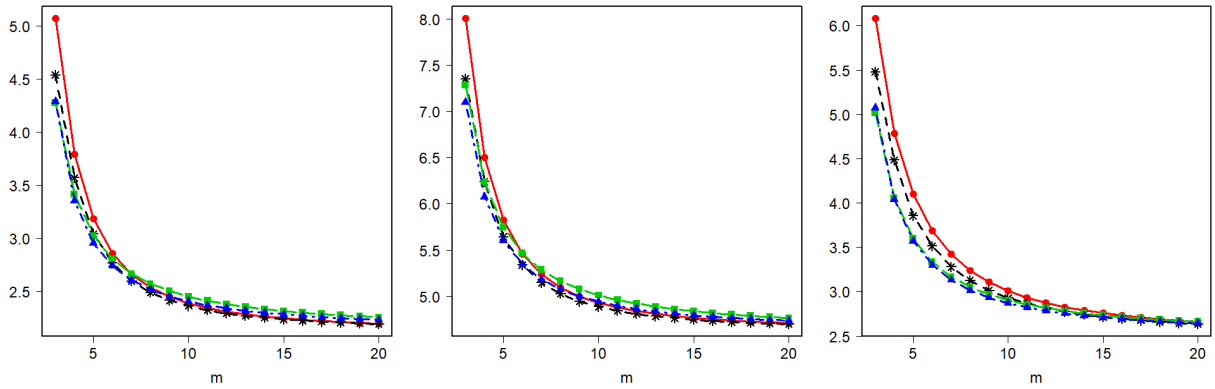


Figure S.1: Asymptotic efficiency in Model 1 with the probit link (left panel), the logit link (central panel) and the complementary log-log link (right panel) when $m$ varies with the $\alpha^{ED}$ thresholds (continuous line with circles), the $\alpha^{ID}$ thresholds (short-dashed line with stars), the $\alpha^{EP}$ thresholds (long-dashed line with squares) and the $\alpha^{DP}$ thresholds (dot-dashed line with triangles)

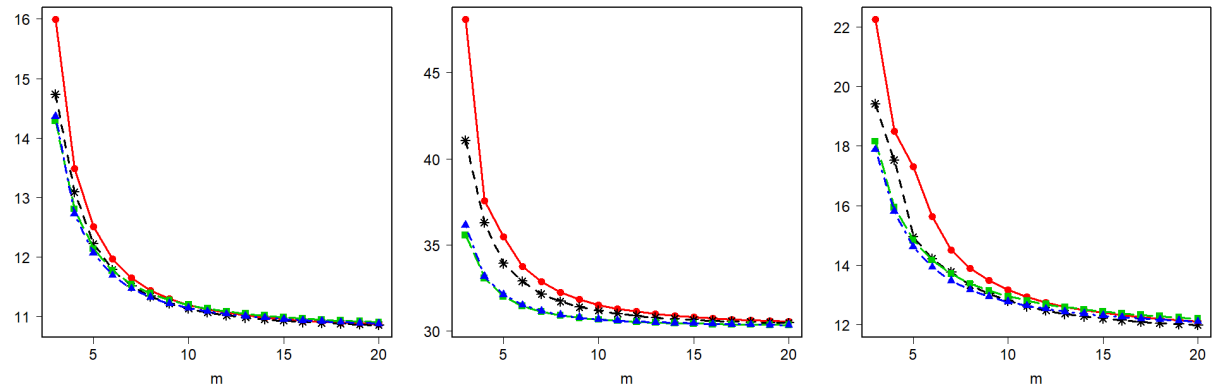


Figure S.2: Asymptotic efficiency in Model 2 with the probit link (left panel), the logit link (central panel) and the complementary log-log link (right panel) when $m$ varies with the $\alpha^{ED}$ thresholds (continuous line with circles), the $\alpha^{ID}$ thresholds (short-dashed line with stars), the $\alpha^{EP}$ thresholds (long-dashed line with squares) and the $\alpha^{DP}$ thresholds (dot-dashed line with triangles)

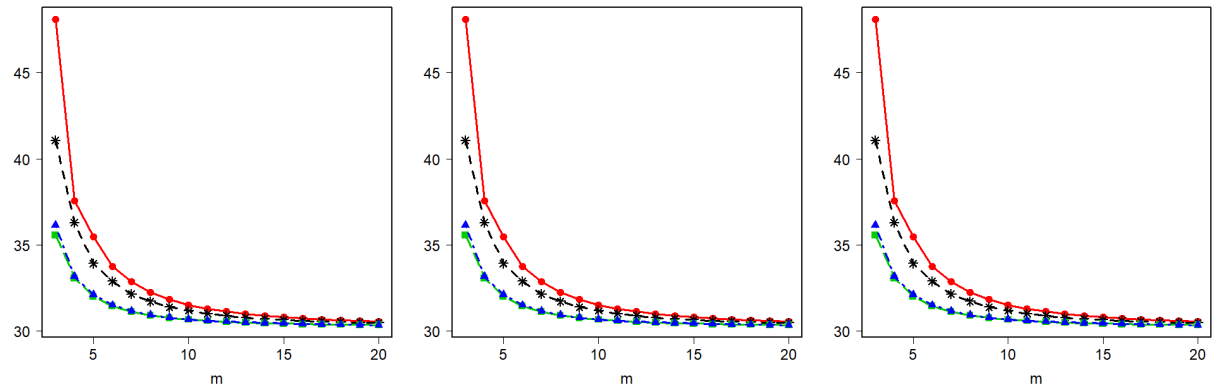


Figure S.3: Asymptotic efficiency in Model 3 with the probit link (left panel), the logit link (central panel) and the complementary log-log link (right panel) when $m$ varies with the $\alpha^{ED}$ thresholds (continuous line with circles), the $\alpha^{ID}$ thresholds (short-dashed line with stars), the $\alpha^{EP}$ thresholds (long-dashed line with squares) and the $\alpha^{DP}$ thresholds (dot-dashed line with triangles)

### S.1.3 Hypothesis Testing

Tables S.1.1, S.1.2 and S.1.3 provide approximations to the power of the test analogue to that of Table 5.1 of the main paper for the $\alpha^{ID}$, the $\alpha^{EP}$ and the $\alpha^{DP}$ thresholds.

Table S.1.1: Power of the test on $\beta_3 = 0$ in Model 3, at the 5% significance level, when $m$ varies and $n = 250, 500$, with the $\alpha^{ID}$ thresholds

| | $n = 250$ | | | $n = 500$ | | |
|---|---|---|---|---|---|---|
| $m$ | Probit | Logit | C. log-log | Probit | Logit | C. log-log |
| 3 | 0.65 | 0.44 | 0.63 | 0.92 | 0.73 | 0.90 |
| 4 | 0.84 | 0.58 | 0.81 | 0.99 | 0.87 | 0.98 |
| 5 | 0.91 | 0.66 | 0.88 | 1.00 | 0.92 | 0.99 |
| 6 | 0.95 | 0.71 | 0.93 | 1.00 | 0.95 | 1.00 |
| 7 | 0.97 | 0.75 | 0.95 | 1.00 | 0.96 | 1.00 |
| 8 | 0.98 | 0.77 | 0.97 | 1.00 | 0.97 | 1.00 |
| 9 | 0.99 | 0.79 | 0.98 | 1.00 | 0.97 | 1.00 |
| 10 | 0.99 | 0.80 | 0.98 | 1.00 | 0.98 | 1.00 |
| 12 | 0.99 | 0.82 | 0.99 | 1.00 | 0.98 | 1.00 |
| 15 | 1.00 | 0.83 | 0.99 | 1.00 | 0.99 | 1.00 |
| 20 | 1.00 | 0.84 | 1.00 | 1.00 | 0.99 | 1.00 |

Table S.1.2: Power of the test on $\beta_3 = 0$ in Model 3, at the 5% significance level, when $m$ varies and $n = 250, 500$, with the $\alpha^{EP}$ thresholds

| | $n = 250$ | | | $n = 500$ | | |
|---|---|---|---|---|---|---|
| $m$ | Probit | Logit | C. log-log | Probit | Logit | C. log-log |
| 3 | 0.44 | 0.31 | 0.41 | 0.72 | 0.55 | 0.69 |
| 4 | 0.67 | 0.44 | 0.63 | 0.92 | 0.73 | 0.90 |
| 5 | 0.80 | 0.53 | 0.76 | 0.98 | 0.82 | 0.97 |
| 6 | 0.87 | 0.59 | 0.84 | 0.99 | 0.87 | 0.99 |
| 7 | 0.92 | 0.64 | 0.89 | 1.00 | 0.91 | 0.99 |
| 8 | 0.94 | 0.67 | 0.92 | 1.00 | 0.93 | 1.00 |
| 9 | 0.96 | 0.70 | 0.94 | 1.00 | 0.94 | 1.00 |
| 10 | 0.97 | 0.72 | 0.96 | 1.00 | 0.95 | 1.00 |
| 12 | 0.98 | 0.75 | 0.97 | 1.00 | 0.96 | 1.00 |
| 15 | 0.99 | 0.77 | 0.98 | 1.00 | 0.97 | 1.00 |
| 20 | 0.99 | 0.80 | 0.99 | 1.00 | 0.98 | 1.00 |

Table S.1.3: Power of the test on $\beta_3 = 0$ in Model 3, at the 5% significance level, when $m$ varies and $n = 250, 500$, with the $\alpha^{DP}$ thresholds

| | $n = 250$ | | | $n = 500$ | | |
|---|---|---|---|---|---|---|
| $m$ | Probit | Logit | C. log-log | Probit | Logit | C. log-log |
| 3 | 0.55 | 0.38 | 0.52 | 0.84 | 0.64 | 0.81 |
| 4 | 0.74 | 0.49 | 0.70 | 0.96 | 0.78 | 0.94 |
| 5 | 0.86 | 0.59 | 0.82 | 0.99 | 0.87 | 0.98 |
| 6 | 0.91 | 0.64 | 0.88 | 1.00 | 0.91 | 0.99 |
| 7 | 0.94 | 0.68 | 0.92 | 1.00 | 0.93 | 1.00 |
| 8 | 0.96 | 0.71 | 0.95 | 1.00 | 0.94 | 1.00 |
| 9 | 0.97 | 0.73 | 0.96 | 1.00 | 0.95 | 1.00 |
| 10 | 0.98 | 0.75 | 0.97 | 1.00 | 0.96 | 1.00 |
| 12 | 0.99 | 0.77 | 0.98 | 1.00 | 0.97 | 1.00 |
| 15 | 0.99 | 0.79 | 0.99 | 1.00 | 0.98 | 1.00 |
| 20 | 0.99 | 0.81 | 0.99 | 1.00 | 0.98 | 1.00 |

### S.1.4 Merging categories

Tables S.1.4, S.1.5 and S.1.6 report analyses analogue to those of Table 6.3 of the main paper on halving the number of categories for the $\alpha^{ID}$, the $\alpha^{EP}$ and the $\alpha^{DP}$ thresholds.

Table S.1.4: Efficiency ratio between $\hat{\boldsymbol{\beta}}_{m/2}$ and $\hat{\boldsymbol{\beta}}_m$ produced by halving the number of categories with the $\alpha^{IC}$ thresholds.

| | Model 1 | | | Model 2 | | | Model 3 | | |
|---|---|---|---|---|---|---|---|---|---|
| $m$ | Probit | Logit | C.log-log | Probit | Logit | C.log-log | Probit | Logit | C.log-log |
| 4 | 2.03 | 1.72 | 1.79 | 1.47 | 1.22 | 1.46 | 5.40 | 3.67 | 5.21 |
| 6 | 1.64 | 1.38 | 1.56 | 1.25 | 1.25 | 1.37 | 2.12 | 1.75 | 1.97 |
| 8 | 1.43 | 1.24 | 1.44 | 1.15 | 1.14 | 1.31 | 1.91 | 1.50 | 1.77 |
| 10 | 1.29 | 1.15 | 1.32 | 1.10 | 1.09 | 1.17 | 1.71 | 1.35 | 1.63 |
| 12 | 1.21 | 1.11 | 1.25 | 1.07 | 1.06 | 1.14 | 1.55 | 1.25 | 1.52 |
| 20 | 1.08 | 1.04 | 1.11 | 1.03 | 1.02 | 1.07 | 1.23 | 1.10 | 1.27 |

Table S.1.5: Efficiency ratio between $\hat{\boldsymbol{\beta}}_{m/2}$ and $\hat{\boldsymbol{\beta}}_m$ produced by halving the number of categories with the $\alpha^{EP}$ thresholds.

| | Model 1 | | | Model 2 | | | Model 3 | | |
|---|---|---|---|---|---|---|---|---|---|
| $m$ | Probit | Logit | C.log-log | Probit | Logit | C.log-log | Probit | Logit | C.log-log |
| 4 | 2.12 | 1.72 | 1.98 | 1.50 | 1.34 | 1.60 | 5.17 | 3.22 | 4.75 |
| 6 | 1.53 | 1.33 | 1.50 | 1.21 | 1.13 | 1.28 | 2.41 | 1.87 | 2.32 |
| 8 | 1.33 | 1.21 | 1.33 | 1.12 | 1.07 | 1.19 | 1.86 | 1.53 | 1.84 |
| 10 | 1.23 | 1.15 | 1.24 | 1.08 | 1.04 | 1.15 | 1.61 | 1.38 | 1.62 |
| 12 | 1.18 | 1.11 | 1.18 | 1.06 | 1.03 | 1.12 | 1.47 | 1.29 | 1.49 |
| 20 | 1.09 | 1.05 | 1.09 | 1.03 | 1.01 | 1.06 | 1.23 | 1.14 | 1.26 |

Table S.1.6: Efficiency ratio between $\hat{\boldsymbol{\beta}}_{m/2}$ and $\hat{\boldsymbol{\beta}}_m$ produced by halving the number of categories with the $\alpha^{DP}$ thresholds.

| | Model 1 | | | Model 2 | | | Model 3 | | |
|---|---|---|---|---|---|---|---|---|---|
| $m$ | Probit | Logit | C.log-log | Probit | Logit | C.log-log | Probit | Logit | C.log-log |
| 4 | 2.16 | 1.77 | 1.99 | 1.51 | 1.34 | 1.62 | 5.50 | 3.47 | 5.11 |
| 6 | 1.56 | 1.33 | 1.54 | 1.23 | 1.15 | 1.28 | 2.27 | 1.77 | 2.14 |
| 8 | 1.33 | 1.20 | 1.34 | 1.12 | 1.07 | 1.20 | 1.84 | 1.49 | 1.80 |
| 10 | 1.23 | 1.13 | 1.24 | 1.08 | 1.05 | 1.14 | 1.59 | 1.33 | 1.58 |
| 12 | 1.17 | 1.10 | 1.19 | 1.06 | 1.03 | 1.11 | 1.45 | 1.25 | 1.47 |
| 20 | 1.08 | 1.04 | 1.09 | 1.02 | 1.01 | 1.06 | 1.21 | 1.12 | 1.24 |

Tables S.1.7, S.1.8 and S.1.9 report analyses analogue to those of Table 6.4 of the main paper on dichothomization for the $\alpha^{ID}$, the $\alpha^{EP}$ and the $\alpha^{DP}$ thresholds.

Table S.1.7: Efficiency ratio between $\hat{\boldsymbol{\beta}}_2$ and $\hat{\boldsymbol{\beta}}_m$ produced by dichotomization with the $\alpha^{ID}$ thresholds

| | Model 1 | | | Model 2 | | | Model 3 | | |
|---|---|---|---|---|---|---|---|---|---|
| $m$ | Probit | Logit | C.log-log | Probit | Logit | C.log-log | Probit | Logit | C.log-log |
| 4 | 2.03 | 1.72 | 1.79 | 1.47 | 1.22 | 1.46 | 5.40 | 3.67 | 5.21 |
| 6 | 2.61 | 2.01 | 2.29 | 1.64 | 1.35 | 1.80 | 8.19 | 4.84 | 7.48 |
| 8 | 2.90 | 2.14 | 2.57 | 1.70 | 1.40 | 1.92 | 10.30 | 5.48 | 9.19 |
| 10 | 3.06 | 2.20 | 2.74 | 1.73 | 1.42 | 1.99 | 11.73 | 5.85 | 10.44 |
| 12 | 3.15 | 2.23 | 2.85 | 1.75 | 1.44 | 2.05 | 12.69 | 6.06 | 11.35 |
| 20 | 3.30 | 2.29 | 3.05 | 1.78 | 1.46 | 2.13 | 14.42 | 6.42 | 13.23 |

Table S.1.8: Efficiency ratio between $\hat{\boldsymbol{\beta}}_2$ and $\hat{\boldsymbol{\beta}}_m$ produced by dichotomization with the $\alpha^{EP}$ thresholds

| | Model 1 | | | Model 2 | | | Model 3 | | |
|---|---|---|---|---|---|---|---|---|---|
| $m$ | Probit | Logit | C.log-log | Probit | Logit | C.log-log | Probit | Logit | C.log-log |
| 4 | 2.12 | 1.72 | 1.98 | 1.50 | 1.34 | 1.60 | 5.17 | 3.22 | 4.75 |
| 6 | 2.58 | 1.97 | 2.41 | 1.64 | 1.41 | 1.80 | 7.90 | 4.32 | 7.17 |
| 8 | 2.81 | 2.08 | 2.63 | 1.69 | 1.44 | 1.91 | 9.63 | 4.93 | 8.77 |
| 10 | 2.95 | 2.14 | 2.77 | 1.72 | 1.45 | 1.97 | 10.79 | 5.30 | 9.88 |
| 12 | 3.04 | 2.18 | 2.85 | 1.74 | 1.45 | 2.01 | 11.61 | 5.56 | 10.68 |
| 20 | 3.20 | 2.25 | 3.02 | 1.77 | 1.46 | 2.09 | 13.32 | 6.06 | 12.44 |

Table S.1.9: Efficiency ratio between $\hat{\boldsymbol{\beta}}_2$ and $\hat{\boldsymbol{\beta}}_m$ produced by dichotomization with the $\alpha^{DP}$ thresholds

| | Model 1 | | | Model 2 | | | Model 3 | | |
|---|---|---|---|---|---|---|---|---|---|
| $m$ | Probit | Logit | C.log-log | Probit | Logit | C.log-log | Probit | Logit | C.log-log |
| 4 | 2.16 | 1.77 | 1.99 | 1.51 | 1.34 | 1.62 | 5.50 | 3.47 | 5.11 |
| 6 | 2.64 | 2.01 | 2.43 | 1.65 | 1.41 | 1.83 | 8.34 | 4.58 | 7.56 |
| 8 | 2.87 | 2.12 | 2.66 | 1.70 | 1.44 | 1.94 | 10.13 | 5.17 | 9.18 |
| 10 | 3.00 | 2.17 | 2.80 | 1.73 | 1.45 | 2.00 | 11.30 | 5.52 | 10.29 |
| 12 | 3.09 | 2.21 | 2.88 | 1.75 | 1.45 | 2.04 | 12.10 | 5.74 | 11.09 |
| 20 | 3.24 | 2.27 | 3.04 | 1.77 | 1.46 | 2.11 | 13.71 | 6.18 | 12.79 |

### S.1.5 Choice of $m$ for given $n$

Tables S.1.10, S.1.11 and S.1.12 report analyses analogue to those of Table 7.1 of the main paper on the percentage of simulated samples with a number of categories smaller than $m$ (in Model 3 with the Logit link), when $n$ and $m$ vary, for the $\alpha^{ID}$, the $\alpha^{EP}$ and the $\alpha^{DP}$ thresholds.

Table S.1.10: Percentage of simulated samples with a number of categories smaller than $m$ in Model 3 with the Logit link, when $n$ and $m$ vary, with the $\alpha^{ID}$ thresholds

| | | | | | $m$ | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $n$ | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| 50 | 0.01 | 0.82 | 3.41 | 8.25 | 14.23 | 22.37 | 30.51 | 39.84 | 49.21 |
| 100 | 0.00 | 0.00 | 0.03 | 0.15 | 0.72 | 1.86 | 3.39 | 5.75 | 8.31 |
| 200 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.03 | 0.17 | 0.41 | 0.76 |
| 300 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.02 | 0.06 | 0.11 |
| 400 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.04 | 0.10 |
| 500 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |

Table S.1.11: Percentage of simulated samples with a number of categories smaller than $m$ in Model 3 with the Logit link, when $n$ and $m$ vary, with the $\alpha^{EP}$ thresholds

| | | | | | $m$ | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $n$ | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| 50 | 0.00 | 0.06 | 0.33 | 3.02 | 4.91 | 8.06 | 15.00 | 22.73 | 45.26 |
| 100 | 0.00 | 0.00 | 0.00 | 0.00 | 0.03 | 0.09 | 0.35 | 1.39 | 2.42 |
| 200 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 300 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 400 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 500 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |

Table S.1.12: Percentage of simulated samples with a number of categories smaller than $m$ in Model 3 with the Logit link, when $n$ and $m$ vary, with the $\alpha^{DP}$ thresholds

| | | | | | $m$ | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $n$ | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| 50 | 0.01 | 0.06 | 0.40 | 2.22 | 5.15 | 10.16 | 16.90 | 39.20 | 52.54 |
| 100 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.12 | 0.75 | 0.72 | 0.88 |
| 200 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 |
| 300 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 400 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 500 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |

Tables S.1.13, S.1.14 and S.1.15 report analyses analogue to those of Table 7.2 of the main paper on the efficiency of the estimators (in Model 3 with the Logit link), when $n$ and $m$ vary, for the $\alpha^{ID}$, the $\alpha^{EP}$ and the $\alpha^{DP}$ thresholds.

Table S.1.13: Simulated efficiency $\left(100 \times \sum_k MSE(\hat{\beta}_k)\right)$ in Model 3 with the Logit link, when $n$ and $m$ vary with the $\alpha^{ID}$ thresholds

| | | | | | $m$ | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $n$ | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| 50 | 1835.07 | 408.86 | 280.22 | 215.09 | 189.58 | 175.70 | 165.87 | 161.66 | 155.05 |
| 100 | 142.38 | 104.39 | 82.35 | 72.30 | 66.27 | 62.23 | 59.89 | 57.46 | 56.72 |
| 200 | 56.84 | 43.46 | 35.15 | 31.07 | 28.62 | 27.38 | 26.12 | 25.26 | 24.89 |
| 300 | 34.97 | 26.74 | 22.26 | 19.63 | 17.92 | 17.12 | 16.48 | 16.12 | 15.76 |
| 400 | 25.15 | 19.54 | 16.03 | 14.10 | 13.20 | 12.44 | 11.89 | 11.59 | 11.37 |
| 500 | 20.16 | 15.73 | 12.87 | 11.38 | 10.52 | 10.01 | 9.57 | 9.33 | 9.09 |

Table S.1.14: Simulated efficiency $\left(100 \times \sum_k MSE(\hat{\beta}_k)\right)$ in Model 3 with the Logit link, when $n$ and $m$ vary with the $\alpha^{ED}$ thresholds

| | | | | | $m$ | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $n$ | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| 50 | 2023.32 | 697.76 | 287.34 | 238.69 | 215.23 | 201.93 | 196.55 | 193.43 | 192.19 |
| 100 | 155.65 | 107.44 | 92.10 | 80.00 | 72.90 | 69.31 | 66.88 | 62.97 | 60.79 |
| 200 | 62.92 | 44.88 | 37.85 | 33.08 | 30.63 | 29.10 | 28.03 | 27.58 | 26.88 |
| 300 | 40.04 | 28.69 | 23.67 | 21.08 | 19.50 | 18.68 | 18.00 | 17.46 | 17.25 |
| 400 | 29.44 | 20.08 | 16.86 | 15.38 | 14.24 | 13.58 | 13.01 | 12.49 | 12.18 |
| 500 | 24.23 | 16.85 | 13.80 | 12.46 | 11.39 | 10.75 | 10.20 | 9.94 | 9.75 |

Table S.1.15: Simulated efficiency $\left(100 \times \sum_k MSE(\hat{\beta}_k)\right)$ in Model 3 with the Logit link, when $n$ and $m$ vary with the $\alpha^{DP}$ thresholds

| | | | | | $m$ | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $n$ | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| 50 | 1159.88 | 338.99 | 248.80 | 220.64 | 204.97 | 199.31 | 194.96 | 193.30 | 192.37 |
| 100 | 135.18 | 103.38 | 83.96 | 75.20 | 68.91 | 64.70 | 61.55 | 59.85 | 58.45 |
| 200 | 55.88 | 42.43 | 34.81 | 31.47 | 29.20 | 28.18 | 27.27 | 26.56 | 26.04 |
| 300 | 36.67 | 26.12 | 22.10 | 19.93 | 18.79 | 17.97 | 17.47 | 17.01 | 16.48 |
| 400 | 25.39 | 18.92 | 16.20 | 14.57 | 13.68 | 12.82 | 12.38 | 12.14 | 11.80 |
| 500 | 21.16 | 15.60 | 13.01 | 11.66 | 10.66 | 10.23 | 9.94 | 9.72 | 9.48 |

## S.2 Merging categories

This section supplies supplementary tables for Section 6 of the main paper.

Tables S.2.1, S.2.2, S.2.3, S.2.4, S.2.5, along with Table 6.1 of the main paper, show the loss in efficiency produced by merging the extreme categories on both sides in the three models with the Probit and the Logit link.

Table S.2.1: Efficiency loss produced by merging of the extreme categories on both sides in Model 1 with the Probit link

| $m$ | $j=1$ | $j=2$ | $j=m-1$ | $j=m$ | $Var(\hat{\beta}_m)$ | $Var(\hat{\beta}_{m-2})$ | Efficiency Ratio |
|---|---|---|---|---|---|---|---|
| 5 | 0.061 | 0.241 | 0.243 | 0.062 | 3.185 | 4.254 | 1.336 |
| 6 | 0.043 | 0.152 | 0.153 | 0.043 | 2.857 | 3.353 | 1.174 |
| 7 | 0.033 | 0.102 | 0.102 | 0.033 | 2.662 | 2.942 | 1.105 |
| 8 | 0.027 | 0.072 | 0.073 | 0.027 | 2.537 | 2.715 | 1.070 |
| 9 | 0.023 | 0.053 | 0.054 | 0.023 | 2.452 | 2.575 | 1.050 |
| 10 | 0.020 | 0.041 | 0.042 | 0.020 | 2.391 | 2.482 | 1.038 |
| 12 | 0.016 | 0.027 | 0.027 | 0.016 | 2.313 | 2.368 | 1.024 |
| 15 | 0.013 | 0.017 | 0.017 | 0.013 | 2.249 | 2.281 | 1.014 |
| 20 | 0.010 | 0.009 | 0.010 | 0.010 | 2.199 | 2.216 | 1.008 |

Table S.2.2: Efficiency loss produced by merging of the extreme categories on both sides in Model 2 with the Probit link

| $m$ | $j=1$ | $j=2$ | $j=m-1$ | $j=m$ | $Var(\hat{\beta}_m)$ | $Var(\hat{\beta}_{m-2})$ | Efficiency Ratio |
|---|---|---|---|---|---|---|---|
| 5 | 0.072 | 0.253 | 0.243 | 0.067 | 12.512 | 14.213 | 1.136 |
| 6 | 0.051 | 0.168 | 0.159 | 0.047 | 11.971 | 12.710 | 1.062 |
| 7 | 0.038 | 0.117 | 0.109 | 0.036 | 11.650 | 12.043 | 1.034 |
| 8 | 0.031 | 0.084 | 0.079 | 0.029 | 11.441 | 11.680 | 1.021 |
| 9 | 0.026 | 0.064 | 0.059 | 0.024 | 11.299 | 11.457 | 1.014 |
| 10 | 0.022 | 0.050 | 0.046 | 0.021 | 11.198 | 11.309 | 1.010 |
| 12 | 0.018 | 0.033 | 0.030 | 0.017 | 11.066 | 11.129 | 1.006 |
| 15 | 0.014 | 0.020 | 0.019 | 0.014 | 10.958 | 10.992 | 1.003 |
| 20 | 0.011 | 0.011 | 0.010 | 0.011 | 10.874 | 10.891 | 1.002 |

Table S.2.3: Efficiency loss produced by merging of the extreme categories on both sides in Model 2 with the Logit link

| $m$ | $j=1$ | $j=2$ | $j=m-1$ | $j=m$ | $Var(\hat{\beta}_m)$ | $Var(\hat{\beta}_{m-2})$ | Efficiency Ratio |
|---|---|---|---|---|---|---|---|
| 5 | 0.044 | 0.231 | 0.229 | 0.044 | 35.460 | 36.513 | 1.030 |
| 6 | 0.031 | 0.131 | 0.129 | 0.031 | 33.763 | 34.002 | 1.007 |
| 7 | 0.024 | 0.081 | 0.080 | 0.024 | 32.861 | 32.934 | 1.002 |
| 8 | 0.020 | 0.054 | 0.054 | 0.020 | 32.238 | 32.266 | 1.001 |
| 9 | 0.017 | 0.039 | 0.038 | 0.017 | 31.822 | 31.835 | 1.000 |
| 10 | 0.015 | 0.029 | 0.029 | 0.015 | 31.520 | 31.527 | 1.000 |
| 12 | 0.013 | 0.019 | 0.018 | 0.013 | 31.128 | 31.131 | 1.000 |
| 15 | 0.010 | 0.011 | 0.011 | 0.011 | 30.807 | 30.808 | 1.000 |
| 20 | 0.009 | 0.006 | 0.006 | 0.009 | 30.557 | 30.557 | 1.000 |

Table S.2.4: Efficiency loss produced by merging of the extreme categories on both sides in Model 3 with the Probit link

| $m$ | $j=1$ | $j=2$ | $j=m-1$ | $j=m$ | $Var(\hat{\beta}_m)$ | $Var(\hat{\beta}_{m-2})$ | Efficiency Ratio |
|---|---|---|---|---|---|---|---|
| 5 | 0.069 | 0.279 | 0.204 | 0.050 | 38.397 | 69.911 | 1.821 |
| 6 | 0.048 | 0.179 | 0.124 | 0.036 | 31.888 | 44.601 | 1.399 |
| 7 | 0.036 | 0.120 | 0.081 | 0.028 | 27.712 | 34.441 | 1.243 |
| 8 | 0.029 | 0.085 | 0.057 | 0.023 | 24.932 | 29.059 | 1.166 |
| 9 | 0.024 | 0.063 | 0.043 | 0.020 | 23.011 | 25.796 | 1.121 |
| 10 | 0.021 | 0.048 | 0.033 | 0.017 | 21.636 | 23.647 | 1.093 |
| 12 | 0.017 | 0.031 | 0.022 | 0.014 | 20.945 | 21.053 | 1.005 |
| 15 | 0.013 | 0.019 | 0.013 | 0.012 | 19.120 | 19.087 | 0.998 |
| 20 | 0.011 | 0.011 | 0.008 | 0.009 | 16.951 | 17.642 | 1.041 |

Table S.2.5: Efficiency loss produced by merging of the extreme categories on both sides in Model 3 with the Logit link

| $m$ | $j=1$ | $j=2$ | $j=m-1$ | $j=m$ | $Var(\hat{\beta}_m)$ | $Var(\hat{\beta}_{m-2})$ | Efficiency Ratio |
|---|---|---|---|---|---|---|---|
| 5 | 0.064 | 0.269 | 0.212 | 0.051 | 60.633 | 101.623 | 1.676 |
| 6 | 0.045 | 0.169 | 0.129 | 0.036 | 53.938 | 71.277 | 1.321 |
| 7 | 0.034 | 0.112 | 0.084 | 0.028 | 49.867 | 59.210 | 1.187 |
| 8 | 0.027 | 0.079 | 0.059 | 0.023 | 47.217 | 52.994 | 1.122 |
| 9 | 0.023 | 0.058 | 0.044 | 0.020 | 45.398 | 49.308 | 1.086 |
| 10 | 0.020 | 0.044 | 0.034 | 0.017 | 44.096 | 46.917 | 1.064 |
| 12 | 0.016 | 0.029 | 0.022 | 0.014 | 42.398 | 44.075 | 1.040 |
| 15 | 0.013 | 0.017 | 0.014 | 0.012 | 41.006 | 41.954 | 1.023 |
| 20 | 0.010 | 0.010 | 0.008 | 0.010 | 39.916 | 40.413 | 1.012 |

Tables S.2.6 and S.2.7 along with Table 6.2 of the main paper show the efficiency loss produced by merging the lowest categories in the three models with the Complementary log-log link.

Table S.2.6: Efficiency loss produced by merging the lowest categories in Model 2 with the Complementary log-log link

| $m$ | $j=1$ | $j=2$ | $Var(\hat{\beta}_m)$ | $Var(\hat{\beta}_{m-1})$ | Efficiency Ratio |
|---|---|---|---|---|---|
| 4 | 0.037 | 0.195 | 18.500 | 18.512 | 1.001 |
| 5 | 0.025 | 0.090 | 17.303 | 17.304 | 1.000 |
| 6 | 0.019 | 0.051 | 15.626 | 15.627 | 1.000 |
| 7 | 0.016 | 0.033 | 14.502 | 14.502 | 1.000 |
| 8 | 0.014 | 0.023 | 13.888 | 13.888 | 1.000 |
| 9 | 0.012 | 0.017 | 13.489 | 13.489 | 1.000 |
| 10 | 0.011 | 0.013 | 13.179 | 13.179 | 1.000 |
| 12 | 0.010 | 0.009 | 12.753 | 12.753 | 1.000 |
| 15 | 0.008 | 0.006 | 12.398 | 12.398 | 1.000 |
| 20 | 0.007 | 0.004 | 12.110 | 12.110 | 1.000 |

Table S.2.7: Efficiency loss produced by merging the lowest categories in Model 3 with the Complementary log-log link

| $m$ | $j=1$ | $j=2$ | $Var(\hat{\beta}_m)$ | $Var(\hat{\beta}_{m-1})$ | Efficiency Ratio |
|---|---|---|---|---|---|
| 4 | 0.099 | 0.426 | 51.400 | 88.621 | 1.724 |
| 5 | 0.060 | 0.257 | 41.465 | 53.786 | 1.297 |
| 6 | 0.042 | 0.159 | 35.356 | 41.092 | 1.162 |
| 7 | 0.032 | 0.105 | 31.327 | 34.506 | 1.101 |
| 8 | 0.026 | 0.073 | 28.519 | 30.490 | 1.069 |
| 9 | 0.022 | 0.054 | 26.479 | 27.802 | 1.050 |
| 10 | 0.019 | 0.041 | 24.947 | 25.889 | 1.038 |
| 12 | 0.015 | 0.027 | 22.831 | 23.375 | 1.024 |
| 15 | 0.012 | 0.016 | 20.960 | 21.257 | 1.014 |
| 20 | 0.010 | 0.009 | 20.304 | 20.467 | 1.008 |

## S.3 Choice of $m$ for given $n$

This section supplies supplementary tables for Section 7 of the main paper.

Tables S.3.1, S.3.2, S.3.3, S.3.4, S.3.5, S.3.6, S.3.7 and S.3.8, along with Table 7.1 of the main paper, illustrate the percentage of simulated samples with a number of observed categories smaller than $m$ for the three models and the three links.

Table S.3.1: Percentage of simulated samples with a number of categories smaller than $m$ in Model 1 with the Probit link, when $n$ and $m$ vary

| $n$ | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|
| 50 | 0.00 | 1.11 | 8.53 | 20.99 | 34.55 | 48.31 | 60.61 | 71.22 | 78.99 |
| 100 | 0.00 | 0.01 | 0.35 | 2.16 | 6.20 | 12.19 | 19.78 | 27.12 | 35.98 |
| 200 | 0.00 | 0.00 | 0.00 | 0.03 | 0.25 | 0.68 | 1.94 | 3.68 | 5.62 |
| 300 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.04 | 0.23 | 0.51 | 1.01 |
| 400 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.01 | 0.13 | 0.25 |
| 500 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.02 |

Table S.3.2: Percentage of simulated samples with a number of categories smaller than $m$ in Model 1 with the Logit link, when $n$ and $m$ vary

| $n$ | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|
| 50 | 0.03 | 2.42 | 14.17 | 30.15 | 45.44 | 58.36 | 69.71 | 78.82 | 86.14 |
| 100 | 0.00 | 0.01 | 1.07 | 5.26 | 12.04 | 20.23 | 29.25 | 39.00 | 48.49 |
| 200 | 0.00 | 0.00 | 0.00 | 0.10 | 0.65 | 2.21 | 4.58 | 7.19 | 10.73 |
| 300 | 0.00 | 0.00 | 0.00 | 0.01 | 0.05 | 0.24 | 0.64 | 1.30 | 2.17 |
| 400 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.04 | 0.14 | 0.33 | 0.53 |
| 500 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.05 | 0.07 |

Table S.3.3: Percentage of simulated samples with a number of categories smaller than $m$ in Model 1 with the Complementary log-log link, when $n$ and $m$ vary

| $n$ | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|
| 50 | 0.17 | 5.55 | 17.14 | 30.42 | 44.73 | 57.73 | 68.87 | 78.39 | 85.49 |
| 100 | 0.00 | 0.32 | 2.68 | 7.20 | 13.21 | 20.11 | 28.56 | 38.01 | 48.02 |
| 200 | 0.00 | 0.00 | 0.08 | 0.49 | 1.37 | 2.71 | 4.65 | 7.56 | 11.14 |
| 300 | 0.00 | 0.00 | 0.00 | 0.03 | 0.10 | 0.43 | 0.99 | 1.83 | 3.05 |
| 400 | 0.00 | 0.00 | 0.00 | 0.00 | 0.04 | 0.04 | 0.18 | 0.41 | 0.73 |
| 500 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.07 | 0.12 | 0.26 |

Table S.3.4: Percentage of simulated samples with a number of categories smaller than $m$ in Model 2 with the Probit link, when $n$ and $m$ vary

| $n$ | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|
| 50 | 0.00 | 0.57 | 5.35 | 16.15 | 28.78 | 41.70 | 53.79 | 64.45 | 73.36 |
| 100 | 0.00 | 0.00 | 0.15 | 1.35 | 4.40 | 9.58 | 15.30 | 22.16 | 29.67 |
| 200 | 0.00 | 0.00 | 0.00 | 0.01 | 0.08 | 0.38 | 1.11 | 2.41 | 4.30 |
| 300 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.05 | 0.12 | 0.38 | 0.68 |
| 400 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.02 | 0.09 |
| 500 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.02 |

Table S.3.5: Percentage of simulated samples with a number of categories smaller than $m$ in Model 2 with the Logit link, when $n$ and $m$ vary

| | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|
| 50 | 0.04 | 4.42 | 19.92 | 37.24 | 51.69 | 65.19 | 76.06 | 84.65 | 90.52 |
| 100 | 0.00 | 0.05 | 2.40 | 8.95 | 17.14 | 26.12 | 35.74 | 45.45 | 55.71 |
| 200 | 0.00 | 0.00 | 0.03 | 0.33 | 1.64 | 3.68 | 6.35 | 9.57 | 13.91 |
| 300 | 0.00 | 0.00 | 0.00 | 0.04 | 0.18 | 0.50 | 1.25 | 2.23 | 3.49 |
| 400 | 0.00 | 0.00 | 0.00 | 0.00 | 0.02 | 0.10 | 0.25 | 0.48 | 0.89 |
| 500 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.04 | 0.06 | 0.14 | 0.23 |

Table S.3.6: Percentage of simulated samples with a number of categories smaller than $m$ in Model 2 with the Complementary log-log link, when $n$ and $m$ vary

| $n$ | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|
| 50 | 3.15 | 16.10 | 30.23 | 44.16 | 57.75 | 69.80 | 78.79 | 86.21 | 91.58 |
| 100 | 0.07 | 2.37 | 7.98 | 15.71 | 24.12 | 34.22 | 43.61 | 53.67 | 63.02 |
| 200 | 0.00 | 0.07 | 0.56 | 1.98 | 4.22 | 7.29 | 10.94 | 16.14 | 22.46 |
| 300 | 0.00 | 0.00 | 0.11 | 0.43 | 0.87 | 1.65 | 2.78 | 4.75 | 7.67 |
| 400 | 0.00 | 0.00 | 0.00 | 0.04 | 0.14 | 0.42 | 0.83 | 1.53 | 2.62 |
| 500 | 0.00 | 0.00 | 0.00 | 0.01 | 0.07 | 0.12 | 0.26 | 0.38 | 0.89 |

Table S.3.7: Percentage of simulated samples with a number of categories smaller than $m$ in Model 3 with the Probit link, when $n$ and $m$ vary

| $n$ | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|
| 50 | 0.05 | 1.68 | 10.17 | 23.39 | 37.57 | 50.86 | 63.08 | 73.37 | 81.38 |
| 100 | 0.00 | 0.01 | 0.71 | 3.49 | 8.32 | 14.57 | 21.71 | 30.28 | 39.17 |
| 200 | 0.00 | 0.00 | 0.00 | 0.07 | 0.40 | 1.19 | 2.68 | 4.75 | 7.24 |
| 300 | 0.00 | 0.00 | 0.00 | 0.00 | 0.04 | 0.13 | 0.36 | 0.68 | 1.31 |
| 400 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.02 | 0.05 | 0.08 | 0.26 |
| 500 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.04 |

Table S.3.8: Percentage of simulated samples with a number of categories smaller than $m$ in Model 3 with the Complementary log-log link, when $n$ and $m$ vary

| $n$ | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|
| 50 | 0.00 | 1.85 | 10.76 | 25.01 | 39.41 | 52.83 | 64.94 | 75.06 | 82.95 |
| 100 | 0.00 | 0.01 | 0.64 | 3.40 | 8.51 | 15.43 | 22.78 | 30.56 | 38.93 |
| 200 | 0.00 | 0.00 | 0.00 | 0.07 | 0.40 | 1.25 | 2.68 | 4.44 | 6.96 |
| 300 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.12 | 0.33 | 0.83 | 1.57 |
| 400 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.02 | 0.09 | 0.21 |
| 500 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.02 | 0.02 |

Tables S.3.9, S.3.10, S.3.11, S.3.12, S.3.13, S.3.14, S.3.15 and S.3.16, along with Table 7.2 of the main paper, show the simulated efficiency of the estimators when $n$ and $m$ vary, for the three models and the three links.

Table S.3.9: Simulated efficiency ($100 \times MSE$) in Model 1 with the Probit link, when $n$ and $m$ vary

| $n$ | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|
| 50 | 24.37 | 14.48 | 11.48 | 9.73 | 9.09 | 8.34 | 7.96 | 7.84 | 7.48 |
| 100 | 7.16 | 5.09 | 4.15 | 3.73 | 3.52 | 3.27 | 3.17 | 3.07 | 3.04 |
| 200 | 2.87 | 2.15 | 1.79 | 1.58 | 1.49 | 1.41 | 1.37 | 1.31 | 1.32 |
| 300 | 1.91 | 1.39 | 1.17 | 1.04 | 0.97 | 0.92 | 0.89 | 0.87 | 0.86 |
| 400 | 1.33 | 1.01 | 0.84 | 0.76 | 0.71 | 0.66 | 0.64 | 0.63 | 0.62 |
| 500 | 1.08 | 0.80 | 0.67 | 0.60 | 0.55 | 0.53 | 0.51 | 0.50 | 0.48 |

Table S.3.10: Simulated efficiency ($100 \times MSE$) in Model 1 with the Logit link, when $n$ and $m$ vary

| $n$ | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|
| 50 | 24.34 | 18.75 | 16.26 | 15.03 | 14.22 | 13.79 | 13.36 | 13.34 | 13.06 |
| 100 | 9.58 | 7.79 | 6.83 | 6.32 | 6.13 | 5.91 | 5.78 | 5.71 | 5.67 |
| 200 | 4.42 | 3.52 | 3.19 | 2.98 | 2.82 | 2.77 | 2.71 | 2.68 | 2.62 |
| 300 | 2.86 | 2.27 | 2.04 | 1.90 | 1.83 | 1.76 | 1.75 | 1.72 | 1.69 |
| 400 | 2.13 | 1.65 | 1.53 | 1.40 | 1.35 | 1.30 | 1.30 | 1.28 | 1.26 |
| 500 | 1.71 | 1.34 | 1.21 | 1.16 | 1.09 | 1.06 | 1.05 | 1.03 | 1.02 |

Table S.3.11: Simulated efficiency ($100 \times MSE$) in Model 1 with the Complementary log-log link, when $n$ and $m$ vary

| $n$ | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|
| 50 | 25.32 | 17.74 | 14.25 | 12.56 | 11.40 | 10.74 | 10.13 | 10.11 | 9.59 |
| 100 | 8.25 | 6.36 | 5.29 | 4.83 | 4.28 | 4.16 | 4.01 | 3.89 | 3.79 |
| 200 | 3.55 | 2.70 | 2.29 | 2.05 | 1.90 | 1.81 | 1.76 | 1.68 | 1.67 |
| 300 | 2.25 | 1.75 | 1.47 | 1.36 | 1.22 | 1.17 | 1.13 | 1.07 | 1.07 |
| 400 | 1.63 | 1.26 | 1.09 | 0.97 | 0.89 | 0.85 | 0.81 | 0.79 | 0.77 |
| 500 | 1.32 | 1.01 | 0.87 | 0.77 | 0.72 | 0.68 | 0.65 | 0.64 | 0.61 |

Table S.3.12: Simulated efficiency $\left(100 \times \sum_k MSE(\hat{\beta}_k)\right)$ in Model 2 with the Probit link, when $n$ and $m$ vary

| $n$ | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|
| 50 | 154.90 | 42.34 | 44.03 | 34.76 | 33.38 | 32.16 | 31.15 | 30.71 | 30.65 |
| 100 | 35.03 | 15.70 | 14.59 | 13.85 | 13.40 | 13.18 | 13.01 | 12.85 | 12.71 |
| 200 | 9.44 | 7.32 | 6.74 | 6.47 | 6.30 | 6.19 | 6.14 | 6.03 | 6.04 |
| 300 | 5.73 | 4.79 | 4.42 | 4.19 | 4.07 | 4.03 | 3.96 | 3.95 | 3.90 |
| 400 | 4.21 | 3.46 | 3.22 | 3.05 | 2.99 | 2.92 | 2.89 | 2.85 | 2.84 |
| 500 | 3.37 | 2.80 | 2.63 | 2.50 | 2.46 | 2.40 | 2.37 | 2.35 | 2.34 |

Table S.3.13: Simulated efficiency $\left(100 \times \sum_k MSE(\hat{\beta}_k)\right)$ in Model 2 with the Logit link, when $n$ and $m$ vary

| $n$ | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|
| 50 | 336.51 | 103.38 | 100.11 | 85.31 | 82.46 | 81.05 | 79.18 | 78.17 | 77.20 |
| 100 | 58.59 | 42.21 | 39.70 | 37.36 | 36.36 | 35.69 | 35.08 | 34.79 | 34.30 |
| 200 | 25.99 | 19.96 | 18.57 | 17.74 | 17.27 | 16.96 | 16.68 | 16.55 | 16.46 |
| 300 | 17.07 | 12.98 | 12.57 | 11.69 | 11.59 | 11.20 | 11.16 | 11.00 | 10.99 |
| 400 | 12.53 | 9.48 | 8.99 | 8.57 | 8.35 | 8.16 | 8.06 | 8.00 | 7.95 |
| 500 | 9.87 | 7.64 | 7.25 | 6.88 | 6.75 | 6.56 | 6.51 | 6.45 | 6.40 |

Table S.3.14: Simulated efficiency $\left(100 \times \sum_k MSE(\hat{\beta}_k)\right)$ in Model 2 with the Complementary log-log link, when $n$ and $m$ vary

| $n$ | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|
| 50 | 192.37 | 234.93 | 68.05 | 46.73 | 48.27 | 42.89 | 39.35 | 37.81 | 37.15 |
| 100 | 26.19 | 29.28 | 21.68 | 18.23 | 17.12 | 16.58 | 16.02 | 15.87 | 15.58 |
| 200 | 11.86 | 10.88 | 9.51 | 8.27 | 7.79 | 7.49 | 7.27 | 7.06 | 7.04 |
| 300 | 7.76 | 6.84 | 6.30 | 5.46 | 5.12 | 4.90 | 4.78 | 4.66 | 4.56 |
| 400 | 5.69 | 4.92 | 4.50 | 3.96 | 3.71 | 3.57 | 3.42 | 3.38 | 3.32 |
| 500 | 4.53 | 3.84 | 3.63 | 3.17 | 2.95 | 2.84 | 2.78 | 2.71 | 2.63 |

Table S.3.15: Simulated efficiency $\left(100 \times \sum_k MSE(\hat{\beta}_k)\right)$ in Model 3 with the Probit link, when $n$ and $m$ vary

| $n$ | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|
| 50 | 13108.17 | 2412.08 | 693.60 | 358.38 | 251.84 | 167.81 | 150.93 | 129.65 | 118.97 |
| 100 | 305.73 | 112.12 | 77.19 | 59.44 | 49.74 | 42.70 | 39.94 | 36.28 | 34.45 |
| 200 | 49.45 | 34.95 | 26.44 | 21.44 | 18.17 | 16.30 | 14.96 | 13.70 | 13.19 |
| 300 | 28.38 | 20.55 | 15.71 | 12.91 | 11.01 | 9.95 | 9.15 | 8.46 | 8.03 |
| 400 | 19.56 | 14.23 | 11.23 | 9.06 | 7.95 | 7.04 | 6.50 | 6.13 | 5.82 |
| 500 | 15.10 | 11.13 | 8.68 | 7.06 | 6.17 | 5.53 | 5.14 | 4.79 | 4.61 |

Table S.3.16: Simulated efficiency $\left(100 \times \sum_k MSE(\hat{\beta}_k)\right)$ in Model 3 with the Complementary log-log link, when $n$ and $m$ vary

| $n$ | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|
| 50 | 6441.78 | 1070.82 | 394.97 | 257.54 | 204.55 | 175.13 | 154.15 | 143.41 | 133.39 |
| 100 | 165.38 | 107.81 | 76.28 | 64.17 | 55.23 | 49.45 | 45.17 | 42.36 | 40.25 |
| 200 | 49.26 | 36.22 | 28.98 | 23.70 | 20.76 | 18.98 | 17.45 | 16.66 | 15.64 |
| 300 | 29.11 | 21.30 | 16.85 | 14.38 | 12.62 | 11.32 | 10.67 | 10.09 | 9.57 |
| 400 | 20.27 | 15.09 | 11.95 | 10.15 | 8.91 | 8.04 | 7.61 | 7.04 | 6.76 |
| 500 | 15.84 | 11.84 | 9.45 | 7.98 | 7.06 | 6.40 | 5.94 | 5.61 | 5.41 |