Appendix C Power Comparison of the Tests of Fixed Effect Models and SKAT/SKAT-O/SKAT-C

For each setting of empirical power calculations, 1,000 datasets were simulated to calculate the empirical power levels as the proportion of p-values which are smaller than a given α level.

Quantitative Trait Cases When the Constants k are Given by Relations (12). For the scenario that some causal variants are rare and some are common, the empirical power levels are plotted in Figures S.1 - S.5 for subregions of size 3, 6, 9, 12, and 15 kb, respectively. For the scenario that all causal variants are rare, the empirical power levels are plotted in Figures S.6 - S.10 for subregions of size 3, 6, 9, 12, and 15 kb, respectively.

For subregions of size 3 kb, the empirical power levels are presented in Figures S.1 and S.6, and we observe the following: (a) the power levels of approximate F-distribution tests of the multivariate additive models (1) are higher than those of the FLM (7), (b) the power levels of F-distribution tests of univariate additive model and FLM for y_{i1} are higher than those of SKAT/SKAT-O/SKAT-C in Figures S.1 and S.6.

For subregions of size 6 kb, the empirical power levels are presented in Figures S.2 and S.7, and we observe the following: (a) the power levels of approximate F-distribution tests of the multivariate additive models (1) are higher than those of the FLM (7), (b) the power levels of F-distribution tests of univariate additive model for y_{i1} are higher than those of SKAT/SKAT-O/SKAT-C in Figures S.2 and S.7, and (c) the power levels of F-distribution tests of univariate FLM for y_{i1} are higher than SKAT and SKAT-O but similar to SKAT-C in Figure S.2 and similar to SKAT/SKAT-O/SKAT-C in Figure S.7.

For subregions of size 9 kb, the empirical power levels are presented in Figures S.3 and S.8, and we note the following: (a) the power levels of approximate F-distribution tests of the multivariate additive models (1) are higher than those of the FLM (7), (b) the power levels of F-distribution tests of both univariate models (1) and (7) for y_{i1} are higher than those of SKAT/SKAT-O but tend to be lower than those SKAT-C in Figure S.3, (c) the power levels of F-distribution tests of univariate additive models (1) are similar to SKAT/SKAT-O/SKAT-C, and the power levels of F-distribution tests of univariate FLM for y_{i1} are lower than SKAT/SKAT-O/SKAT-C in Figure S.8.

For subregions of size 12 kb, the empirical power levels are presented in Figures S.4 and S.9. The power levels of approximate F-distribution tests of the multivariate additive models (1) are similar to or lower than those of the FLM (7) in Figure S.4 when some causal variants are rare and some are common, and are higher than those of the FLM (7) in Figure S.9 when all causal variants are rare. In Figure S.4, the power levels of F-distribution tests univariate functional (or additive) linear model for y_{i1} are higher than (or similar to) SKAT/SKAT-O but lower than SKAT-C; and in Figure S.9, the power levels of F-distribution tests of both univariate fixed effect models for y_{i1} are lower than SKAT/SKAT-O/SKAT-C.

For subregions of size 15 kb, the empirical power levels are presented in Figures S.5 and S.10. The power levels of approximate F-distribution tests of the multivariate additive models (1) are generally lower than those of the FLM (7) in Figure S.5 when some causal variants are rare and some are common, and are higher than those of the FLM (7) in Figure S.10 when all causal variants are rare. In Figure S.5, the power levels of F-distribution tests univariate functional (or additive) linear model for y_{i1} are higher (or lower) than SKAT and SKAT-O but lower than SKAT-C; and in Figure S.10, the power levels of F-distribution tests of both univariate fixed effect models for y_{i1} are lower than SKAT/SKAT-O/SKAT-C.

Dichotomous Trait Cases When the Constants k are Given by Relations (14). For the Rao's efficient score tests of additive effect logistic regression model (2) and GFLM (8) and SKAT/SKAT-O/SKAT-C, the empirical power levels are presented in Figures S.11 - S.20. When the sizes of subregions are 3 or 6 kb, the Rao's efficient score test of additive effect logistic regression model (2) and GFLM (8) perform well and similarly in Figures S.11, S.12, S.16, and S.17. When the size of subregions is 9 kb, the Rao's efficient score test of both additive effect logistic regression model (2) and GFLM (8) perform well in Figure S.13, and the Rao's efficient score test of additive effect logistic regression model (2) is generally similar to or less powerful than that of GFLM (8) in Figure S.18. When the size of subregions is 12 or 15 kb, the Rao's efficient score test of additive effect logistic regression model (2) is less powerful than that of GFLM (8) in Figures S.14, S.15, S.19 and S.20, respectively.

When some causal variants are rare and some are common, the Rao's efficient score tests of both fixed effect logistic regression models (2) and GFLM (8) in Figures S.11, S.12, and S.13 are generally similar to or more powerful than SKAT/SKAT-O/SKAT-C when the sizes of subregions are 3, 6, or 9 kb. The Rao's efficient score test of additive effect logistic regression model (2) is less powerful than SKAT/SKAT-O/SKAT-C, and the Rao's efficient score tests of GFLM (8) in Figures S.14 and S.15 are generally similar to or more powerful than SKAT/SKAT-O but less powerful than SKAT-C if size of subregions is 12 or 15 kb.

When all causal variants are rare and size of the subregions is 3, 6, or 9 kb, SKAT/SKAT-O/SKAT-C are generally similar to or slightly more powerful than the Rao's efficient score tests in Figures S.16, S.17, and S.18, respectively. When all causal variants are rare and the size of subregions is 12 or 15 kb, SKAT/SKAT-O/SKAT-C are more powerful than the Rao's efficient score tests in Figures S.19 and S.20.

Summary of Power Performance. In Tables S.17 and S.18, we provide a summary of power performance comparisons among the tests of two fixed models and SKAT/SKAT-O/SKAT-C. The comparison is made for each Figure separately. For instance, the test of additive models (1) performs the best, the test of the FLM (7) performs better, and SKAT/KAT-O/SKAT-C perform well in Figures S.1 and S.2. In Figure S.7, the test of additive models (1) performs better than the test of the FLM (7) and SKAT/SKAT-O/SKAT-C, while the test of the FLM (7) perform similarly to SKAT/SKAT-O/SKAT-C.

The tests of fixed effect models perform similarly to or better than SKAT/SKAT-O/SKAT-C in most cases. Either one or both of the two fixed effect models perform better than SKAT/SKAT-O/SKAT-C of the mixed models except for the 12 and 15 kb region sizes. When some causal variants are rare and some are common, the multivariate additive model (1) of quantitative traits performs better than the FLM (7) if genetic region size is 3, 6, or 9 kb, while the FLM (7) performs better if genetic region size is 12 or 15 kb. When all causal variants are rare, the multivariate additive model (1) of quantitative traits performs better than the FLM (7). The additive effect logistic regression model (2) of dichotomous traits performs well if genetic region size is 3 or 6 kb. As the genetic region sizes increase from 9 kb to 15 kb, the GFLM (8) tends to perform better than the additive effect model (2).

Table S.17: **Power performance comparison of the approximate F-distribution tests of fixed models and SKAT/SKAT-O/SKAT-C of quantitative traits.** The comparison is made Figure by Figure to compare the performance of three models: (I) linear additive effect model (1), (II) FLM (7), and (III) mixed models of SKAT/SKAT-O/SKAT-C. Since SKAT/SKAT-O/SKAT-C do not have multivariate versions, the comparison of the two fixed models of (1) and (7) with the mixed model is made only for the univariate case.

Type of Causal Variants	**Model or Test**	**Region Size and Figure**				
		3 kb	**6 kb**	**9 kb**	**12 kb**	**15 kb**
		Figure S.1	**Figure S.2**	**Figure S.3**	**Figure S.4**	**Figure S.5**
Common and Rare	**Additive Model (1)**	best	best	best	good	good
	FLM (7)	better	better	better	better	better
	SKAT and SKAT-O	good	good	good	good	better
	SSKAT-C	good	better	best	best	best
		Figure S.6	**Figure S.7**	**Figure S.8**	**Figure S.9**	**Figure S.10**
Rare	**Additive Model (1)**	best	better	better	better	better
	FLM (7)	better	good	good	good	good
	SKAT and SKAT-O	good	good	better	best	best
	SKAT-C	good	good	better	best	best

Table S.18: **Power performance comparison of the Rao's efficient score tests of fixed models and SKAT/SKAT-O/SKAT-C of dichotomous traits.** The comparison is made Figure by Figure to compare the performance of three models: (I) logistic additive effect model (2), (II) GFLM (8), and (III) mixed models of SKAT/SKAT-O/SKAT-C.

Type of Causal Variants	**Model or Test**	**Region Size and Figure**				
		3 kb	**6 kb**	**9 kb**	**12 kb**	**15 kb**
		Figure S.11	**Figure S.12**	**Figure S.13**	**Figure S.14**	**Figure S.15**
Common and Rare	**Additive Model (2)**	best	better	better	good	good
	GFLM (8)	good	better	better	better	better
	SKAT and SKAT-O	good	good	good	better	better
	SKAT-C	better	better	better	best	best
		Figure S.16	**Figure S.17**	**Figure S.18**	**Figure S.19**	**Figure S.20**
Rare	**Additive Model (2)**	good	good	good	good	good
	GFLM (8)	good	good	better	better	better
	SKAT and SKAT-O	good	good	better	best	best
	SKAT-C	good	good	better	best	best

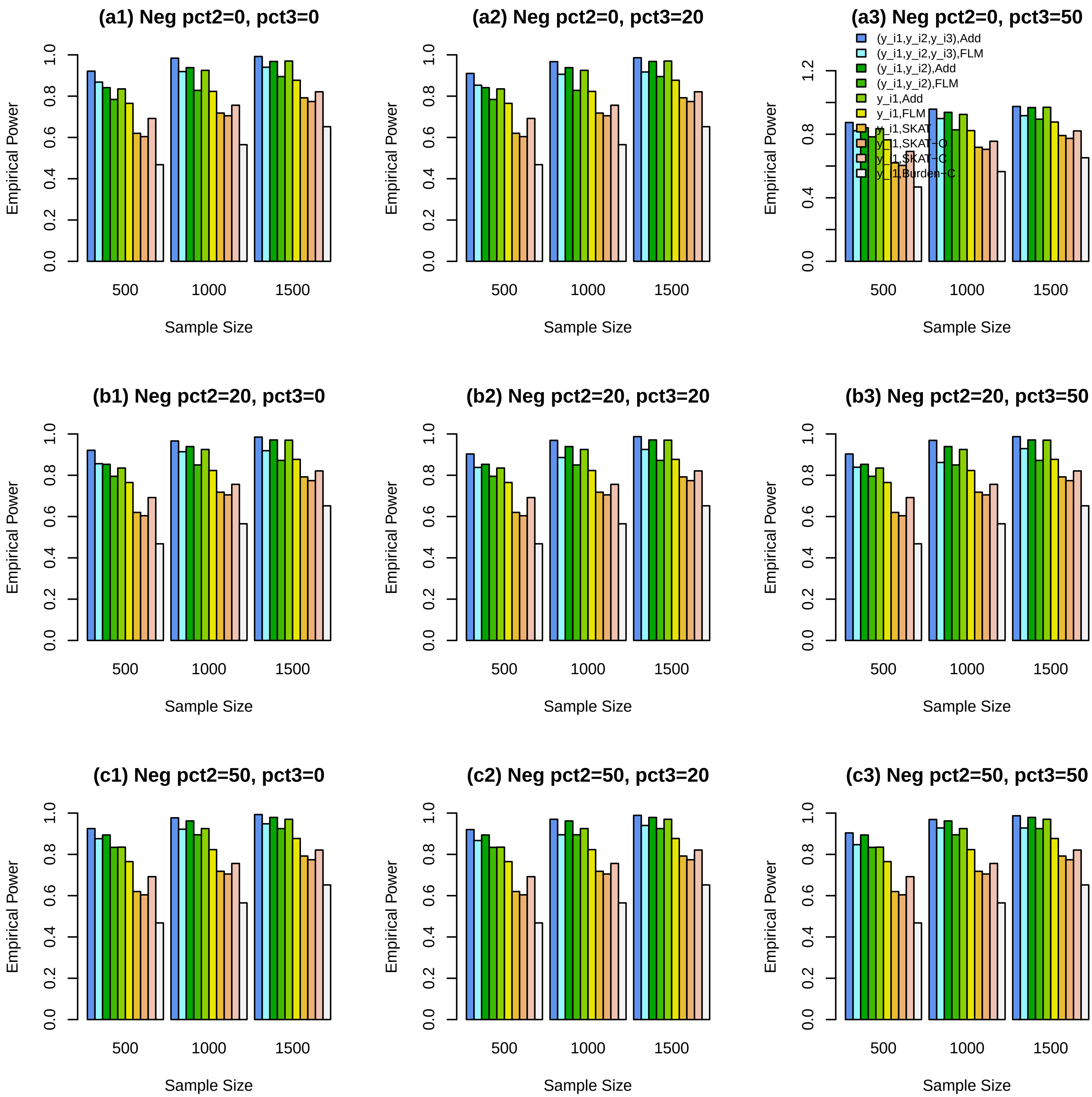

Figure S.1: **The Empirical Power of the Approximate F-distribution Tests of the Multivariate Additive Model (1) and the Functional Linear Model (7) Using B-spline Basis Based on Pillai-Bartlett Trace, SKAT, SKAT-O, SKAT-C, and Burden-C at $\alpha = 0.01$, When Some Causal Variants are Rare and Some are Common, 10% of the Variants Were Causal, the Region Size is 3 kb, and the Constant $k = 1.0$ in Genetic Effect Size $|\beta_{ij}|$.** For the trait y_{i1}, 20%/80% causal variants had negative/positive effects; pct2 represents the percentage of negative effect causal variants for trait y_{i2}; and pct3 represents the percentage of negative effect causal variants for trait y_{i3}.

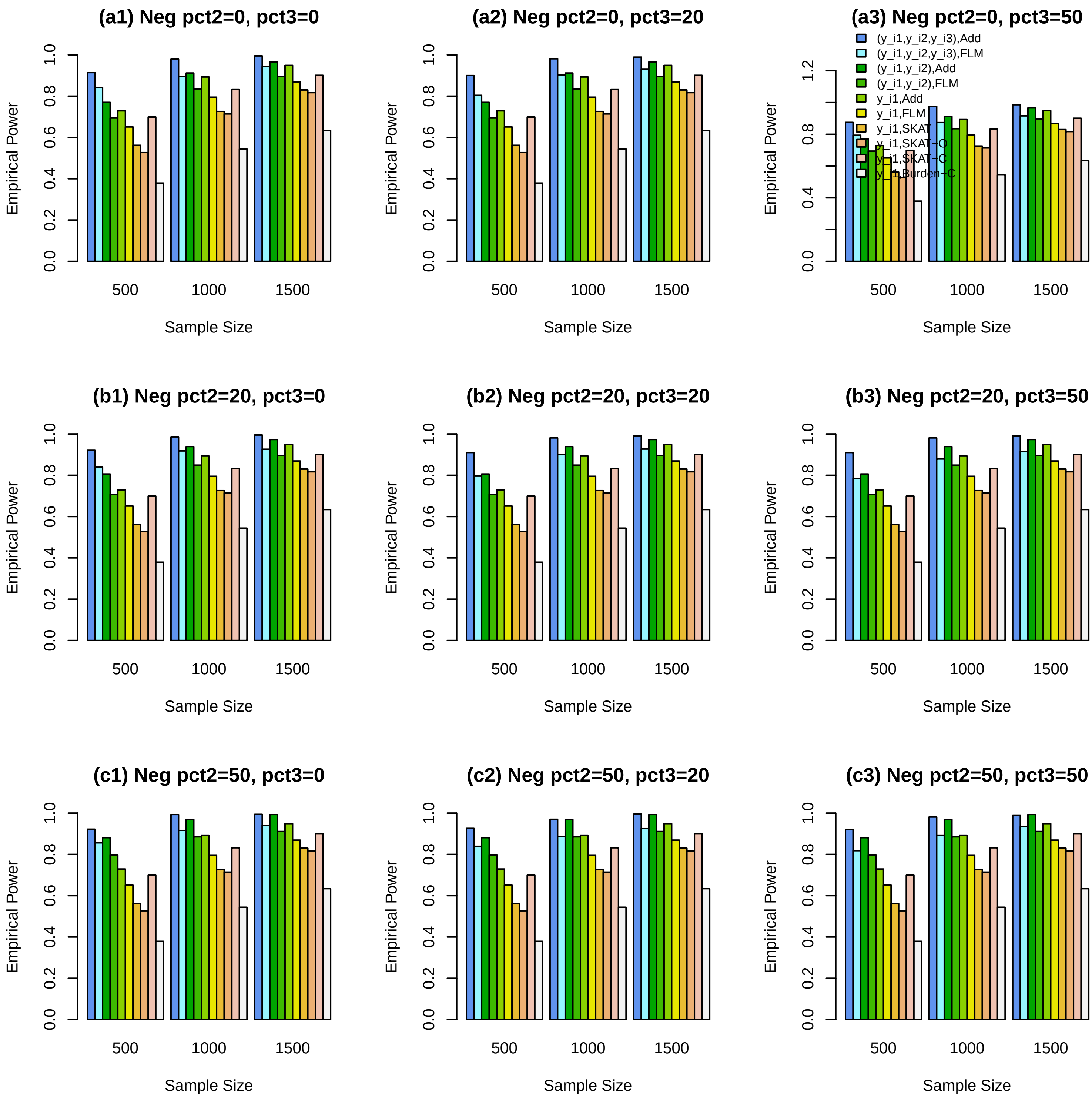

Figure S.2: **The Empirical Power of the Approximate F-distribution Tests of the Multivariate Additive Model (1) and the Functional Linear Model (7) Using B-spline Basis Based on Pillai-Bartlett Trace, SKAT, SKAT-O, SKAT-C, and Burden-C at $\alpha = 0.01$, When Some Causal Variants are Rare and Some are Common, 10% of the Variants Were Causal, the Region Size is 6 kb, and the Constant $k = 2.0$ in Genetic Effect Size $|\beta_{ij}|$.** For the trait y_{i1}, 20%/80% causal variants had negative/positive effects; pct2 represents the percentage of negative effect causal variants for trait y_{i2}; and pct3 represents the percentage of negative effect causal variants for trait y_{i3}.

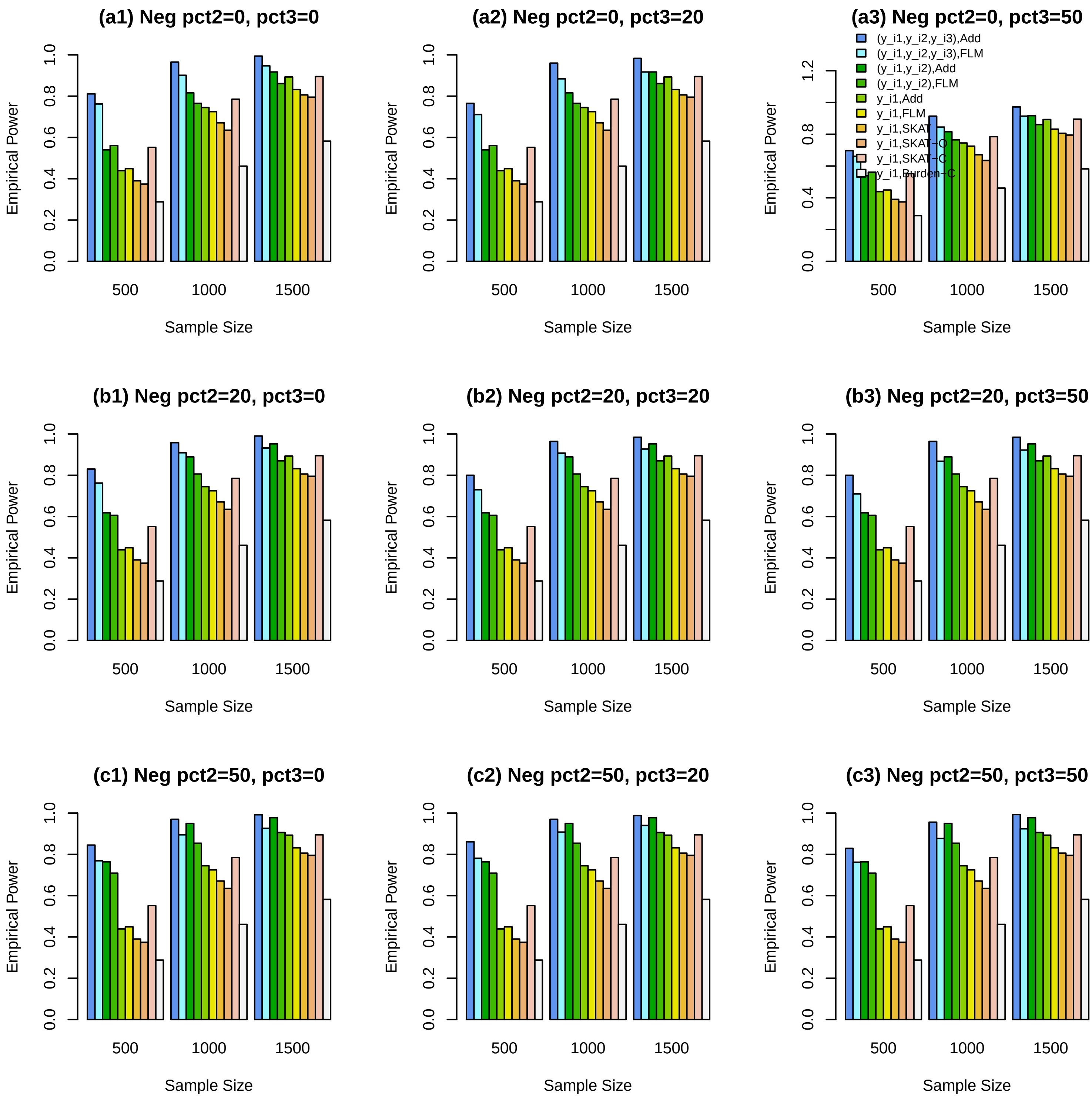

Figure S.3: **The Empirical Power of the Approximate F-distribution Tests of the Multivariate Additive Model (1) and the Functional Linear Model (7) Using B-spline Basis Based on Pillai-Bartlett Trace, SKAT, SKAT-O, SKAT-C, and Burden-C at $\alpha = 0.01$, When Some Causal Variants are Rare and Some are Common, 10% of the Variants Were Causal, the Region Size is 9 kb, and the Constant $k = 3.0$ in Genetic Effect Size $|\beta_{ij}|$.** For the trait y_{i1}, 20%/80% causal variants had negative/positive effects; pct2 represents the percentage of negative effect causal variants for trait y_{i2}; and pct3 represents the percentage of negative effect causal variants for trait y_{i3}.

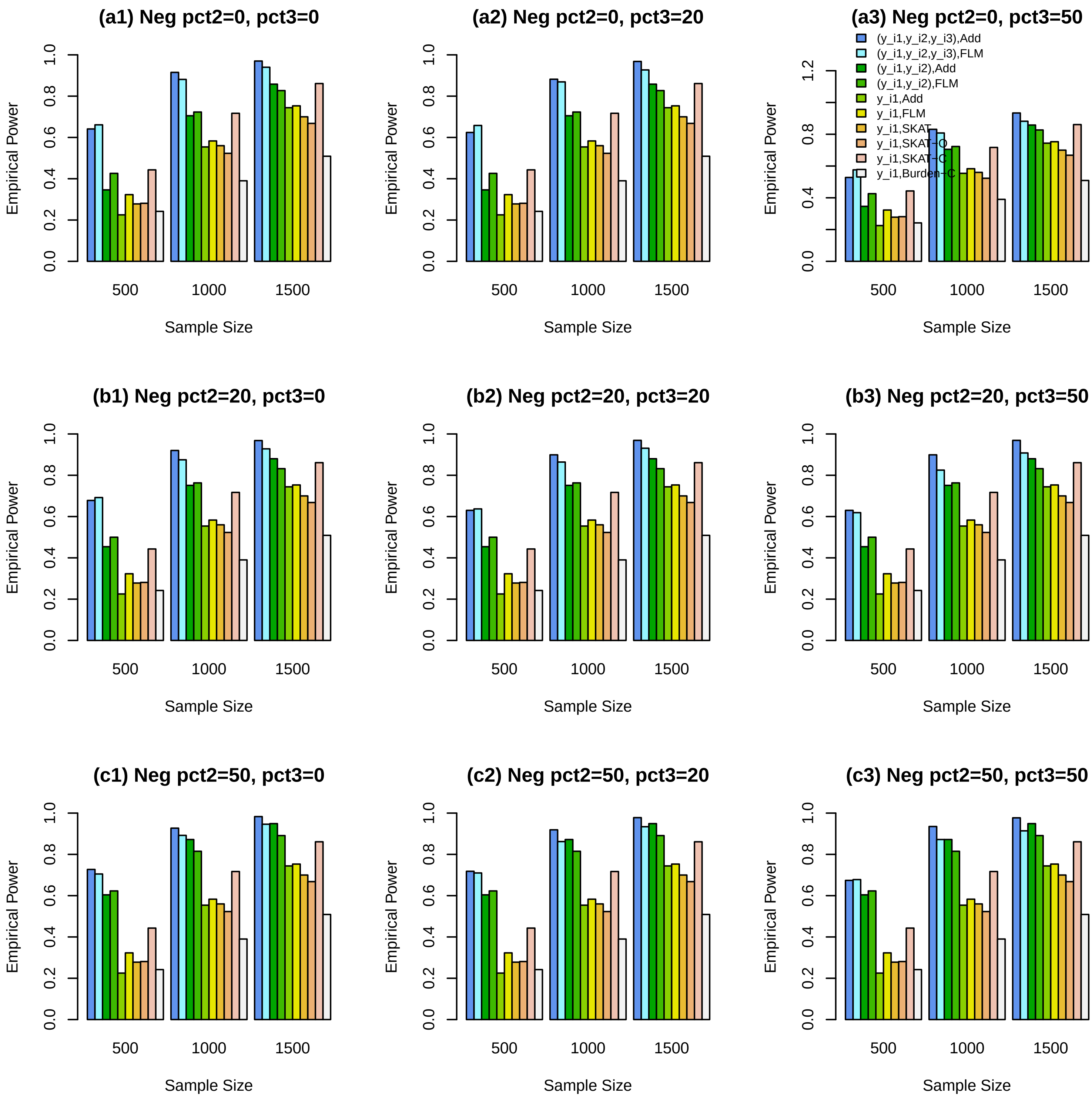

Figure S.4: **The Empirical Power of the Approximate F-distribution Tests of the Multivariate Additive Model (1) and the Functional Linear Model (7) Using B-spline Basis Based on Pillai-Bartlett Trace, SKAT, SKAT-O, SKAT-C, and Burden-C at $\alpha = 0.01$, When Some Causal Variants are Rare and Some are Common, 10% of the Variants Were Causal, the Region Size is 12 kb, and the Constant $k = 4.0$ in Genetic Effect Size $|\beta_{ij}|$.** For the trait y_{i1}, 20%/80% causal variants had negative/positive effects; pct2 represents the percentage of negative effect causal variants for trait y_{i2}; and pct3 represents the percentage of negative effect causal variants for trait y_{i3}.

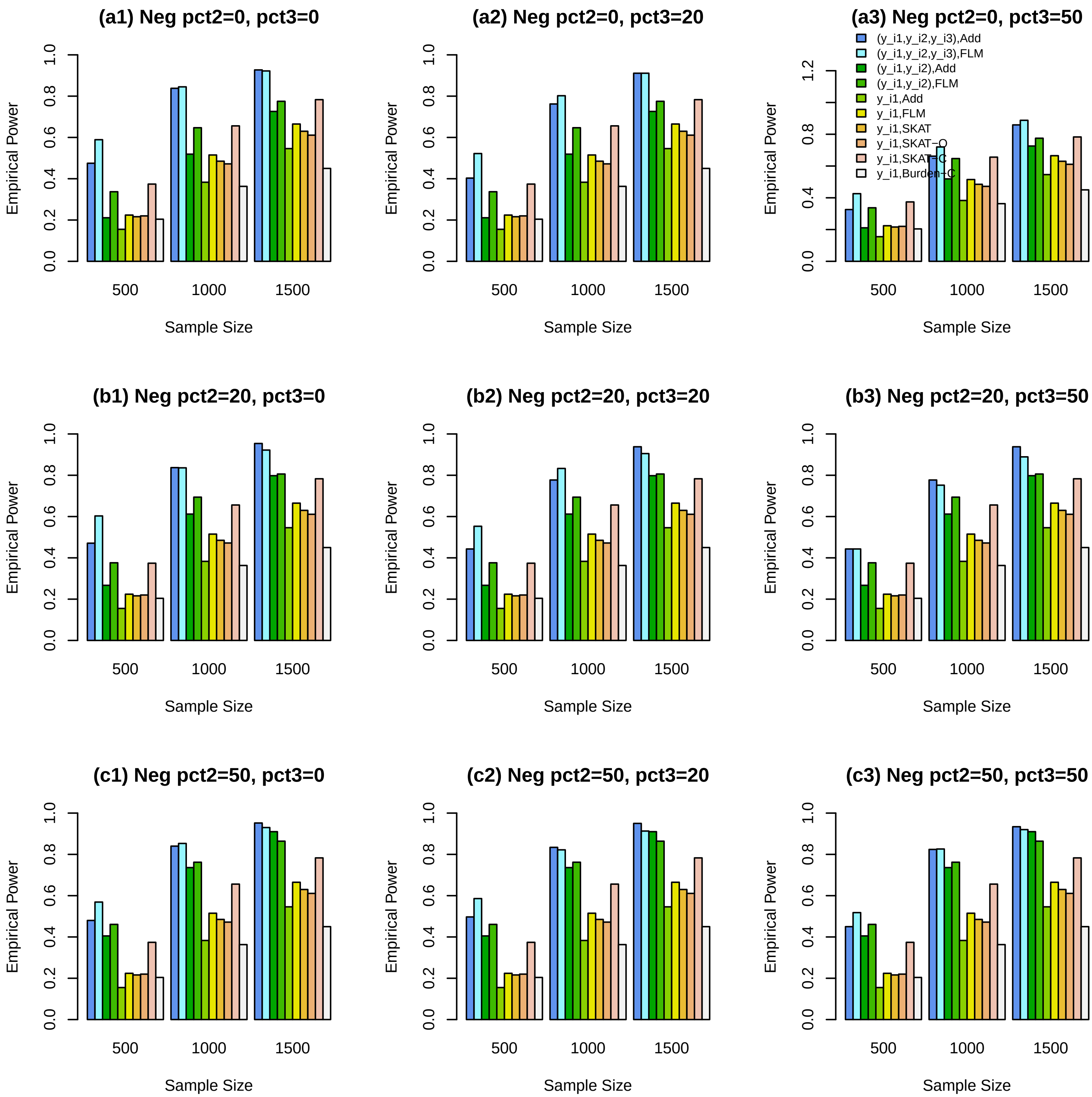

Figure S.5: **The Empirical Power of the Approximate F-distribution Tests of the Multivariate Additive Model (1) and the Functional Linear Model (7) Using B-spline Basis Based on Pillai-Bartlett Trace, SKAT, SKAT-O, SKAT-C, and Burden-C at $\alpha = 0.01$, When Some Causal Variants are Rare and Some are Common, 10% of the Variants Were Causal, the Region Size is 15 kb, and the Constant $k = 5.0$ in Genetic Effect Size $|\beta_{ij}|$.** For the trait y_{i1}, 20%/80% causal variants had negative/positive effects; pct2 represents the percentage of negative effect causal variants for trait y_{i2}; and pct3 represents the percentage of negative effect causal variants for trait y_{i3}.

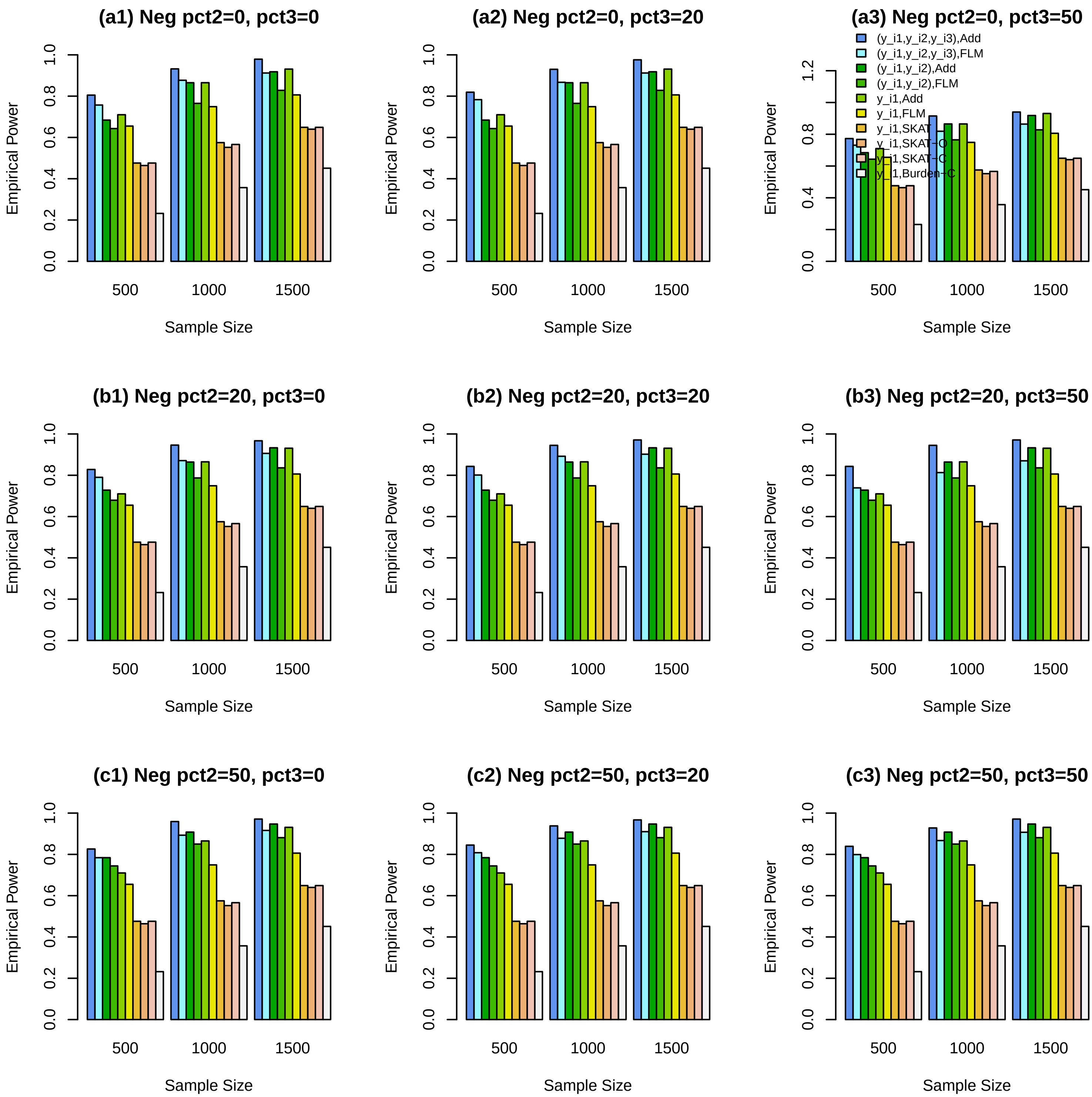

Figure S.6: **The Empirical Power of the Approximate F-distribution Tests of the Multivariate Additive Model (1) and the Functional Linear Model (7) Using B-spline Basis Based on Pillai-Bartlett Trace, SKAT, SKAT-O, SKAT-C, and Burden-C at $\alpha = 0.01$, When All Causal Variants are Rare, 10% of the Variants Were Causal, the Region Size is 3 kb, and the Constant $k = 1.0$ in Genetic Effect Size $|\beta_{ij}|$.** For the trait y_{i1}, 20%/80% causal variants had negative/positive effects; pct2 represents the percentage of negative effect causal variants for trait y_{i2}; and pct3 represents the percentage of negative effect causal variants for trait y_{i3}.

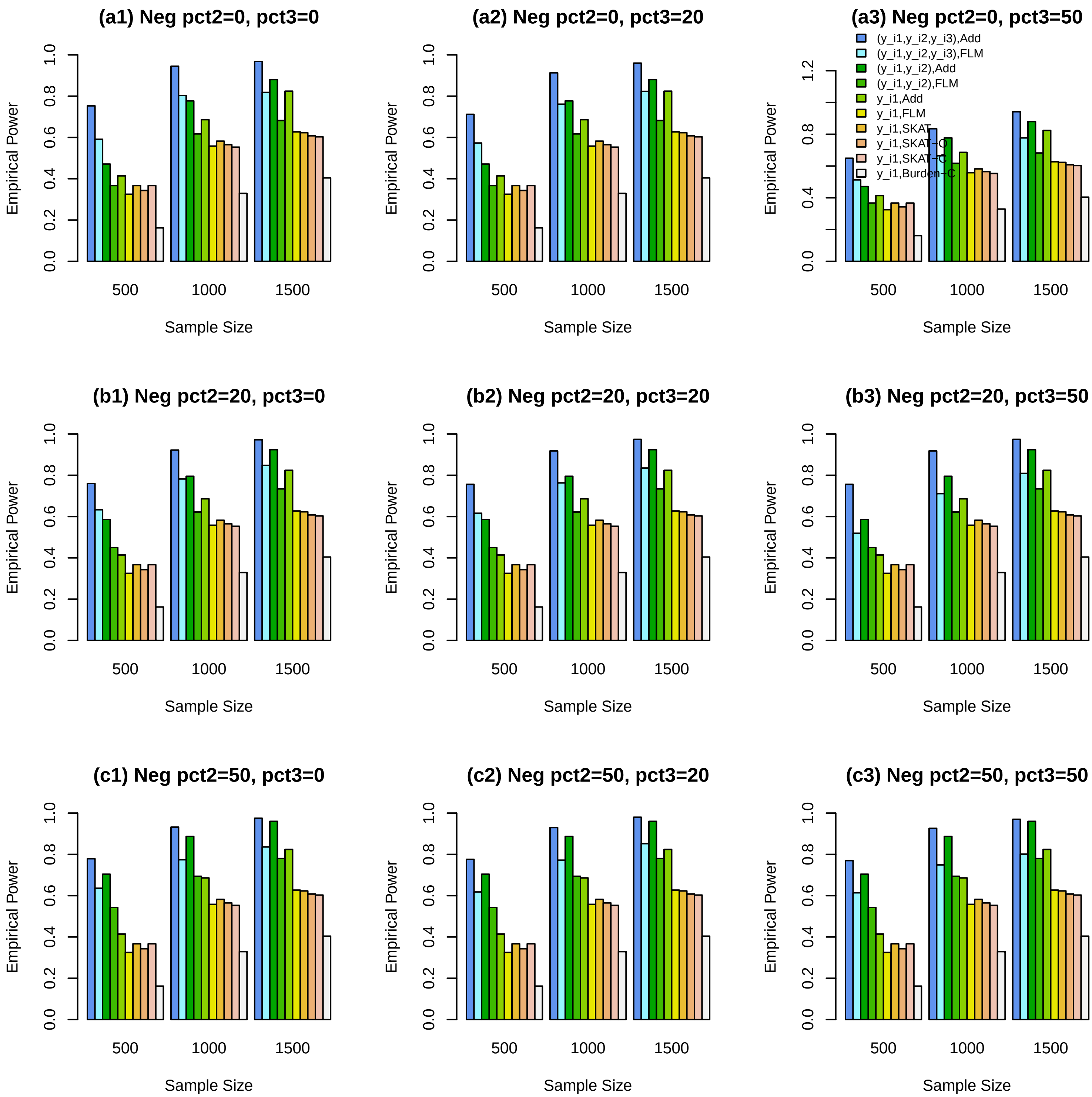

Figure S.7: **The Empirical Power of the Approximate F-distribution Tests of the Multivariate Additive Model (1) and the Functional Linear Model (7) Using B-spline Basis Based on Pillai-Bartlett Trace, SKAT, SKAT-O, SKAT-C, and Burden-C at $\alpha = 0.01$, When All Causal Variants are Rare, 10% of the Variants Were Causal, the Region Size is 6 kb, and the Constant $k = 2.0$ in Genetic Effect Size $|\beta_{ij}|$.** For the trait y_{i1}, 20%/80% causal variants had negative/positive effects; pct2 represents the percentage of negative effect causal variants for trait y_{i2}; and pct3 represents the percentage of negative effect causal variants for trait y_{i3}.

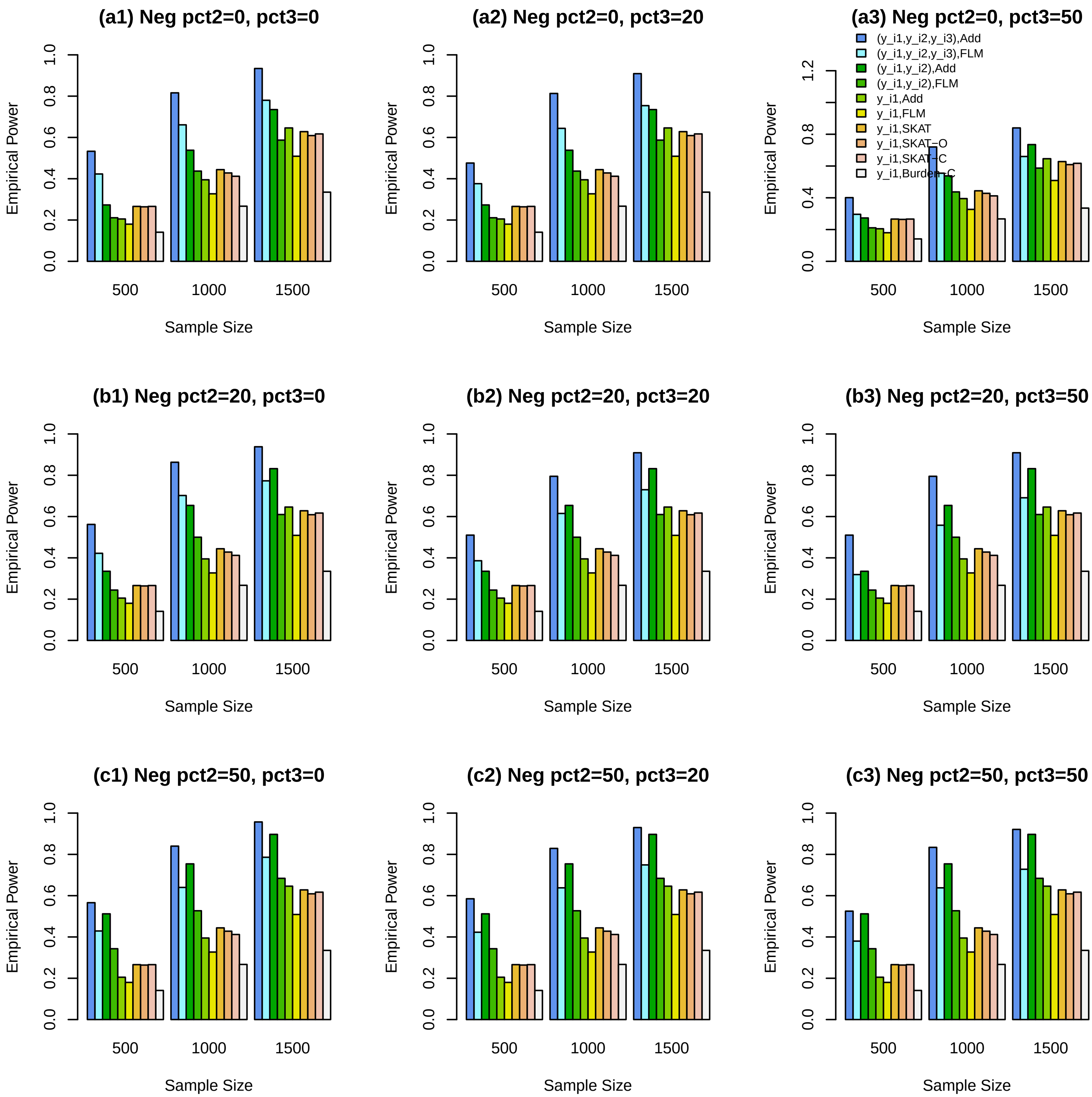

Figure S.8: **The Empirical Power of the Approximate F-distribution Tests of the Multivariate Additive Model (1) and the Functional Linear Model (7) Using B-spline Basis Based on Pillai-Bartlett Trace, SKAT, SKAT-O, SKAT-C, and Burden-C at $\alpha = 0.01$, When All Causal Variants are Rare, 10% of the Variants Were Causal, the Region Size is 9 kb, and the Constant $k = 3.0$ in Genetic Effect Size $|\beta_{ij}|$.** For the trait y_{i1}, 20%/80% causal variants had negative/positive effects; pct2 represents the percentage of negative effect causal variants for trait y_{i2}; and pct3 represents the percentage of negative effect causal variants for trait y_{i3}.

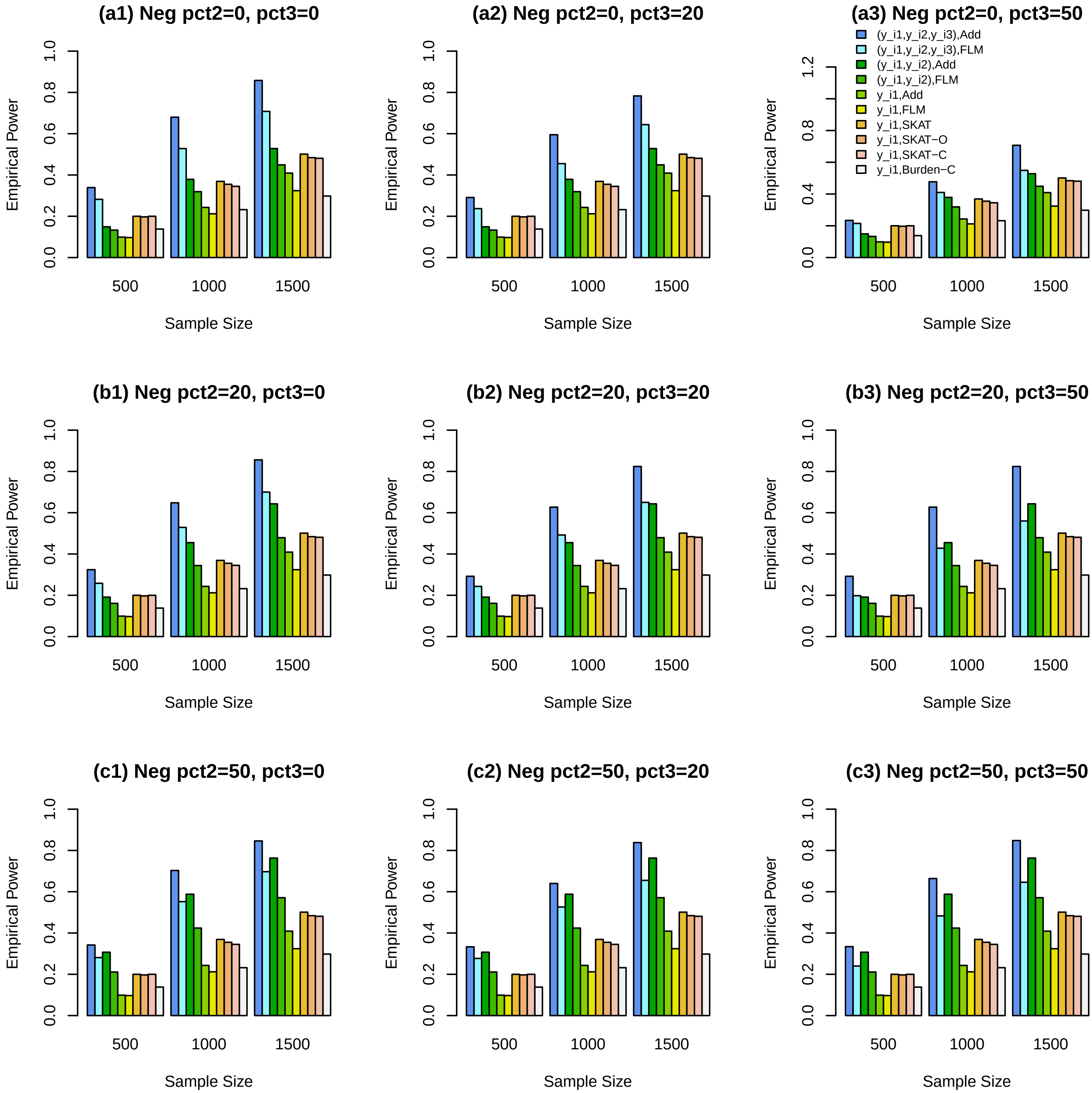

Figure S.9: **The Empirical Power of the Approximate F-distribution Tests of the Multivariate Additive Model (1) and the Functional Linear Model (7) Using B-spline Basis Based on Pillai-Bartlett Trace, SKAT, SKAT-O, SKAT-C, and Burden-C at $\alpha = 0.01$, When All Causal Variants are Rare, 10% of the Variants Were Causal, the Region Size is 12 kb, and the Constant $k = 4.0$ in Genetic Effect Size $|\beta_{ij}|$.** For the trait y_{i1}, 20%/80% causal variants had negative/positive effects; pct2 represents the percentage of negative effect causal variants for trait y_{i2}; and pct3 represents the percentage of negative effect causal variants for trait y_{i3}.

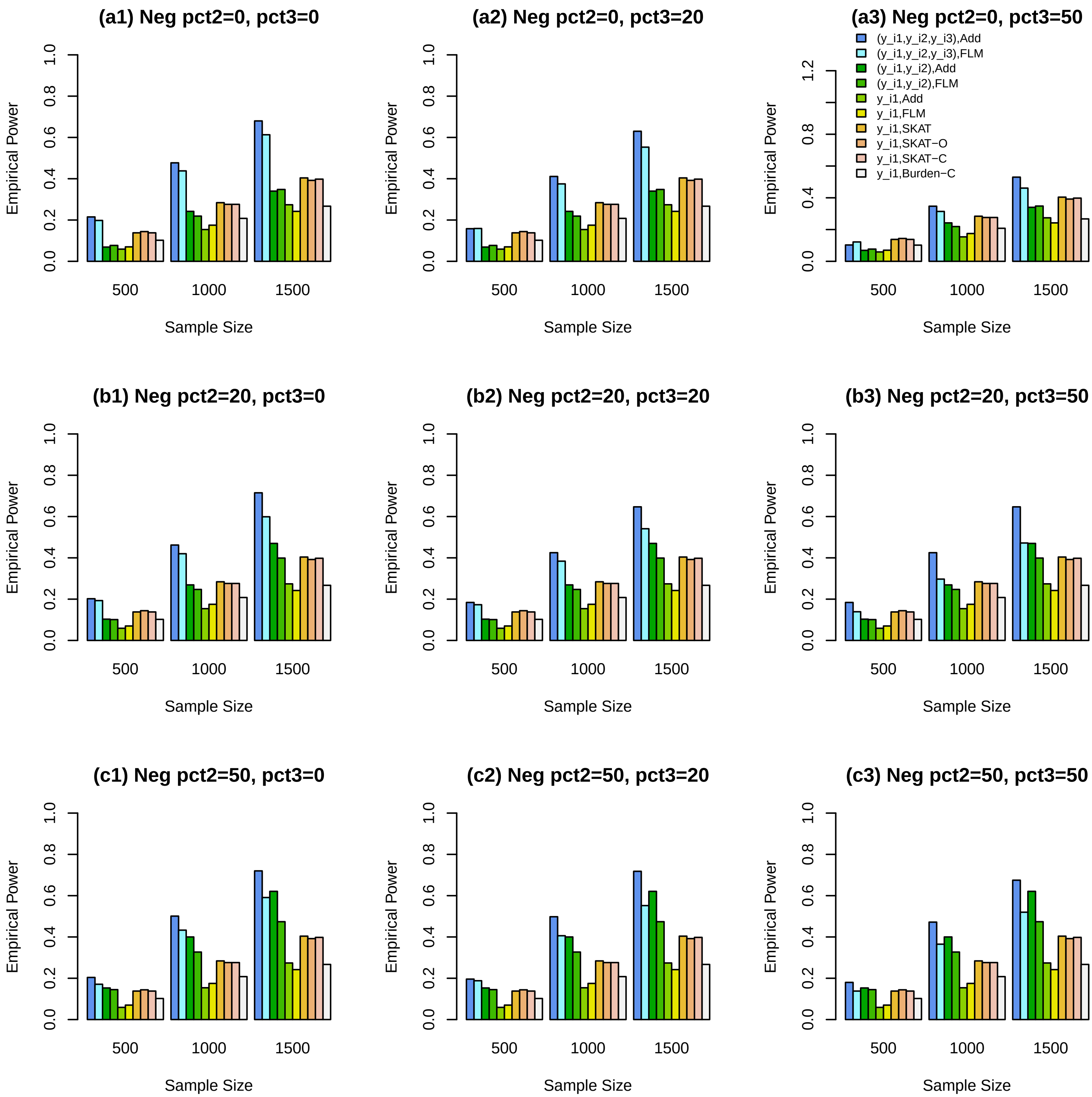

Figure S.10: **The Empirical Power of the Approximate F-distribution Tests of the Multivariate Additive Model (1) and the Functional Linear Model (7) Using B-spline Basis Based on Pillai-Bartlett Trace, SKAT, SKAT-O, SKAT-C, and Burden-C at $\alpha = 0.01$, When All Causal Variants are Rare, 10% of the Variants Were Causal, the Region Size is 15 kb, and the Constant $k = 5.0$ in Genetic Effect Size $|\beta_{ij}|$.** For the trait y_{i1}, 20%/80% causal variants had negative/positive effects; pct2 represents the percentage of negative effect causal variants for trait y_{i2}; and pct3 represents the percentage of negative effect causal variants for trait y_{i3}.

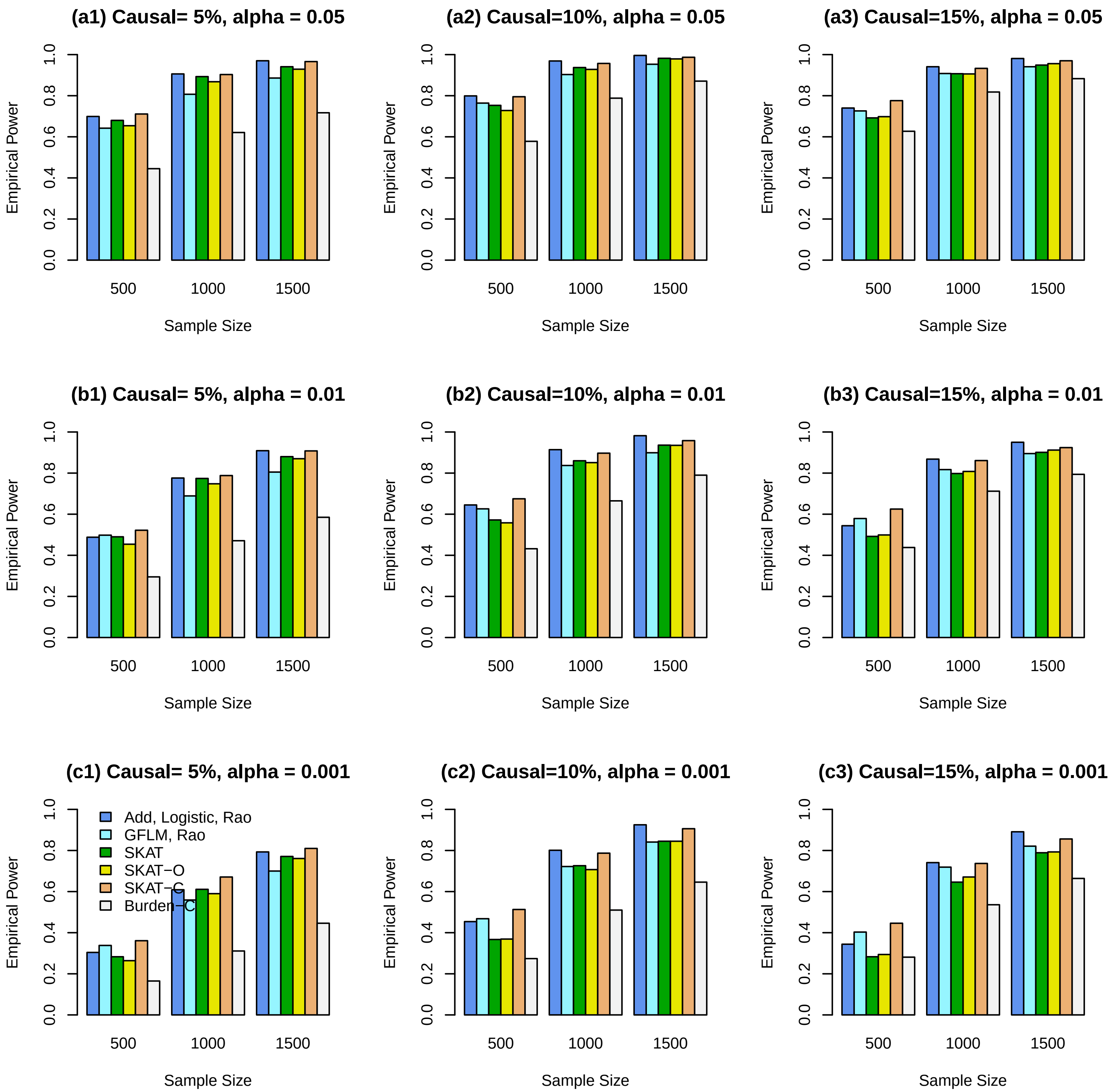

Figure S.11: **The Empirical Power of the Rao's Efficient Score Tests of Additive Logistic Regression Model (2) and the GFLM (8) Using B-spline Basis, SKAT, SKAT-O, SKAT-C, and Burden-C, When Some Causal Variants are Rare and Some are Common, the Region Size is 3 kb, the Constant $k = 1.0$ in Genetic Effect Size $|\beta_{ij}|$, and 20%/80% Causal Variants Had Negative/Positive Effects.** The order of B-spline basis was 4, and the number of B-spline basis functions was $K = K_\beta = 10$.

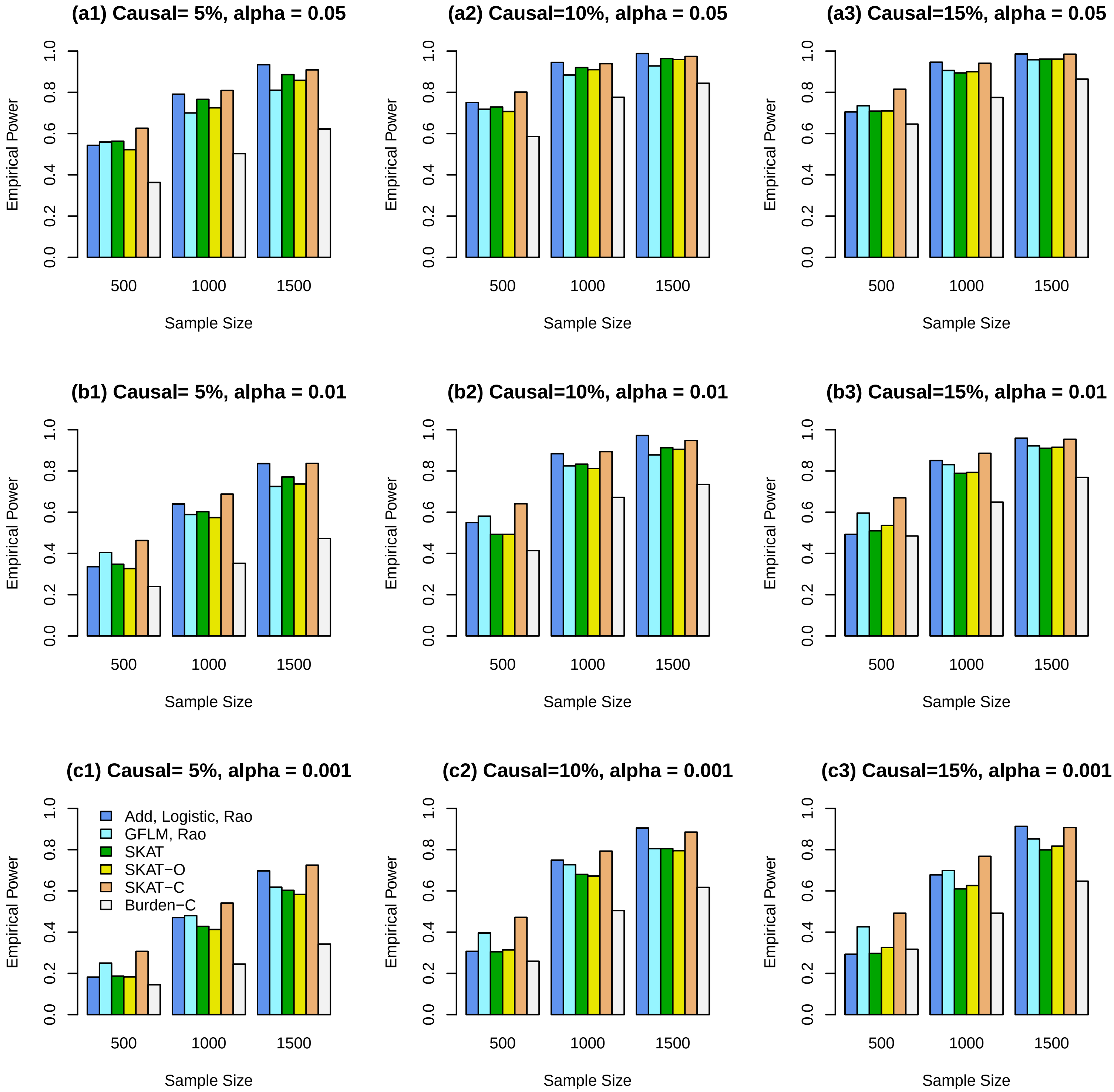

Figure S.12: **The Empirical Power of the Rao's Efficient Score Tests of Additive Logistic Regression Model (2) and the GFLM (8) Using B-spline Basis, SKAT, SKAT-O, SKAT-C, and Burden-C, When Some Causal Variants are Rare and Some are Common, the Region Size is 6 kb, the Constant $k = 1.25$ in Genetic Effect Size $|\beta_{ij}|$, and 20%/80% Causal Variants Had Negative/Positive Effects.** The order of B-spline basis was 4, and the number of B-spline basis functions was $K = K_\beta = 10$.

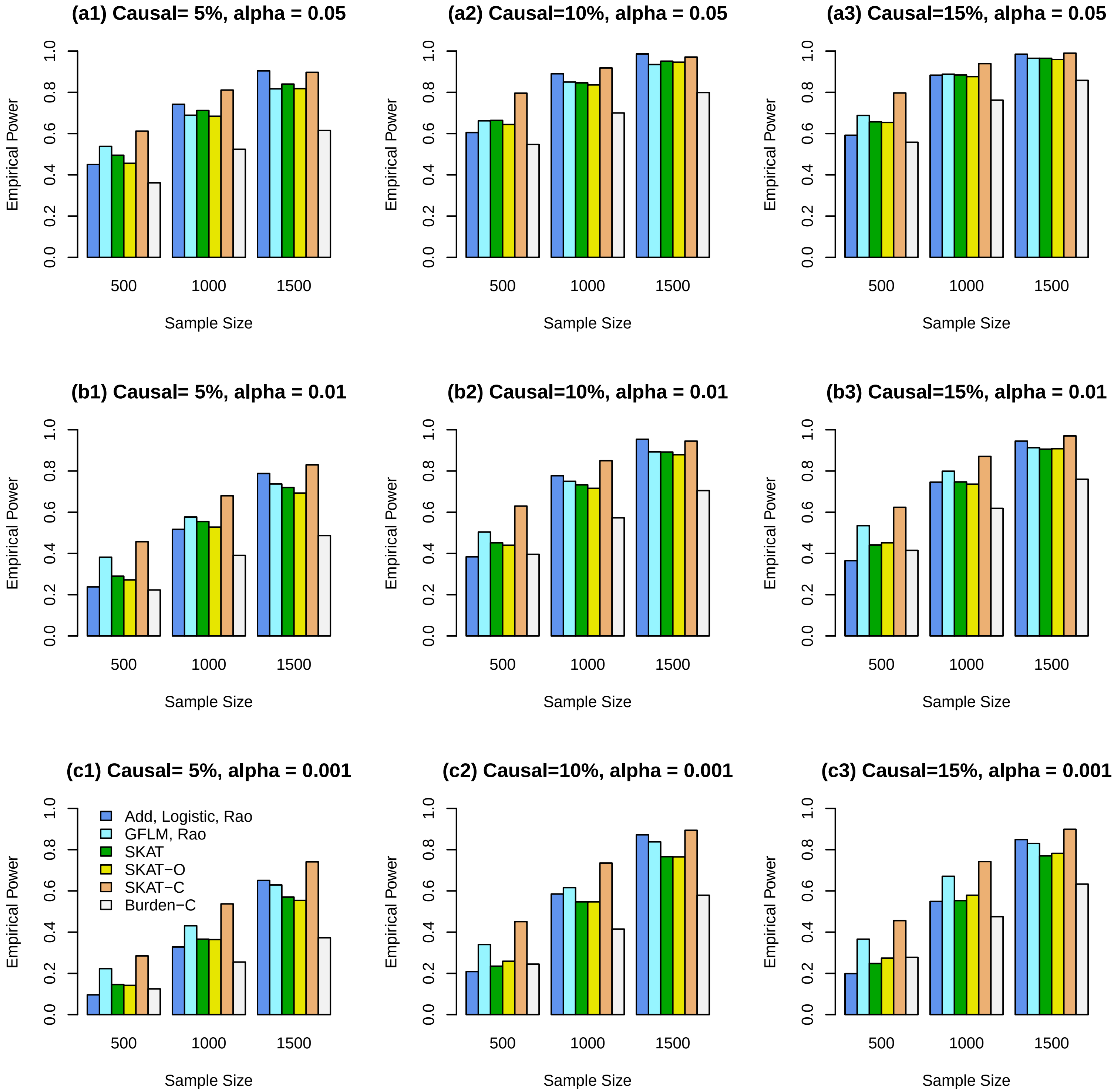

Figure S.13: **The Empirical Power of the Rao's Efficient Score Tests of Additive Logistic Regression Model (2) and the GFLM (8) Using B-spline Basis, SKAT, SKAT-O, SKAT-C, and Burden-C, When Some Causal Variants are Rare and Some are Common, the Region Size is 9 kb, the Constant $k = 1.5$ in Genetic Effect Size $|\beta_{ij}|$, and 20%/80% Causal Variants Had Negative/Positive Effects.** The order of B-spline basis was 4, and the number of B-spline basis functions was $K = K_\beta = 10$.

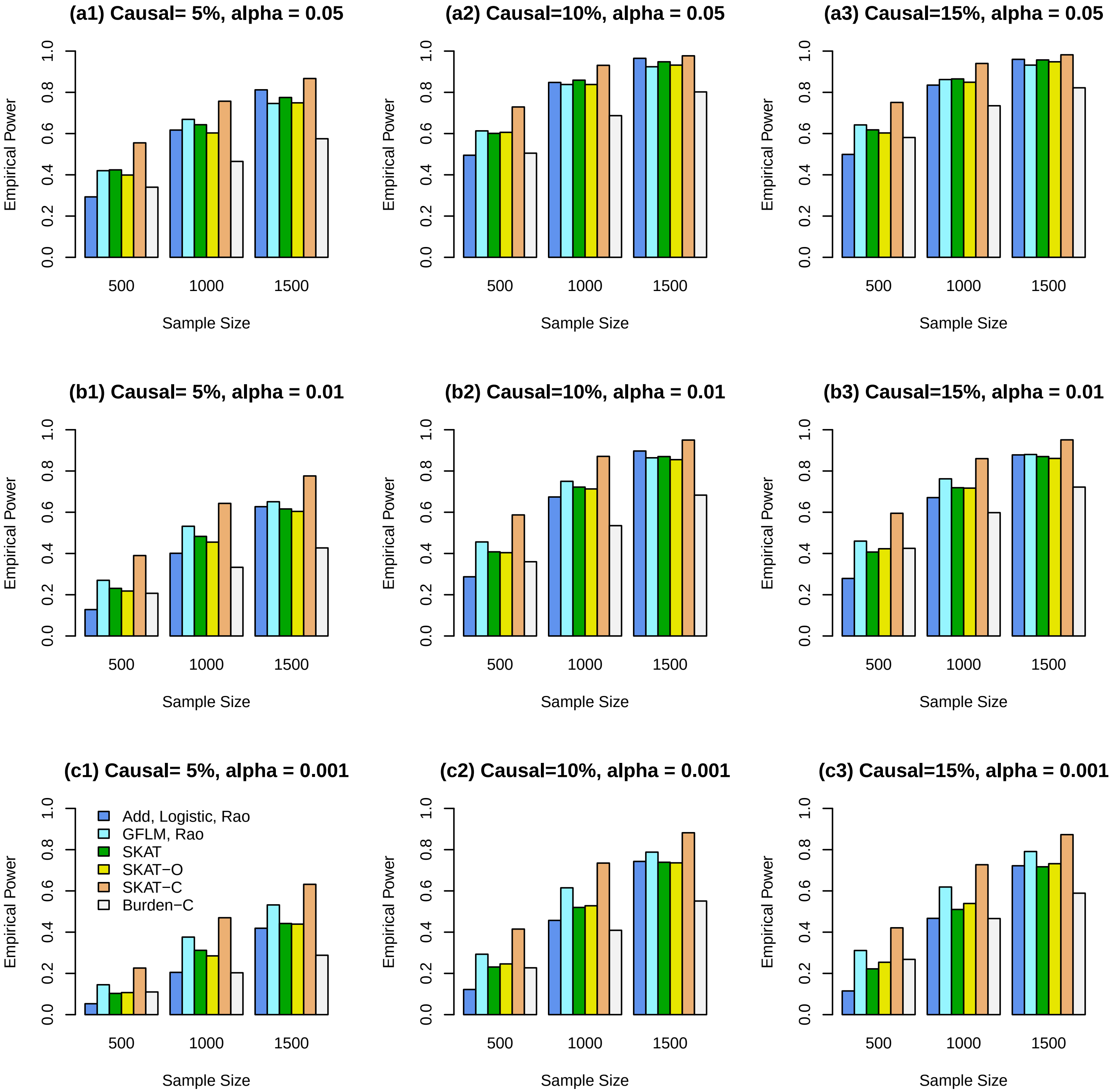

Figure S.14: **The Empirical Power of the Rao's Efficient Score Tests of Additive Logistic Regression Model (2) and the GFLM (8) Using B-spline Basis, SKAT, SKAT-O, SKAT-C, and Burden-C, When Some Causal Variants are Rare and Some are Common, the Region Size is 12 kb, the Constant $k = 1.75$ in Genetic Effect Size $|\beta_{ij}|$, and 20%/80% Causal Variants Had Negative/Positive Effects.** The order of B-spline basis was 4, and the number of B-spline basis functions was $K = K_\beta = 10$.

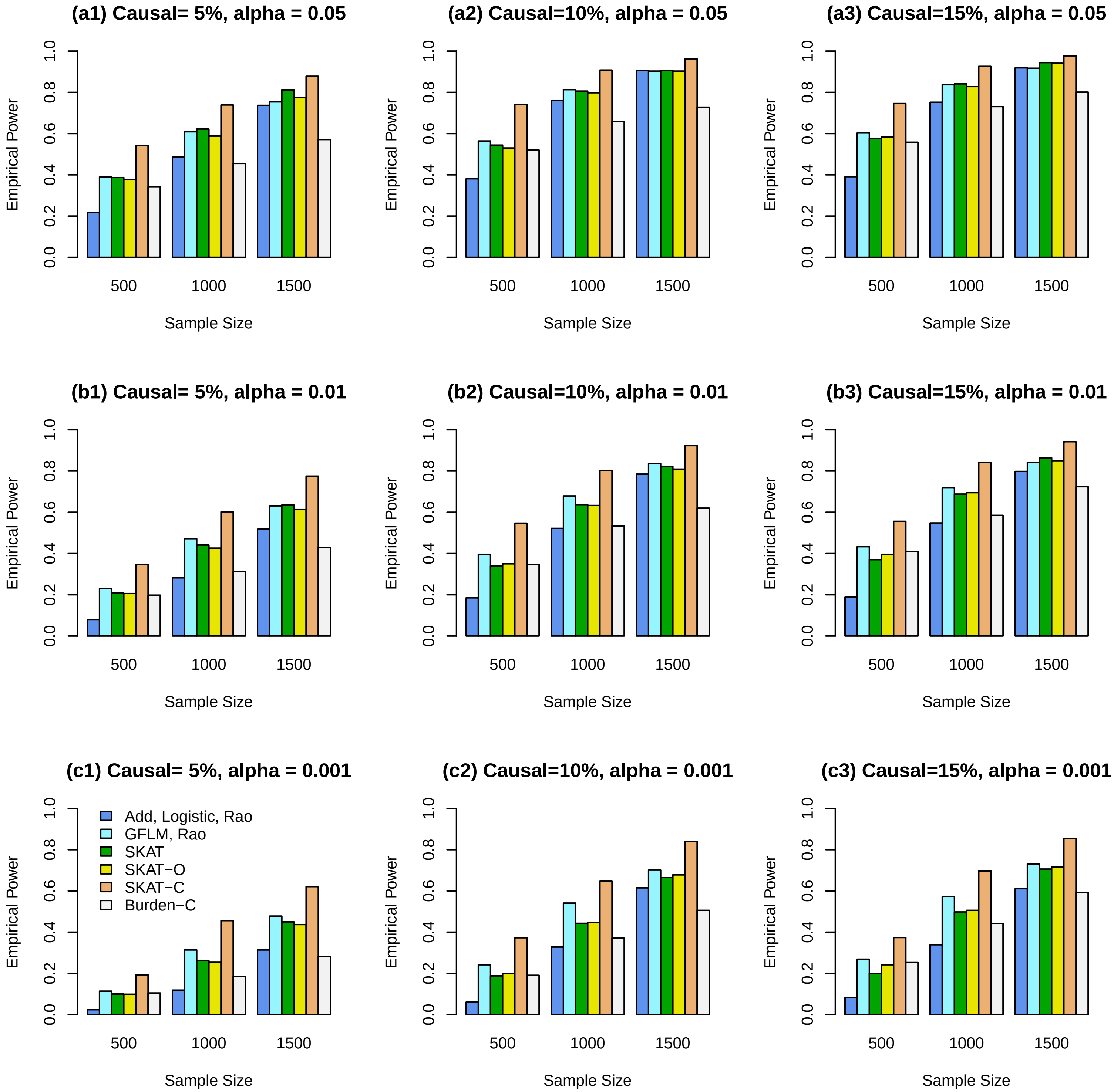

Figure S.15: **The Empirical Power of the Rao's Efficient Score Tests of Additive Logistic Regression Model (2) and the GFLM (8) Using B-spline Basis, SKAT, SKAT-O, SKAT-C, and Burden-C, When Some Causal Variants are Rare and Some are Common, the Region Size is 15 kb, the Constant $k = 2.0$ in Genetic Effect Size $|\beta_{ij}|$, and 20%/80% Causal Variants Had Negative/Positive Effects.** The order of B-spline basis was 4, and the number of B-spline basis functions was $K = K_\beta = 10$.