# Appendix A

**Table A1** Performance comparison between SEM and MICE under simulated missingness scenarios with 5 variables and 5000 samples.

| Mechanism | Proportion | Metrics | SEM | MICE | Wilcoxon $p$-value | Bootstrap CI |
|---|---|---|---|---|---|---|
| MCAR | 0.15 | Precision | 0.7587 | 0.7543 | 0.8339 | (-0.0043, 0.029) |
| | | Recall | 0.7606 | 0.7569 | 0.8339 | (-0.0046,0.0022) |
| | | F1-score | 0.7596 | 0.7555 | 0.8339 | (-0.0044,0.026) |
| MCAR | 0.3 | Precision | 0.7772 | 0.7821 | 0.5169 | (-0.023,0.0107) |
| | | Recall | 0.7816 | 0.7875 | 0.323 | (-0.0208,0.0072) |
| | | F1-score | 0.7792 | 0.7845 | 0.5004 | (-0.0204,0.0094) |
| MCAR | 0.5 | Precision | 0.7695 | 0.7707 | 0.777 | (-0.028,0.032) |
| | | Recall | 0.7704 | 0.7771 | 0.8719 | (-0.027,0.029) |
| | | F1-score | 0.7699 | 0.7706 | 0.778 | (-0.0262,0.0322) |
| MAR | 0.15 | Precision | 0.7879 | 0.7804 | 0.7213 | (-0.0161,0.0435) |
| | | Recall | 0.7915 | 0.7847 | 0.7213 | (-0.0145,0.0421) |
| | | F1-score | 0.7894 | 0.7822 | 0.7213 | (-0.0163,0.0432) |
| MAR | 0.3 | Precision | 0.736 | 0.7568 | 0.213 | (-0.0542,0.0143) |
| | | Recall | 0.7915 | 0.7847 | 0.7213 | (-0.0145,0.0421) |
| | | F1-score | 0.7894 | 0.7822 | 0.7213 | (-0.0163,0.0432) |
| MAR | 0.5 | Precision | 0.7399 | 0.7445 | 0.7476 | (-0.0448,0.0342) |
| | | Recall | 0.7442 | 0.7489 | 0.7987 | (-0.0445,0.0307) |
| | | F1-score | 0.7418 | 0.7464 | 0.7817 | (-0.0437,0.0332) |
| MNAR | 0.15 | Precision | 0.7229 | 0.7185 | 0.476 | (-0.0077,0.0162) |
| | | Recall | 0.7304 | 0.725 | 0.439 | (-0.0049,0.0184) |
| | | F1-score | 0.7264 | 0.7215 | 0.476 | (-0.0068,0.0169) |
| MNAR | 0.3 | Precision | 0.7476 | 0.7342 | 0.4201 | (-0.0085,0.0437) |
| | | Recall | 0.7518 | 0.7384 | 0.3922 | (-0.0075,0.0419) |
| | | F1-score | 0.7494 | 0.7326 | 0.3932 | (-0.00087,0.0431) |
| MNAR | 0.5 | Precision | 0.7463 | 0.752 | 0.875 | (-0.046,0.0316) |
| | | Recall | 0.7531 | 0.7576 | 0.967 | (-0.0417,0.033) |
| | | F1-score | 0.7494 | 0.7544 | 0.923 | (-0.0452,0.031) |

**Table A2** Performance comparison between SEM and MICE under simulated missingness scenarios with 10 variables and 1000 samples.

| Mechanism | Proportion | Metrics | SEM | MICE | Wilcoxon $p$-value | Bootstrap CI |
|---|---|---|---|---|---|---|
| MCAR | 0.15 | Precision | 0.8265 | 0.8037 | 0.0322 | (0.0032, 0.0439) |
| | | Recall | 0.8162 | 0.79 | 0.0073 | (0.0074,0.0487) |
| | | F1-score | 0.821 | 0.7964 | 0.0188 | (0.0058,0.0471) |
| MCAR | 0.3 | Precision | 0.8441 | 0.8264 | 0.049 | (-0.0076,0.041) |
| | | Recall | 0.829 | 0.7922 | 0.0008 | (0.0095,0.0615) |
| | | F1-score | 0.8361 | 0.8082 | 0.006 | (0.0023,0.0514) |
| MCAR | 0.5 | Precision | 0.8361 | 0.8139 | 0.055 | (-0.003,0.0505) |
| | | Recall | 0.8162 | 0.7711 | 0.0011 | (0.0165,0.074) |
| | | F1-score | 0.8256 | 0.7909 | 0.0084 | (0.007,0.0626) |
| MAR | 0.15 | Precision | 0.8247 | 0.8231 | 0.6094 | (-0.0265,0.0246) |
| | | Recall | 0.8174 | 0.8086 | 0.3273 | (-0.0177,0.0341) |
| | | F1-score | 0.8207 | 0.8155 | 0.4362 | (-0.0201,0.0303) |
| MAR | 0.3 | Precision | 0.8278 | 0.8299 | 0.639 | (-0.0225,0.0186) |
| | | Recall | 0.8165 | 0.8036 | 0.117 | (-0.0091,0.0349) |
| | | F1-score | 0.8219 | 0.816 | 0.242 | (-0.0165,0.0255) |
| MAR | 0.5 | Precision | 0.8416 | 0.7951 | 0.0002 | (0.0144,0.075) |
| | | Recall | 0.8246 | 0.749 | 0.4.9E-06 | (0.0429,0.1099) |
| | | F1-score | 0.8327 | 0.7704 | 0.1.63E-05 | (0.0322,0.0924) |
| MNAR | 0.15 | Precision | 0.8317 | 0.8121 | 0.045 | (0.0002,0.0405) |
| | | Recall | 0.8201 | 0.7994 | 0.049 | (0.0013,0.0426) |
| | | F1-score | 0.8256 | 0.8053 | 0.052 | (0.0004,0.0408) |
| MNAR | 0.3 | Precision | 0.8422 | 0.8204 | 0.078 | (-0.0013,0.0476) |
| | | Recall | 0.8314 | 0.7949 | 0.004 | (0.0145,0.0618) |
| | | F1-score | 0.8365 | 0.8069 | 0.014 | (0.0058,0.0557) |
| MNAR | 0.5 | Precision | 0.8464 | 0.8151 | 0.0228 | (0.0073,0.0578) |
| | | Recall | 0.8266 | 0.7707 | 0.0001 | (0.0323,0.0822) |
| | | F1-score | 0.836 | 0.7917 | 0.0008 | (0.0225,0.0703) |

**Table A3** Performance comparison between SEM and MICE under simulated missingness scenarios with 10 variables and 5000 samples.

| Mechanism | Proportion | Metrics | SEM | MICE | Wilcoxon $p$-value | Bootstrap CI |
|---|---|---|---|---|---|---|
| MCAR | 0.15 | Precision | 0.7961 | 0.7902 | 0.7241 | (-0.0125, 0.0248) |
| | | Recall | 0.8165 | 0.8161 | 0.9885 | (-0.0145,0.0144) |
| | | F1-score | 0.8056 | 0.8023 | 0.8644 | (-0.0122,0.02) |
| MCAR | 0.3 | Precision | 0.7937 | 0.7932 | 0.5663 | (-0.0311,0.0262) |
| | | Recall | 0.8153 | 0.8157 | 0.6811 | (-0.0288,0.0206) |
| | | F1-score | 0.8038 | 0.8038 | 0.5505 | (-0.0289,0.0226) |
| MCAR | 0.5 | Precision | 0.8128 | 0.8167 | 0.8125 | (-0.0305,0.0185) |
| | | Recall | 0.833 | 0.8348 | 0.8875 | (-0.0234,0.0189) |
| | | F1-score | 0.8222 | 0.8252 | 0.911 | (-0.0292,0.0195) |
| MAR | 0.15 | Precision | 0.7666 | 0.7756 | 0.2641 | (-0.0329,0.0104) |
| | | Recall | 0.795 | 0.8017 | 0.2886 | (-0.024,0.0113) |
| | | F1-score | 0.7799 | 0.7878 | 0.2758 | (-0.0259,0.012) |
| MAR | 0.3 | Precision | 0.7923 | 0.7827 | 0.2977 | (-0.0087,0.0309) |
| | | Recall | 0.8144 | 0.8069 | 0.2355 | (-0.0075,0.0251) |
| | | F1-score | 0.8026 | 0.794 | 0.2326 | (-0.0081,0.0278) |
| MAR | 0.5 | Precision | 0.8085 | 0.7993 | 0.4337 | (-0.0178,0.0372) |
| | | Recall | 0.824 | 0.8183 | 0.054 | (-0.0187,0.0293) |
| | | F1-score | 0.8158 | 0.8082 | 0.04 | (-0.0165,0.0334) |
| MNAR | 0.15 | Precision | 0.7969 | 0.7812 | 0.0359 | (-0.0038,0.0344) |
| | | Recall | 0.82 | 0.8047 | 0.04 | (-0.0012,0.0314) |
| | | F1-score | 0.8077 | 0.7222 | 0.035 | (-0.0028,0.0338) |
| MNAR | 0.3 | Precision | 0.7816 | 0.7733 | 0.3388 | (-0.0115,0.03) |
| | | Recall | 0.8066 | 0.8019 | 0.5718 | (-0.0148,0.0232) |
| | | F1-score | 0.7932 | 0.7867 | 0.3876 | (-0.0135,0.0253) |
| MNAR | 0.5 | Precision | 0.8103 | 0.7866 | 0.1611 | (-0.0128,0.0587) |
| | | Recall | 0.8293 | 0.8074 | 0.1195 | (-0.0072,0.0516) |
| | | F1-score | 0.8193 | 0.7964 | 0.1449 | (-0.0087,0.0565) |

**Table A4** Performance comparison between SEM and MICE under simulated missingness scenarios with 20 variables and 1000 samples.

| Mechanism | Proportion | Metrics | SEM | MICE | Wilcoxon $p$-value | Bootstrap CI |
|---|---|---|---|---|---|---|
| MCAR | 0.15 | Precision | 0.8812 | 0.8697 | 0.1178 | (-0.0029, 0.0248) |
|  |  | Recall | 0.8573 | 0.8464 | 0.1366 | (-0.0036,0.0261) |
|  |  | F1-score | 0.8688 | 0.8575 | 0.1505 | (-0.0036,0.0254) |
| MCAR | 0.3 | Precision | 0.8842 | 0.874 | 0.036 | (-0.0044,0.0238) |
|  |  | Recall | 0.8675 | 0.8485 | 0.0072 | (0.0002,0.0353) |
|  |  | F1-score | 0.8754 | 0.8607 | 0.0132 | (-0.0017,0.03) |
| MCAR | 0.5 | Precision | 0.8883 | 0.8769 | 0.047 | (-0.0032,0.0253) |
|  |  | Recall | 0.8723 | 0.8458 | 0.001 | (0.0104,0.0443) |
|  |  | F1-score | 0.88 | 0.8606 | 0.005 | (0.0031,0.0341) |
| MAR | 0.15 | Precision | 0.8854 | 0.7824 | 0.034 | (-0.001,0.028) |
|  |  | Recall | 0.869 | 0.8502 | 0.022 | (0.0027,0.0371) |
|  |  | F1-score | 0.8768 | 0.8606 | 0.029 | (0.0014,0.0325) |
| MAR | 0.3 | Precision | 0.884 | 0.8757 | 0.06 | (-0.0055,0.0211) |
|  |  | Recall | 0.8671 | 0.8483 | 0.0014 | (0.0011,0.034) |
|  |  | F1-score | 0.8751 | 0.8613 | 0.0069 | (-0.0005,0.0281) |
| MAR | 0.5 | Precision | 0.8847 | 0.8702 | 0.021 | (0.0027,0.0285) |
|  |  | Recall | 0.862 | 0.8308 | 9.06E-05 | (0.0172,0.0456) |
|  |  | F1-score | 0.8728 | 0.8494 | 6.0E-04 | (0.0098,0.0365) |
| MNAR | 0.15 | Precision | 0.8889 | 0.8894 | 0.5517 | (-0.0147,0.0118) |
|  |  | Recall | 0.8639 | 0.8628 | 0.2118 | (-0.0152,0.0166) |
|  |  | F1-score | 0.8758 | 0.8755 | 0.295 | (-0.0152,0.0148) |
| MNAR | 0.3 | Precision | 0.8856 | 0.8722 | 0.065 | (-0.0057,0.032) |
|  |  | Recall | 0.8573 | 0.8361 | 0.248 | (-0.0016,0.0445) |
|  |  | F1-score | 0.8708 | 0.8532 | 0.045 | (-0.0041,0.0376) |
| MNAR | 0.5 | Precision | 0.9012 | 0.8587 | 5.01E-06 | (0.024,0.0629) |
|  |  | Recall | 0.8774 | 0.8142 | 1.03E-07 | (0.0414,0.0865) |
|  |  | F1-score | 0.8888 | 0.8352 | 2.34E-07 | (0.0334,0.0763) |

**Table A5** Performance comparison between SEM and MICE under simulated missingness scenarios with 20 variables and 5000 samples.

| Mechanism | Proportion | Metrics | SEM | MICE | Wilcoxon $p$-value | Bootstrap CI |
|---|---|---|---|---|---|---|
| MCAR | 0.15 | Precision | 0.8615 | 0.8514 | 0.198 | (0.0014, 0.0249) |
|  |  | Recall | 0.8794 | 0.8704 | 0.148 | (0.0012,0.0215) |
|  |  | F1-score | 0.8701 | 0.8606 | 0.138 | (0.0013,0.0226) |
| MCAR | 0.3 | Precision | 0.8876 | 0.8705 | 0.2642 | (-0.0006,0.0416) |
|  |  | Recall | 0.9002 | 0.8843 | 0.1843 | (0.001,0.0377) |
|  |  | F1-score | 0.8937 | 0.8772 | 0.2026 | (0.0007,0.0408) |
| MCAR | 0.5 | Precision | 0.879 | 0.8789 | 0.723 | (-0.0221,0.0225) |
|  |  | Recall | 0.8924 | 0.8921 | 0.79 | (-0.0193,0.0183) |
|  |  | F1-score | 0.8855 | 0.8853 | 0.787 | (-0.0215,0.0202) |
| MAR | 0.15 | Precision | 0.8414 | 0.843 | 0.93 | (-0.0155,0.0112) |
|  |  | Recall | 0.8593 | 0.8606 | 0.94 | (-0.0157,0.0102) |
|  |  | F1-score | 0.85 | 0.8515 | 0.94 | (-0.0146,0.0108) |
| MAR | 0.3 | Precision | 0.8809 | 0.8713 | 0.222 | (-0.0068,0.0268) |
|  |  | Recall | 0.894 | 0.8826 | 0.182 | (-0.0024,0.0272) |
|  |  | F1-score | 0.8873 | 0.8768 | 0.208 | (-0.0049,0.0266) |
| MAR | 0.5 | Precision | 0.8785 | 0.873 | 0.524 | (-0.014,0.0259) |
|  |  | Recall | 0.8921 | 0.8842 | 0.037 | (-0.0093,0.0265) |
|  |  | F1-score | 0.8851 | 0.8784 | 0.058 | (-0.0113,0.0261) |
| MNAR | 0.15 | Precision | 0.8768 | 0.8423 | 0.0262 | (0.0021,0.0281) |
|  |  | Recall | 0.888 | 0.8782 | 0.043 | (0.0001,0.0226) |
|  |  | F1-score | 0.8826 | 0.87 | 0.034 | (0.0012,0.0268) |
| MNAR | 0.3 | Precision | 0.8961 | 0.8568 | 0.0008 | (0.0198,0.0671) |
|  |  | Recall | 0.9064 | 0.8704 | 0.0007 | (0.0186,0.0616) |
|  |  | F1-score | 0.9011 | 0.8634 | 0.001 | (0.0207,0.0653) |
| MNAR | 0.5 | Precision | 0.895 | 0.8808 | 0.046 | (-0.0021,0.0346) |
|  |  | Recall | 0.9056 | 0.8911 | 0.008 | (0.0018,0.0318) |
|  |  | F1-score | 0.9001 | 0.8858 | 0.012 | (-0.0007,0.0329) |