Supplementary Tables for Clark and Barr (2017): Root Mean Squared Errors of Relative Risk Estimators in Simulation Study when $n = 100$

Note that there are some very large root mean squared errors in some tables, for the logistic and other models, when the true relative risk is large. This is because the parameter estimates are effectively infinite in some samples. This can occur in binary regression when all of the $Y = 1$ instances occur at higher values of a covariate than all of the $Y = 0$ instances (Hosmer et. al 2013, section 4.4). This typically occurs only when the sample size is small and the covariates have strong predictive power.

Table 1: Expected values of relative risk estimators for $x_1 = 0.5$ vs $x_1 = -0.5$ when logit$(P[Y = 1]) = \alpha + \beta_1 x_1 + \beta_2 x_2$. Sample size is 100 and $x_1, x_2 \sim U(-2, 2)$. *Median prob* refers to the probability that $Y = 1$ when both covariates are at their median value. *Logit, log, poisson and log-logit* refer to binary regression with logit link (logistic regression) and log link, Poisson regression with log link and binary regression with blended log-logit link, respectively.

median prob.	true RR	RMSE of estimated RR logit	log	poisson	log-logit	prob.reject
0.3	1.200	0.202	0.172	0.176	0.175	0.1
0.3	1.500	0.292	0.198	0.211	0.212	0.8
0.3	3.000	2.131	1.122	0.783	0.914	28.5
0.5	1.200	0.126	0.099	0.107	0.112	1.2
0.5	1.500	0.212	0.179	0.144	0.177	9.3
0.5	3.000	1.798	1.428	0.952	1.102	40.9

Table 2: Expected values of relative risk estimators for $x_1 = 2$ vs $x_1 = 1$ when logit$(P[Y = 1]) = \alpha + \beta_1 x_1 + \beta_2 x_2$. Sample size is 100 and $x_1, x_2 \sim U(-2, 2)$. *Median prob* refers to the probability that $Y = 1$ when both covariates are at their median value. *Logit, log, poisson and log-logit* refer to binary regression with logit link (logistic regression) and log link, Poisson regression with log link and binary regression with blended log-logit link, respectively.

median prob.	true RR	RMSE of estimated RR logit	log	poisson	log-logit	prob.reject
0.3	1.173	0.131	0.174	0.180	0.177	0.1
0.3	1.332	0.108	0.231	0.261	0.243	0.8
0.3	1.346	0.100	0.580	0.982	0.371	28.5
0.5	1.098	0.046	0.489	0.991	0.087	40.9
0.5	1.143	0.059	0.098	0.117	0.108	1.2
0.5	1.206	0.044	0.186	0.282	0.164	9.3

Table 3: Expected values of relative risk estimators for $x_1 = -1$ vs $x_1 = -2$ when logit$(P[Y = 1]) = \alpha + \beta_1 x_1 + \beta_2 x_2$. Sample size is 100 and $x_1, x_2 \sim U(-2, 2)$. *Median prob* refers to the probability that $Y = 1$ when both covariates are at their median value. *Logit, log, poisson and log-logit* refer to binary regression with logit link (logistic regression) and log link, Poisson regression with log link and binary regression with blended log-logit link, respectively.

median prob.	true RR	RMSE of estimated RR				prob.reject
		logit	log	poisson	log-logit	
0.3	1.224	0.260	0.173	0.177	0.176	0.1
0.3	1.634	0.434	0.262	0.257	0.265	0.8
0.3	4.583	4.262	2.695	2.325	2.279	28.5
0.5	1.260	0.218	0.131	0.126	0.132	1.2
0.5	1.865	0.557	0.520	0.434	0.440	9.3
0.5	8.200	16.072	6.625	6.134	5.277	40.9

Table 4: Expected values of relative risk estimators for $x_1 = 0.5$ vs $x_1 = -0.5$ when logit$(P[Y = 1]) = \alpha + \beta_1 x_1 + \beta_2 x_2$. Sample size is 100 and $x_1, x_2 \sim N\left(0, \frac{4}{3}\right)$. *Median prob* refers to the probability that $Y = 1$ when both covariates are at their median value. *Logit, log, poisson and log-logit* refer to binary regression with logit link (logistic regression) and log link, Poisson regression with log link and binary regression with blended log-logit link, respectively.

median prob.	true RR	RMSE of estimated RR				prob.reject
		logit	log	poisson	log-logit	
0.3	1.200	0.214	0.147	0.170	0.171	0.4
0.3	1.500	0.305	0.198	0.185	0.202	2.2
0.3	3.000	1.627	1.485	1.041	0.855	25.4
0.5	1.200	0.136	0.106	0.105	0.116	1.4
0.5	1.500	0.235	0.260	0.146	0.180	10.9
0.5	3.000	1.711	1.667	1.196	0.913	36.3

Table 5: Expected values of relative risk estimators for $x_1 = 2$ vs $x_1 = 1$ when logit$(P[Y = 1]) = \alpha + \beta_1 x_1 + \beta_2 x_2$. Sample size is 100 and $x_1, x_2 \sim N\left(0, \frac{4}{3}\right)$. *Median prob* refers to the probability that $Y = 1$ when both covariates are at their median value. *Logit, log, poisson and log-logit* refer to binary regression with logit link (logistic regression) and log link, Poisson regression with log link and binary regression with blended log-logit link, respectively.

median prob.	true RR	RMSE of estimated RR				prob.reject
		logit	log	poisson	log-logit	
0.3	1.173	0.135	0.147	0.175	0.174	0.4
0.3	1.332	0.106	0.152	0.222	0.223	2.2
0.3	1.346	0.092	0.234	0.684	0.311	25.4
0.5	1.098	0.045	0.256	0.741	0.080	36.3
0.5	1.143	0.060	0.083	0.118	0.114	1.4
0.5	1.206	0.042	0.098	0.242	0.148	10.9

Table 6: Expected values of relative risk estimators for $x_1 = -1$ vs $x_1 = -2$ when logit$(P[Y = 1]) = \alpha + \beta_1 x_1 + \beta_2 x_2$. Sample size is 100 and $x_1, x_2 \sim N\left(0, \frac{4}{3}\right)$. *Median prob* refers to the probability that $Y = 1$ when both covariates are at their median value. *Logit, log, poisson and log-logit* refer to binary regression with logit link (logistic regression) and log link, Poisson regression with log link and binary regression with blended log-logit link, respectively.

median prob.	true RR	RMSE of estimated RR				prob.reject
		logit	log	poisson	log-logit	
0.3	1.224	0.273	0.151	0.170	0.171	0.4
0.3	1.634	0.451	0.304	0.251	0.266	2.2
0.3	4.583	3.417	3.064	2.607	2.239	25.4
0.5	1.260	0.237	0.150	0.123	0.131	1.4
0.5	1.865	0.618	0.617	0.465	0.442	10.9
0.5	8.200	54.848	6.865	6.385	5.454	36.3