**SUPPLEMENTARY MATERIALS**

**Table S1. Results of five models incorporating covariate selection stability (with selection thresholds set from x = 0 to x=10), conducted on 4 simulated high signal datasets (Table 1), to illustrate the number of false positive and false negative covariates selected in final models for each method at different threshold values of x.**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | sAIC | | | | mBIC | | | | enet | | | | MCP | | | | Combi | | |
| Dataset HS1 | FV | FP | FN | FV | | FP | FN | FV | | FP | FN | FV | | FP | FN | FV | | FP | FN |
| x = 0 | 0 | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 |
| x =1 | 0 | 0 | 0 | 1 | | 1 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 |
| x =3 | 0 | 0 | 0 | 2 | | 2 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 |
| x =5 | 0 | 0 | 0 | 2 | | 2 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 |
| x =10 | 1 | 1 | 0 | 8 | | 8 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 |
| Dataset HS2 |  |  |  |  | |  |  |  | |  |  |  | |  |  |  | |  |  |
| x = 0 | 0 | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 |
| x =1 | 1 | 1 | 0 | 1 | | 1 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 1 | | 1 | 0 |
| x =3 | 1 | 1 | 0 | 3 | | 3 | 0 | 1 | | 1 | 0 | 2 | | 2 | 0 | 1 | | 1 | 0 |
| x =5 | 1 | 1 | 0 | 3 | | 3 | 0 | 1 | | 1 | 0 | 3 | | 3 | 0 | 1 | | 1 | 0 |
| x =10 | 3 | 3 | 0 | 9 | | 9 | 0 | 1 | | 1 | 0 | 4 | | 4 | 0 | 2 | | 2 | 0 |
| Dataset HS3 |  |  |  |  | |  |  |  | |  |  |  | |  |  |  | |  |  |
| x = 0 | 0 | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 |
| x =1 | 0 | 0 | 0 | 2 | | 2 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 |
| x =3 | 1 | 1 | 0 | 4 | | 4 | 0 | 0 | | 0 | 0 | 1 | | 1 | 0 | 1 | | 1 | 0 |
| x =5 | 1 | 1 | 0 | 4 | | 4 | 0 | 1 | | 1 | 0 | 1 | | 1 | 0 | 1 | | 1 | 0 |
| x =10 | 2 | 2 | 0 | 9 | | 9 | 0 | 1 | | 1 | 0 | 3 | | 3 | 0 | 2 | | 2 | 0 |
| Dataset HS4 |  |  |  |  | |  |  |  | |  |  |  | |  |  |  | |  |  |
| x = 0 | 0 | 0 | 0 | 1 | | 1 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 |
| x =1 | 0 | 0 | 0 | 2 | | 2 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 1 | | 1 | 0 |
| x =3 | 1 | 1 | 0 | 2 | | 2 | 0 | 1 | | 1 | 0 | 0 | | 0 | 0 | 1 | | 1 | 0 |
| x =5 | 3 | 3 | 0 | 3 | | 3 | 0 | 2 | | 2 | 0 | 1 | | 1 | 0 | 2 | | 2 | 0 |
| x =10 | 8 | 8 | 0 | 4 | | 4 | 0 | 3 | | 3 | 0 | 2 | | 2 | 0 | 3 | | 3 | 0 |

Key; sAIC – covariate selection based on stepwise Akaike information criterion, mBIC - covariate selection based on a stepwise on a modified Bayesian Information Criterion, enet – covariate section based on elastic net regression, MCP – covariate selection based on minimax convex penalty regression, Combi – covariate selection based on a synthesised combination of sAIC, mBIC, enet and MCP, FV – total number of false variables (false positive and negative covariates) selected, FP – number of false positive covariates selected, FN – number of false negative variables selected.

**Table S2. Results of five models incorporating covariate selection stability (with selection thresholds set from x = 0 to x=10), conducted on 4 simulated medium signal datasets (Table 1), to illustrate the number of false positive and false negative covariates selected in final models for each method at different threshold values of x.**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | sAIC | | | | mBIC | | | | enet | | | | MCP | | | | Combi | | |
| Dataset MS1 | FV | FP | FN | FV | | FP | FN | FV | | FP | FN | FV | | FP | FN | FV | | FP | FN |
| x = 0 | 0 | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 |
| x =1 | 0 | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 |
| x =3 | 0 | 0 | 0 | 3 | | 3 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 |
| x =5 | 0 | 0 | 0 | 5 | | 5 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 |
| x =10 | 1 | 1 | 0 | 9 | | 9 | 0 | 0 | | 0 | 0 | 2 | | 2 | 0 | 0 | | 0 | 0 |
| Dataset MS2 |  |  |  |  | |  |  |  | |  |  |  | |  |  |  | |  |  |
| x = 0 | 0 | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 |
| x =1 | 0 | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 |
| x =3 | 0 | 0 | 0 | 1 | | 1 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 |
| x =5 | 0 | 0 | 0 | 2 | | 2 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 |
| x =10 | 1 | 1 | 0 | 5 | | 5 | 0 | 0 | | 0 | 0 | 1 | | 1 | 0 | 1 | | 1 | 0 |
| Dataset MS3 |  |  |  |  | |  |  |  | |  |  |  | |  |  |  | |  |  |
| x = 0 | 0 | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 |
| x =1 | 2 | 2 | 0 | 2 | | 2 | 0 | 0 | | 0 | 0 | 1 | | 1 | 0 | 1 | | 1 | 0 |
| x =3 | 6 | 6 | 0 | 4 | | 4 | 0 | 2 | | 2 | 0 | 4 | | 4 | 0 | 3 | | 3 | 0 |
| x =5 | 6 | 6 | 0 | 8 | | 8 | 0 | 3 | | 3 | 0 | 5 | | 5 | 0 | 6 | | 6 | 0 |
| x =10 | 7 | 7 | 0 | 11 | | 11 | 0 | 7 | | 7 | 0 | 8 | | 8 | 0 | 8 | | 8 | 0 |
| Dataset MS4 |  |  |  |  | |  |  |  | |  |  |  | |  |  |  | |  |  |
| x = 0 | 1 | 1 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 |
| x =1 | 1 | 1 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 |
| x =3 | 4 | 4 | 0 | 2 | | 2 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 |
| x =5 | 5 | 5 | 0 | 6 | | 6 | 0 | 1 | | 1 | 0 | 0 | | 0 | 0 | 1 | | 1 | 0 |
| x =10 | 13 | 13 | 0 | 14 | | 14 | 0 | 3 | | 3 | 0 | 0 | | 0 | 0 | 4 | | 4 | 0 |

Key; sAIC – covariate selection based on stepwise Akaike information criterion, mBIC - covariate selection based on a stepwise on a modified Bayesian Information Criterion, enet – covariate section based on elastic net regression, MCP – covariate selection based on minimax convex penalty regression, Combi – covariate selection based on a synthesised combination of sAIC, mBIC, enet and MCP, FV – total number of false variables (false positive and negative covariates) selected, FP – number of false positive covariates selected, FN – number of false negative variables selected.

**Table S3. Results of five models incorporating covariate selection stability (with selection thresholds set from x = 0 to x=10), conducted on 4 simulated low signal datasets (Table 1), to illustrate the number of false positive and false negative covariates selected in final models for each method at different threshold values of x.**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | sAIC | | | | mBIC | | | | enet | | | | MCP | | | | Combi | | |
| Dataset LS1 | FV | FP | FN | FV | | FP | FN | FV | | FP | FN | FV | | FP | FN | FV | | FP | FN |
| x = 0 | 2 | 0 | 2 | 1 | | 0 | 1 | 2 | | 0 | 2 | 1 | | 0 | 1 | 0 | | 0 | 0 |
| x =1 | 0 | 0 | 0 | 0 | | 0 | 0 | 1 | | 0 | 1 | 0 | | 0 | 0 | 0 | | 0 | 0 |
| x =3 | 3 | 3 | 0 | 3 | | 3 | 0 | 2 | | 2 | 0 | 3 | | 3 | 0 | 2 | | 2 | 0 |
| x =5 | 3 | 3 | 0 | 5 | | 5 | 0 | 4 | | 4 | 0 | 5 | | 5 | 0 | 5 | | 5 | 0 |
| x =10 | 2 | 0 | 2 | 1 | | 0 | 1 | 2 | | 0 | 2 | 1 | | 0 | 1 | 0 | | 0 | 0 |
| Dataset LS2 |  |  |  |  | |  |  |  | |  |  |  | |  |  |  | |  |  |
| x = 0 | 2 | 0 | 2 | 1 | | 1 | 0 | 2 | | 0 | 2 | 2 | | 0 | 2 | 0 | | 0 | 0 |
| x =1 | 1 | 0 | 1 | 2 | | 2 | 0 | 1 | | 0 | 1 | 2 | | 1 | 1 | 1 | | 1 | 0 |
| x =3 | 3 | 3 | 0 | 3 | | 3 | 0 | 0 | | 0 | 0 | 3 | | 3 | 0 | 3 | | 3 | 0 |
| x =5 | 3 | 3 | 0 | 12 | | 12 | 0 | 3 | | 3 | 0 | 5 | | 5 | 0 | 7 | | 7 | 0 |
| x =10 | 7 | 7 | 0 | 15 | | 15 | 0 | 9 | | 9 | 0 | 8 | | 8 | 0 | 11 | | 11 | 0 |
| Dataset LS3 |  |  |  |  | |  |  |  | |  |  |  | |  |  |  | |  |  |
| x = 0 | 3 | 0 | 3 | 2 | | 2 | 0 | 2 | | 0 | 2 | 1 | | 0 | 1 | 1 | | 1 | 0 |
| x =1 | 3 | 2 | 1 | 2 | | 2 | 0 | 1 | | 0 | 1 | 1 | | 0 | 1 | 1 | | 1 | 0 |
| x =3 | 3 | 3 | 0 | 8 | | 8 | 0 | 4 | | 4 | 0 | 1 | | 1 | 0 | 4 | | 4 | 0 |
| x =5 | 6 | 6 | 0 | 11 | | 11 | 0 | 6 | | 6 | 0 | 2 | | 2 | 0 | 7 | | 7 | 0 |
| x =10 | 8 | 8 | 0 | 15 | | 15 | 0 | 9 | | 9 | 0 | 9 | | 9 | 0 | 9 | | 9 | 0 |
| Dataset LS4 |  |  |  |  | |  |  |  | |  |  |  | |  |  |  | |  |  |
| x = 0 | 2 | 1 | 1 | 2 | | 1 | 1 | 2 | | 0 | 2 | 1 | | 0 | 1 | 1 | | 0 | 1 |
| x =1 | 3 | 2 | 1 | 4 | | 3 | 1 | 2 | | 1 | 1 | 1 | | 0 | 1 | 3 | | 2 | 1 |
| x =3 | 6 | 5 | 1 | 4 | | 3 | 1 | 4 | | 3 | 1 | 3 | | 2 | 1 | 4 | | 3 | 1 |
| x =5 | 8 | 7 | 1 | 4 | | 4 | 0 | 6 | | 5 | 1 | 4 | | 3 | 1 | 6 | | 5 | 1 |
| x =10 | 16 | 15 | 1 | 13 | | 13 | 0 | 9 | | 8 | 1 | 10 | | 9 | 1 | 13 | | 12 | 1 |

Key; sAIC – covariate selection based on stepwise Akaike information criterion, mBIC - covariate selection based on a stepwise on a modified Bayesian Information Criterion, enet – covariate section based on elastic net regression, MCP – covariate selection based on minimax convex penalty regression, Combi – covariate selection based on a synthesised combination of sAIC, mBIC, enet and MCP, FV – total number of false variables (false positive and negative covariates) selected, FP – number of false positive covariates selected, FN – number of false negative variables selected.

**Table S4. Results of five models incorporating covariate selection stability (with selection thresholds set from x = 0 to x=10), conducted on 4 simulated extra-wide datasets (Table 1), to illustrate the number of false positive and false negative covariates selected in final models for each method at different threshold values of x.**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | sAIC | | | | mBIC | | | | enet | | | | MCP | | | | Combi | | |
| Dataset WID1 | FV | FP | FN | FV | | FP | FN | FV | | FP | FN | FV | | FP | FN | FV | | FP | FN |
| x = 0 | 0 | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 |
| x =1 | 0 | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 |
| x =3 | 0 | 0 | 0 | 2 | | 2 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 |
| x =5 | 0 | 0 | 0 | 2 | | 2 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 |
| x =10 | 0 | 0 | 0 | 7 | | 7 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 |
| Dataset WID2 |  |  |  |  | |  |  |  | |  |  |  | |  |  |  | |  |  |
| x = 0 | 0 | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 |
| x =1 | 0 | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 |
| x =3 | 1 | 1 | 0 | 2 | | 2 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 |
| x =5 | 2 | 2 | 0 | 3 | | 3 | 0 | 1 | | 1 | 0 | 0 | | 0 | 0 | 1 | | 1 | 0 |
| x =10 | 3 | 3 | 0 | 7 | | 7 | 0 | 1 | | 1 | 0 | 1 | | 1 | 0 | 2 | | 2 | 0 |
| Dataset WID3 |  |  |  |  | |  |  |  | |  |  |  | |  |  |  | |  |  |
| x = 0 | 0 | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 |
| x =1 | 1 | 1 | 0 | 1 | | 1 | 0 | 1 | | 1 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 |
| x =3 | 3 | 3 | 0 | 2 | | 2 | 0 | 2 | | 2 | 0 | 2 | | 2 | 0 | 2 | | 2 | 0 |
| x =5 | 3 | 3 | 0 | 3 | | 3 | 0 | 2 | | 2 | 0 | 2 | | 2 | 0 | 2 | | 2 | 0 |
| x =10 | 8 | 8 | 0 | 6 | | 6 | 0 | 5 | | 5 | 0 | 4 | | 4 | 0 | 6 | | 6 | 0 |
| Dataset WID4 |  |  |  |  | |  |  |  | |  |  |  | |  |  |  | |  |  |
| x = 0 | 0 | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 |
| x =1 | 1 | 1 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 |
| x =3 | 3 | 3 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 |
| x =5 | 3 | 3 | 0 | 3 | | 3 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 |
| x =10 | 5 | 5 | 0 | 9 | | 9 | 0 | 1 | | 1 | 0 | 0 | | 0 | 0 | 2 | | 2 | 0 |

Key; sAIC – covariate selection based on stepwise Akaike information criterion, mBIC - covariate selection based on a stepwise on a modified Bayesian Information Criterion, enet – covariate section based on elastic net regression, MCP – covariate selection based on minimax convex penalty regression, Combi – covariate selection based on a synthesised combination of sAIC, mBIC, enet and MCP, FV – total number of false variables (false positive and negative covariates) selected, FP – number of false positive covariates selected, FN – number of false negative variables selected.

**Table S5. Results of five models incorporating covariate selection stability (with selection thresholds set from x = 0 to x=10), conducted on 4 simulated binary outcome datasets (Table 1), to illustrate the number of false positive and false negative covariates selected in final models for each method at different threshold values of x.**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | sAIC | | | | mBIC | | | | enet | | | | MCP | | | | Combi | | |
| Dataset BIN1 | FV | FP | FN | FV | | FP | FN | FV | | FP | FN | FV | | FP | FN | FV | | FP | FN |
| x = 0 | 0 | 0 | 0 | 1 | | 1 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 |
| x =1 | 1 | 1 | 0 | 2 | | 2 | 0 | 0 | | 0 | 0 | 1 | | 1 | 0 | 1 | | 1 | 0 |
| x =3 | 1 | 1 | 0 | 4 | | 4 | 0 | 1 | | 1 | 0 | 1 | | 1 | 0 | 1 | | 1 | 0 |
| x =5 | 3 | 3 | 0 | 4 | | 4 | 0 | 2 | | 2 | 0 | 2 | | 2 | 0 | 1 | | 1 | 0 |
| x =10 | 4 | 4 | 0 | 12 | | 12 | 0 | 5 | | 5 | 0 | 2 | | 2 | 0 | 6 | | 6 | 0 |
| Dataset BIN2 |  |  |  |  | |  |  |  | |  |  |  | |  |  |  | |  |  |
| x = 0 | 1 | 0 | 1 | 1 | | 1 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 |
| x =1 | 1 | 0 | 1 | 1 | | 1 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 |
| x =3 | 1 | 0 | 1 | 2 | | 2 | 0 | 0 | | 0 | 0 | 1 | | 1 | 0 | 1 | | 1 | 0 |
| x =5 | 1 | 0 | 1 | 2 | | 2 | 0 | 0 | | 0 | 0 | 1 | | 1 | 0 | 1 | | 1 | 0 |
| x =10 | 4 | 4 | 0 | 9 | | 9 | 0 | 2 | | 2 | 0 | 2 | | 2 | 0 | 2 | | 2 | 0 |
| Dataset BIN3 |  |  |  |  | |  |  |  | |  |  |  | |  |  |  | |  |  |
| x = 0 | 1 | 1 | 0 | 1 | | 1 | 0 | 0 | | 0 | 0 | 0 | | 0 | 0 | 1 | | 1 | 0 |
| x =1 | 1 | 1 | 0 | 2 | | 2 | 0 | 1 | | 1 | 0 | 1 | | 1 | 0 | 1 | | 1 | 0 |
| x =3 | 1 | 1 | 0 | 6 | | 6 | 0 | 1 | | 1 | 0 | 1 | | 1 | 0 | 1 | | 1 | 0 |
| x =5 | 2 | 2 | 0 | 8 | | 8 | 0 | 2 | | 2 | 0 | 1 | | 1 | 0 | 2 | | 2 | 0 |
| x =10 | 3 | 3 | 0 | 8 | | 8 | 0 | 7 | | 7 | 0 | 3 | | 3 | 0 | 4 | | 4 | 0 |
| Dataset BIN4 |  |  |  |  | |  |  |  | |  |  |  | |  |  |  | |  |  |
| x = 0 | 1 | 0 | 1 | 2 | | 1 | 1 | 1 | | 0 | 1 | 1 | | 0 | 1 | 1 | | 0 | 1 |
| x =1 | 1 | 0 | 1 | 3 | | 2 | 1 | 1 | | 0 | 1 | 1 | | 0 | 1 | 1 | | 0 | 1 |
| x =3 | 5 | 4 | 1 | 4 | | 4 | 0 | 2 | | 1 | 1 | 1 | | 0 | 1 | 1 | | 0 | 1 |
| x =5 | 5 | 4 | 1 | 6 | | 6 | 0 | 2 | | 1 | 1 | 1 | | 0 | 1 | 4 | | 3 | 1 |
| x =10 | 9 | 9 | 0 | 20 | | 20 | 0 | 3 | | 2 | 1 | 5 | | 4 | 1 | 8 | | 7 | 1 |

Key; sAIC – covariate selection based on stepwise Akaike information criterion, mBIC - covariate selection based on a stepwise on a modified Bayesian Information Criterion, enet – covariate section based on elastic net regression, MCP – covariate selection based on minimax convex penalty regression, Combi – covariate selection based on a synthesised combination of sAIC, mBIC, enet and MCP, FV – total number of false variables (false positive and negative covariates) selected, FP – number of false positive covariates selected, FN – number of false negative variables selected.

**Table S6. Results of five models incorporating covariate selection stability (with selection thresholds set from x = 0 to x=10), conducted on 4 simulated datasets of increased complexity (Table 1), to illustrate the number of false positive and false negative covariates selected in final models for each method at different threshold values of x.**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | sAIC | | | | mBIC | | | | enet | | | | MCP | | | | Combi | | |
| Dataset IC1 | FV | FP | FN | FV | | FP | FN | FV | | FP | FN | FV | | FP | FN | FV | | FP | FN |
| x = 0 | 4 | 0 | 4 | 1 | | 1 | 0 | 2 | | 0 | 2 | 3 | | 0 | 3 | 3 | | 0 | 3 |
| x =1 | 3 | 0 | 3 | 2 | | 2 | 0 | 1 | | 0 | 1 | 3 | | 0 | 3 | 1 | | 0 | 1 |
| x =3 | 4 | 1 | 3 | 5 | | 5 | 0 | 2 | | 1 | 1 | 3 | | 0 | 3 | 4 | | 3 | 1 |
| x =5 | 4 | 4 | 0 | 6 | | 6 | 0 | 2 | | 2 | 0 | 6 | | 3 | 3 | 4 | | 4 | 0 |
| x =10 | 6 | 6 | 0 | 6 | | 6 | 0 | 5 | | 5 | 0 | 8 | | 5 | 3 | 7 | | 7 | 0 |
| Dataset IC2 |  |  |  |  | |  |  |  | |  |  |  | |  |  |  | |  |  |
| x = 0 | 3 | 0 | 3 | 2 | | 0 | 2 | 3 | | 0 | 3 | 4 | | 0 | 4 | 3 | | 0 | 3 |
| x =1 | 2 | 0 | 2 | 1 | | 0 | 1 | 3 | | 0 | 3 | 4 | | 0 | 4 | 3 | | 0 | 3 |
| x =3 | 4 | 2 | 2 | 5 | | 4 | 1 | 4 | | 1 | 3 | 5 | | 1 | 4 | 2 | | 0 | 2 |
| x =5 | 10 | 8 | 2 | 11 | | 10 | 1 | 4 | | 1 | 3 | 4 | | 1 | 3 | 2 | | 1 | 1 |
| x =10 | 20 | 19 | 1 | 18 | | 18 | 0 | 6 | | 5 | 1 | 9 | | 7 | 2 | 6 | | 5 | 1 |

Key; sAIC – covariate selection based on stepwise Akaike information criterion, mBIC - covariate selection based on a stepwise on a modified Bayesian Information Criterion, enet – covariate section based on elastic net regression, MCP – covariate selection based on minimax convex penalty regression, Combi – covariate selection based on a synthesised combination of sAIC, mBIC, enet and MCP, FV – total number of false variables (false positive and negative covariates) selected, FP – number of false positive covariates selected, FN – number of false negative variables selected.