

Subtypes of the Missing Not At Random Missing Data Mechanism: Supplemental Material

Brenna Gomer and Ke-Hai Yuan

Comparison of MNAR Generation Methods

This supplemental material is an extension of the “Generating Missing Values from MNAR Subtypes” section of the article. Specifically, the MNAR generation methods of Table 1 of the article are compared side by side in a small-scale simulation study. Readers may assume the notation, terminology, and set-up here are equivalent to the main text of the article in all respects unless explicitly noted otherwise.

Design

The simulation study of this supplemental section is analogous to Study 1 and conducted in a regression model context. However, a few adjustments were made to diversify the scenarios considered by the article. The predictor variables $\mathbf{x}_i = (x_{i1}, x_{i2}, x_{i3})'$ in the model $y_i = \mathbf{x}_i' \boldsymbol{\beta} + \varepsilon_i$ were generated from a multivariate normal distribution with $N = 400$ alongside an auxiliary variable z , such that

$$(x_{i1}, x_{i2}, x_{i3}, z)' \sim N(\mathbf{0}, \boldsymbol{\Sigma}), \quad \varepsilon_i \sim N(0, 2^2)$$

where $\boldsymbol{\Sigma}$ is a correlation matrix consisting of the correlation $\boldsymbol{\rho}_x = (\rho_{21}, \rho_{31}, \rho_{32})'$ between the x variables and the correlation $\boldsymbol{\rho}_z = (\rho_{z1}, \rho_{z2}, \rho_{z3})'$ between the auxiliary variable and predictors. The values of $\boldsymbol{\rho}_x$ were randomly generated from $Uniform(.1, .3)$ and set to $(.2146, .2816, .1403)'$. Values of $\boldsymbol{\rho}_z$ were set to $(.1531, .1744, .75)'$, where ρ_{z1} and ρ_{z2} were also generated from $Uniform(.1, .3)$. The last value ρ_{z3} was chosen so that z strongly correlated with x_3 , which is the variable subject to missingness. In conditions that did not use an auxiliary variable approach to generate MNAR values, z was ignored. Regression coefficients were set to $\boldsymbol{\beta} = (0, 2, 2, 2)'$.

The missingness rate for the predictor x_3 was set to 20%. The methods used to handle missing values were pattern-mixture models (PMM), selection models (SM), multiple imputation (MI), normal-distribution-based maximum likelihood (ML), and listwise deletion (LD) as implemented in Study 1 of the main text of the article. Four versions of diffuse MNAR processes (Diffuse 1a, Diffuse 1b, Diffuse 2a, and Diffuse 2b) were considered to allow the presence of interactions and variation of the strength of diffuse MNAR. As in the main text, Diffuse 1a and Diffuse 1b were weakly diffuse MNAR processes while Diffuse 2a and Diffuse 2b were strongly diffuse MNAR processes. Diffuse 1b and Diffuse 2b included interaction

terms, while Diffuse 1a and Diffuse 2a did not. Interactions were specified between x_2 and x_3 or between x_2 and z in auxiliary variable conditions. Although they share the same names, the MNAR missingness processes of this supplemental section were specified differently from those in the main text of the article and obtained via different criteria.

Comparability of MNAR Values Across MNAR Generation Methods

In general, comparisons between MNAR generation methods are confounded with other differences in how the MNAR mechanism is specified that are not inherently related to the generation methods. Thus, ensuring comparability of MNAR values is important in the context of this supplemental section. Great care was taken to minimize the differences between MNAR values from the MNAR generation methods – for example, this involved ensuring that Diffuse 1a was a weakly diffuse MNAR process of equivalent strength across the interval selection, logistic model, probit model, and auxiliary variable approaches. Efforts were also taken to match how deterministic the missingness processes were across the generation methods and how much each individual variable contributed to the missingness processes.

We sought to obtain comparable sets of MNAR values between MNAR generation methods by choosing sets of missingness parameters that 1) yielded similar proportions of missingness; 2) had the same strength of diffuse MNAR under a particular diffuse MNAR condition; 3) resulted in missingness processes that were equally deterministic; and 4) matched with respect to how much each individual variable contributed to the missingness processes. However, these aims were not (and cannot be) perfectly achieved.

Implementation of MNAR Generation Methods

The implementation of MNAR generation methods used in this supplemental section is as follows. Particular missingness parameter specifications and constraints are given in Table M1. Note that the parameter ψ_{int} is set to 0 for diffuse MNAR processes without interaction terms.

Interval selection The missingness criteria in the interval selection approach considered a linear combination of variables in the dataset rather than specifying criteria for each variable individually. That is, a new variable q was defined such that its values are given by

$$q_i = (y_i, x_{i1})\psi_s + x_{i2}\psi_L + x_{i3}\psi_k + (x_{i2} \times x_{i3})\psi_{int}$$

The ψ coefficients in q are missingness parameters whose values determine which MNAR subtype is present, the strength of diffuse MNAR, and presence of interactions as described by Table M1. Note there is no intercept ψ_0 for the interval selection approach. To set the probability of missingness, the values of q were first divided into three regions according to the 40th and 80th percentile values $q_{.40}$ and $q_{.80}$. To yield an overall missingness rate of 20%, the probability of missingness for the three regions was specified as

Table M1: Specifications and Constraints for Missingness Parameter Values $\boldsymbol{\psi} = (\psi_0, \psi_L, \boldsymbol{\psi}_s, \psi_k, \psi_{int})'$ across MNAR Generation Methods

Generation Method	Mechanism	Missingness Parameters				
		ψ_0	ψ_L	ψ_s	ψ_k	ψ_{int}
Interval Selection	Focused	–	0	0	1	0
	Diffuse	–	$1.5/*$	$1/*$	1	$c/*$
Logistic Model	Focused	-6.9068	0	0	ψ_v	0
	Diffuse	-6.9068	$1.5\psi_v/*$	$\psi_v/*$	ψ_v	$c\psi_v/*$
Probit Model	Focused	-3.0902	0	0	ψ_v	0
	Diffuse	-3.0902	$1.5\psi_v/*$	$\psi_v/*$	ψ_v	$c\psi_v/*$
Auxiliary Variables	Focused	ψ_0	0	0	ψ_{mis}	0
	Diffuse	ψ_0	ψ_L	ψ_s	ψ_k	ψ_{int}

Note. ψ_v was chosen so as to obtain the desired rate of missingness. ‘*’ and c are constants that varied between conditions. These granted additional flexibility to missingness parameters and were used to match the strength between diffuse MNAR processes. Diffuse MNAR processes without interaction terms set ψ_{int} to 0.

$$p_i = \begin{cases} .0005 & q_i \in (q_0, q_{.40}) \\ .001 & q_i \in (q_{.40}, q_{.80}) \\ .997 & q_i \in (q_{.80}, q_1) \end{cases} \quad (1)$$

Logistic and probit models Similar to the formulation used in the main text of the article, the logistic model of this appendix was specified as

$$\text{logit}(p_i) = \psi_0 + (y_i, x_{i1})\boldsymbol{\psi}_s + x_{i2}\psi_L + x_{i3}\psi_k + (x_{i2} \times x_{i3})\psi_{int} \quad (2)$$

The intercept value of -6.9068 in Table M1 was chosen so as to obtain a 0.1% missingness rate if all other parameters in $\boldsymbol{\psi}$ were set to 0¹. That is, under missingness that is MCAR, equation (2) yields $\text{logit}(.001) = -6.9068$. This procedure was used to find intercept values that are equivalent between the logistic and probit model approaches, so that their missingness processes are equally deterministic.

The probit model was specified as

$$p_i = \Phi[\psi_0 + (y_i, x_{i1})\boldsymbol{\psi}_s + x_{i2}\psi_L + x_{i3}\psi_k + (x_{i2} \times x_{i3})\psi_{int}] \quad (3)$$

The intercept value of -3.0902 in Table M1 was found using the same criteria as for the logistic model approach. The specific calculations for obtaining this value were $\Phi^{-1}(.001) = -3.0902$.

¹Thus, missingness was almost completely determined by the values of the variables and highly deterministic.

Auxiliary variables A total of three auxiliary variable conditions were considered to generate MNAR values, with the missingness process specified generally as

$$p_i = h[\psi_0 + (y_i, x_{i1})\boldsymbol{\psi}_s + x_{i2}\psi_L + z_i\psi_k + (x_{i2} \times z_i)\psi_{int}] \quad (4)$$

where the form of $h(\cdot)$ was specified according to the interval selection approach, the logistic model approach, or the probit model approach. Note that the use of these methods with auxiliary variables has replaced the term $x_{i3}\psi_k$ with $z_i\psi_k$. Unlike the other methods, the interval selection approach with auxiliary variables did not include the intercept ψ_0 . This approach followed the same procedure that was described previously for interval selection, except that q was redefined so that its values are given by

$$q_i = (y_i, x_{i1})\boldsymbol{\psi}_s + x_{i2}\psi_L + z_i\psi_k + (x_{i2} \times z_i)\psi_{int}$$

The missingness parameter constraints for a given method that used auxiliary variables were equivalent to the ones specified in Table M1 for the same given method without auxiliary variables. Note that the implementation of the auxiliary variable approach here treats z as a distinct variable with random error, but there are many other valid choices of z .

Results

Results are presented for the relative bias and RMSE (involving the absolute bias) of $\hat{\beta}_2$ and $\hat{\beta}_3$ in Tables M2-M5. For the sake of brevity, discussion is limited to MNAR generation methods and MNAR subtypes as these are the topics of this supplemental section and article respectively.

Comparing MNAR Generation Methods

In general Results in Tables M2-M5 are generally similar across the MNAR generation methods but with minor differences between methods according to whether or not they used auxiliary variables. For example, the logistic model, probit model, and interval selection approaches in the first three rows of Table M2 result in similar levels of bias in $\hat{\beta}_2$. However, these are somewhat different than the levels of bias seen for the approaches that used auxiliary variables which are shown in the fourth through sixth rows of the table. This occurs because MNAR values are inherently less comparable when they come from missingness processes that involve different sets of variables. Recall that the auxiliary variable approaches used z instead of x_3 in their missingness criteria.

Thus, the logistic model, probit model, and interval selection approaches tend to have similar levels of bias in $\hat{\beta}_2$ and $\hat{\beta}_3$ across the conditions in Tables M2 and M4, and similar RMSE values for the estimated regression coefficients across the conditions in Tables M3 and M5. Likewise, the three auxiliary variable approaches have similar levels of bias and RMSE values across the conditions in those tables respectively.

Results are most similar between the logistic and probit model approaches. For example, the levels of bias in $\hat{\beta}_2$ in the first and second rows of Table M2 for the logistic and probit model approaches are very close to each other.

Table M2: Empirical Bias for the Estimated Regression Coefficient $\hat{\beta}_2$

Method	MNAR Function	Focused	Diffuse1a	Diffuse1b	Diffuse2a	Diffuse2b
PMM	Logistic	-0.0032	-0.1205	-0.1220	-0.2077	-0.2077
	Probit	-0.0032	-0.1157	-0.1180	-0.2012	-0.2031
	Interval	-0.0026	-0.1279	-0.1298	-0.2223	-0.2223
	Aux. Logistic	-0.0101	-0.1169	-0.1183	-0.2091	-0.2091
	Aux. Probit	-0.0102	-0.1124	-0.1138	-0.2042	-0.2017
	Aux. Interval	-0.0105	-0.1410	-0.1426	-0.2196	-0.2224
SM	Logistic	0.0112	0.0035	0.0045	0.0017	0.0089
	Probit	0.0113	0.0040	0.0143	0.0026	0.0055
	Interval	0.0091	0.0021	0.0031	-0.0020	0.0033
	Aux. Logistic	0.0316	0.0177	0.0176	0.0024	0.0043
	Aux. Probit	0.0307	0.0213	0.0213	0.0027	0.0135
	Aux. Interval	0.0332	0.0078	0.0088	0.0018	0.0027
MI	Logistic	0.0148	0.0663	0.0686	0.0420	0.0542
	Probit	0.0139	0.0630	0.0886	0.0402	0.0451
	Interval	0.0156	0.0718	0.0741	0.0426	0.0509
	Aux. Logistic	0.0193	0.0561	0.0577	0.0389	0.0435
	Aux. Probit	0.0187	0.0539	0.0559	0.0360	0.0596
	Aux. Interval	0.0207	0.0591	0.0609	0.0431	0.0448
ML	Logistic	0.0078	0.0582	0.0607	0.0341	0.0489
	Probit	0.0074	0.0555	0.0836	0.0331	0.0386
	Interval	0.0077	0.0624	0.0656	0.0347	0.0444
	Aux. Logistic	0.0171	0.0503	0.0521	0.0319	0.0368
	Aux. Probit	0.0163	0.0484	0.0501	0.0292	0.0558
	Aux. Interval	0.0183	0.0524	0.0547	0.0352	0.0374
LD	Logistic	-0.0042	-0.0457	-0.0458	-0.0879	-0.0854
	Probit	-0.0042	-0.0444	-0.0389	-0.0847	-0.0848
	Interval	-0.0045	-0.0490	-0.0489	-0.0949	-0.0930
	Aux. Logistic	-0.0026	-0.0450	-0.0446	-0.0878	-0.0861
	Aux. Probit	-0.0031	-0.0429	-0.0428	-0.0859	-0.0743
	Aux. Interval	-0.0024	-0.0552	-0.0546	-0.0920	-0.0922

Note. Aux. = MNAR generation methods that use auxiliary variables; Interval = interval selection; PMM = pattern-mixture model; SM = selection model; MI = multiple imputation; ML = normal-distribution-based maximum likelihood; LD = listwise deletion.

Table M3: Empirical Root Mean Square Error (RMSE) of $\hat{\beta}_2$

Method	MNAR Function	Focused	Diffuse1a	Diffuse1b	Diffuse2a	Diffuse2b
PMM	Logistic	0.1058	0.2647	0.2676	0.4309	0.4310
	Probit	0.1056	0.2558	0.2610	0.4185	0.4221
	Interval	0.1058	0.2779	0.2818	0.4586	0.4588
	Aux. Logistic	0.1084	0.2590	0.2618	0.4334	0.4341
	Aux. Probit	0.1088	0.2508	0.2532	0.4240	0.4212
	Aux. Interval	0.1083	0.3033	0.3067	0.4538	0.4594
SM	Logistic	0.1179	0.1130	0.1129	0.1164	0.1167
	Probit	0.1192	0.1145	0.1165	0.1183	0.1174
	Interval	0.1172	0.1090	0.1092	0.1129	0.1124
	Aux. Logistic	0.1324	0.1344	0.1343	0.1180	0.1187
	Aux. Probit	0.1309	0.1393	0.1397	0.1180	0.1235
	Aux. Interval	0.1331	0.1207	0.1211	0.1172	0.1170
MI	Logistic	0.1180	0.1769	0.1798	0.1444	0.1590
	Probit	0.1195	0.1712	0.2112	0.1424	0.1478
	Interval	0.1194	0.1840	0.1872	0.1438	0.1537
	Aux. Logistic	0.1175	0.1612	0.1626	0.1402	0.1452
	Aux. Probit	0.1166	0.1581	0.1612	0.1366	0.1656
	Aux. Interval	0.1193	0.1647	0.1672	0.1441	0.1465
ML	Logistic	0.1147	0.1641	0.1673	0.1347	0.1514
	Probit	0.1157	0.1605	0.2025	0.1341	0.1396
	Interval	0.1152	0.1691	0.1737	0.1344	0.1449
	Aux. Logistic	0.1164	0.1528	0.1550	0.1319	0.1370
	Aux. Probit	0.1153	0.1501	0.1522	0.1287	0.1599
	Aux. Interval	0.1168	0.1551	0.1582	0.1354	0.1376
LD	Logistic	0.1149	0.1499	0.1499	0.2132	0.2093
	Probit	0.1148	0.1483	0.1423	0.2089	0.2092
	Interval	0.1155	0.1541	0.1546	0.2252	0.2223
	Aux. Logistic	0.1167	0.1491	0.1486	0.2130	0.2106
	Aux. Probit	0.1171	0.1461	0.1461	0.2096	0.1920
	Aux. Interval	0.1157	0.1636	0.1633	0.2209	0.2215

Note. Aux. = MNAR generation methods that use auxiliary variables; Interval = interval selection; PMM = pattern-mixture model; SM = selection model; MI = multiple imputation; ML = normal-distribution-based maximum likelihood; LD = listwise deletion.

Likewise, the levels of bias in $\hat{\beta}_2$ for the logistic and probit model approaches with auxiliary variables in the fourth and fifth rows of Table M2 are very much alike. This similarity persists across the rest of the conditions in Table M2 as well as across Tables M3-M5 and is a result of the rough mathematical equivalence between the logistic and probit models.

Across MNAR mechanisms The differences in results between MNAR generation methods are consistent across focused and diffuse MNAR processes. For example, the difference in the levels of bias in $\hat{\beta}_2$ between the probit model and interval selection approach in the second and third rows of Table M2 is roughly the same across the MNAR mechanisms in the columns. This also holds for other pairs of MNAR generation methods and across missing-data handling methods for the results in Tables M2-M5.

There is one exception to this pattern: the difference between the interval selection approach with auxiliary variables and other MNAR generation methods is larger under Diffuse 1a and Diffuse 1b processes than it is under focused MNAR, Diffuse 2a, and Diffuse 2b processes. For example, comparing this approach to the probit model approach with auxiliary variables, the difference in the levels of bias in $\hat{\beta}_2$ in Table M2 is 0.0286 and 0.0288 under Diffuse 1a and Diffuse 1b but 0.0154 and 0.0207 under Diffuse 2a and Diffuse 2b. This pattern also occurs for the interval selection approach with auxiliary variables and the other MNAR generation methods. This finding persists across the results in Tables M2-M5. This happens because the set of ψ missingness parameters for weakly diffuse MNAR under the auxiliary interval selection method was not as well-matched to the other methods in this case.

Across missing-data handling methods The magnitude of the differences in results between MNAR generation methods are not constant across methods used to handle missing values. For example, the difference in the bias of $\hat{\beta}_2$ in Table M2 between the logistic and probit model approaches under diffuse 1b is 0.004 for PMM but is 0.02 for MI. This type of pattern can also be seen in Tables M3-M5. This pattern occurs because the reaction to the particular balance of variables involved in the missingness process differs between missing-data handling methods. In addition, the influence of a specific variable on the results was not necessarily equivalent between MNAR generation methods despite our best efforts.

Comparing MNAR Processes

Like the results of Study 1 in the main text of the article, there are substantial differences between focused and diffuse MNAR in terms of their adverse impact on estimated regression coefficients as shown in Tables M2-M5.

In general Across the tables, the strength of diffuse MNAR has a modest influence on $\hat{\beta}_2$ and $\hat{\beta}_3$ with respect to bias and RMSE. The inclusion of interaction terms has almost no effect on these results in most cases. However, the inclusion of interaction terms does have a minor impact on results under weakly diffuse MNAR (Diffuse 1a and Diffuse 1b) values that were generated from the probit model approach. This is most apparent under MI and ML.

Table M4: Empirical Bias for the Estimated Regression Coefficient $\hat{\beta}_3$

Method	MNAR Function	Focused	Diffuse1a	Diffuse1b	Diffuse2a	Diffuse2b
PMM	Logistic	0.0061	-0.2060	-0.2069	-0.2511	-0.2533
	Probit	0.0043	-0.1982	-0.1913	-0.2464	-0.2481
	Interval	0.0002	-0.2167	-0.2181	-0.2600	-0.2606
	Aux. Logistic	0.0032	-0.1644	-0.1651	-0.2438	-0.2437
	Aux. Probit	0.0031	-0.1574	-0.1581	-0.2375	-0.2310
	Aux. Interval	0.0033	-0.1949	-0.1952	-0.2526	-0.2541
SM	Logistic	0.0635	0.0074	0.0069	0.0010	-0.0001
	Probit	0.0632	0.0073	0.0045	0.0002	-0.0002
	Interval	0.0610	0.0082	0.0078	0.0046	0.0037
	Aux. Logistic	0.0269	0.0030	0.0029	-0.0005	-0.0010
	Aux. Probit	0.0257	0.0019	0.0017	-0.0012	-0.0027
	Aux. Interval	0.0291	0.0032	0.0033	0.0011	0.0008
MI	Logistic	0.1382	0.0161	0.0157	-0.0071	-0.0048
	Probit	0.1323	0.0157	0.0251	-0.0067	-0.0063
	Interval	0.1445	0.0147	0.0152	-0.0088	-0.0064
	Aux. Logistic	0.0636	0.0131	0.0136	-0.0073	-0.0064
	Aux. Probit	0.0614	0.0127	0.0123	-0.0073	-0.0016
	Aux. Interval	0.0669	0.0085	0.0090	-0.0073	-0.0071
ML	Logistic	0.1305	0.0317	0.0313	0.0051	0.0053
	Probit	0.1250	0.0308	0.0353	0.0049	0.0047
	Interval	0.1366	0.0331	0.0325	0.0045	0.0046
	Aux. Logistic	0.0629	0.0226	0.0225	0.0041	0.0041
	Aux. Probit	0.0611	0.0215	0.0215	0.0035	0.0049
	Aux. Interval	0.0663	0.0193	0.0194	0.0050	0.0047
LD	Logistic	-0.0022	-0.0935	-0.0927	-0.0892	-0.0845
	Probit	-0.0028	-0.0896	-0.0740	-0.0858	-0.0845
	Interval	-0.0038	-0.0996	-0.0986	-0.0938	-0.0894
	Aux. Logistic	-0.0010	-0.0725	-0.0719	-0.0836	-0.0818
	Aux. Probit	-0.0006	-0.0698	-0.0693	-0.0798	-0.0651
	Aux. Interval	-0.0014	-0.0815	-0.0801	-0.0885	-0.0870

Note. Aux. = MNAR generation methods that use auxiliary variables; Interval = interval selection; PMM = pattern-mixture model; SM = selection model; MI = multiple imputation; ML = normal-distribution-based maximum likelihood; LD = listwise deletion.

Table M5: Empirical Root Mean Square Error (RMSE) of $\hat{\beta}_3$

Method	MNAR Function	Focused	Diffuse1a	Diffuse1b	Diffuse2a	Diffuse2b
PMM	Logistic	0.2249	0.4359	0.4379	0.5215	0.5265
	Probit	0.2178	0.4214	0.4113	0.5135	0.5169
	Interval	0.2818	0.4592	0.4620	0.5377	0.5393
	Aux. Logistic	0.1306	0.3616	0.3631	0.5086	0.5084
	Aux. Probit	0.1298	0.3476	0.3492	0.4967	0.4876
	Aux. Interval	0.1300	0.4164	0.4172	0.5242	0.5272
SM	Logistic	0.2041	0.1156	0.1158	0.1103	0.1105
	Probit	0.2025	0.1160	0.1148	0.1107	0.1102
	Interval	0.1987	0.1132	0.1130	0.1101	0.1092
	Aux. Logistic	0.1553	0.1207	0.1204	0.1101	0.1098
	Aux. Probit	0.1529	0.1210	0.1206	0.1095	0.1089
	Aux. Interval	0.1583	0.1164	0.1164	0.1103	0.1099
MI	Logistic	0.3208	0.1518	0.1507	0.1309	0.1273
	Probit	0.3098	0.1506	0.1476	0.1291	0.1291
	Interval	0.3318	0.1554	0.1555	0.1338	0.1274
	Aux. Logistic	0.1827	0.1401	0.1401	0.1268	0.1266
	Aux. Probit	0.1797	0.1390	0.1387	0.1258	0.1190
	Aux. Interval	0.1887	0.1376	0.1368	0.1320	0.1307
ML	Logistic	0.2972	0.1413	0.1413	0.1155	0.1161
	Probit	0.2874	0.1397	0.1431	0.1155	0.1150
	Interval	0.3078	0.1445	0.1436	0.1164	0.1154
	Aux. Logistic	0.1785	0.1321	0.1319	0.1140	0.1138
	Aux. Probit	0.1756	0.1306	0.1302	0.1127	0.1128
	Aux. Interval	0.1838	0.1276	0.1276	0.1154	0.1150
LD	Logistic	0.1488	0.2280	0.2271	0.2146	0.2072
	Probit	0.1472	0.2215	0.1975	0.2094	0.2071
	Interval	0.1483	0.2399	0.2382	0.2238	0.2161
	Aux. Logistic	0.1341	0.1950	0.1940	0.2054	0.2024
	Aux. Probit	0.1332	0.1910	0.1901	0.1992	0.1764
	Aux. Interval	0.1343	0.2068	0.2043	0.2141	0.2115

Note. Aux. = MNAR generation methods that use auxiliary variables; Interval = interval selection; PMM = pattern-mixture model; SM = selection model; MI = multiple imputation; ML = normal-distribution-based maximum likelihood; LD = listwise deletion.

This occurs because the criteria for comparable MNAR values were simultaneously applied across MNAR generation methods within Diffuse 1a and Diffuse 1b, as well as between Diffuse 1a and Diffuse 1b. In other words, the set of chosen probit missingness parameters sacrificed comparability between Diffuse 1a and Diffuse 1b in favor of comparability between MNAR generation methods.

Estimates of β_2 Under focused MNAR, the relative bias of $\hat{\beta}_2$ in Table M2 is negligible and its RMSE in Table M3 tends to be small across missing-data handling methods. Under diffuse MNAR, the relative bias and RMSE of $\hat{\beta}_2$ tend to be larger – but the severity of this increase differs substantially between missing-data handling methods. These results are similar to the results of Study 1 in the main text of the article.

Estimates of β_3 The bias and RMSE of $\hat{\beta}_3$ in Tables M4 and M5 are larger under diffuse MNAR than under focused MNAR when PMM and LD are used. However, the bias and RMSE of $\hat{\beta}_3$ are larger under focused MNAR than under diffuse MNAR when SM, MI, and ML are used. This finding differs from the pattern of results in Study 1, and likely results from the stronger relationships between variables in this scenario. The data generated in this supplemental section had larger regression coefficients, larger correlations between variables, and smaller random error variance compared to the data generated for Study 1. As relationships between variables become stronger, the observed variables have greater influence on the missingness process under both focused and diffuse MNAR. When this effect was added to the contribution of missingness coefficients for completely observed variables under diffuse MNAR, the cumulative influence of the observed variables became much greater than the influence of x_3 . This allowed the strategies of SM, MI, and ML to better estimate the variance of x_3 and lead to improved estimates of β_3 .

Discussion

As two kinds of results were presented that match the intent of this supplemental section and main text of the article, we will discuss each of these separately.

MNAR generation methods The choice of MNAR generation method does not affect the pattern of statistical results under focused and diffuse MNAR. The results in Tables M2-M5 are very similar across methods. Differences in results between MNAR generation methods occur when MNAR values are no longer comparable. For example, the interval selection approach with auxiliary variables differed more from the other MNAR generation methods under Diffuse 1a and 1b than under the other MNAR processes. Large differences can arise due to inconsistencies regarding the strength of diffuse MNAR, sets of variables involved in the missingness process, or how deterministic/probabilistic the missingness processes are.

Distinctions of MNAR This supplemental section provides additional evidence that focused and diffuse MNAR consistently differ in the severity of their impact on statistical inference. Like Study 1 in the main text of the article, this supplemental section found

substantial differences between focused and diffuse MNAR in terms of their impact on estimated regression coefficients. Also, the strength of diffuse MNAR had a minor impact on results while the presence of interactions had very little impact. Thus, the distinction between focused and diffuse MNAR is still valuable in light of those considerations. This is also supported by the findings in the main text of the article.

Conclusion

This supplemental section has shown that simulation study findings are not dependent on the MNAR generation method. Findings depend on the specification of the MNAR process itself – such as the MNAR subtype, strength of diffuse MNAR, or presence of interaction terms. The results of this supplemental section also support the claims in the main text of the article – that the distinction between focused and diffuse MNAR is a valuable distinction of the MNAR mechanism.

In general, MNAR values are not comparable between methods unless great pains are taken to ensure otherwise. Achieving comparable MNAR missingness can be difficult even in a small-scale simulation study like this one. Thus, we recommend generating MNAR values using the same method consistently throughout a simulation study if comparable sets of MNAR values are required. The choice of MNAR generation method should be based on what fits the research question and setting.