Supplementary Figures to A General Framework for Integrative Analysis of Incomplete Multi-omics Data

Dan-Yu Lin, Donglin Zeng, and David Couper

Department of Biostatistics, University of North Carolina, Chapel Hill, NC 27599-7420

Figure Legends

Figure S1. Quantile-quantile plots for testing $\beta = 0$ under the Joint Model when the detection limit is 1.

Figure S2. Simulation results under the Joint Model when the biomarker has the t-distribution with 5 degrees of freedom. The red, black, green, and blue curves pertain to the proposed method, complete-case analysis, imputation at limit, and imputation at mid-point, respectively.

Figure S3. Simulation results under the Joint Model when the biomarker is the logarithm of the standard exponential random variable. The red, black, green, and blue curves pertain to the proposed method, complete-case analysis, imputation at limit, and imputation at mid-point, respectively.

Figure S4. Simulation results under the Joint Model when the biomarker has the standard normal distribution and the inverse-normal transformation is performed on the observed values. The red, black, green, and blue curves pertain to the proposed method, complete-case analysis, imputation at limit, and imputation at mid-point, respectively.

Figure S5. Simulation results under the Joint Model with $\beta = 0$ when the biomarker has 30% missing values: (A) effect of the SNP genotype on the quantitative omics variable (i.e., α_G); (B) effect of the quantitative omics variable on the phenotype (i.e., γ); and (C) effect of the genotype on the phenotype (i.e., β_G). The bias and standard error of the

parameter estimator and the power or type I error of the association test are plotted against the detection limit of the quantitative omics variable. The red, black, green, and blue curves pertain to the proposed method, complete-case analysis, imputation at limit, and imputation at mid-point, respectively. The silver curve pertains to the mean of the standard error estimator for the proposed method.

Figure S6. Simulation results under the Joint Model with $\beta = 0.2$ when the biomarker has 30% missing values: (A) effect of the SNP genotype on the quantitative omics variable (i.e., α_G); (B) effect of the quantitative omics variable on the phenotype (i.e., γ); and (C) effect of the genotype on the phenotype (i.e., β). The bias and standard error of the parameter estimator and the power of the association test are plotted against the lower detection limit of the quantitative omics variable. The red, black, green, and blue curves pertain to the proposed method, complete-case analysis, imputation at limit, and imputation at mid-point, respectively. The silver curve pertains to the mean of the standard error estimator for the proposed method.

Figure S7. Simulation results under the Joint Model with $\beta = 0.2$ when there are two SNPs in the model: (A) effect of the SNP with MAF of 0.4 on the quantitative omics variable (i.e., α_G); (B) effect of the quantitative omics variable on the phenotype (i.e., γ); and (C) effect of the SNP with MAF of 0.4 on the phenotype (i.e., β). The bias and standard error of the parameter estimator and the power of the association test are plotted against the lower detection limit of the quantitative omics variable. The red, black, green, and blue curves pertain to the proposed method, complete-case analysis, imputation at limit, and imputation at mid-point, respectively. The silver curve pertains to the mean of the standard error estimator for the proposed method.

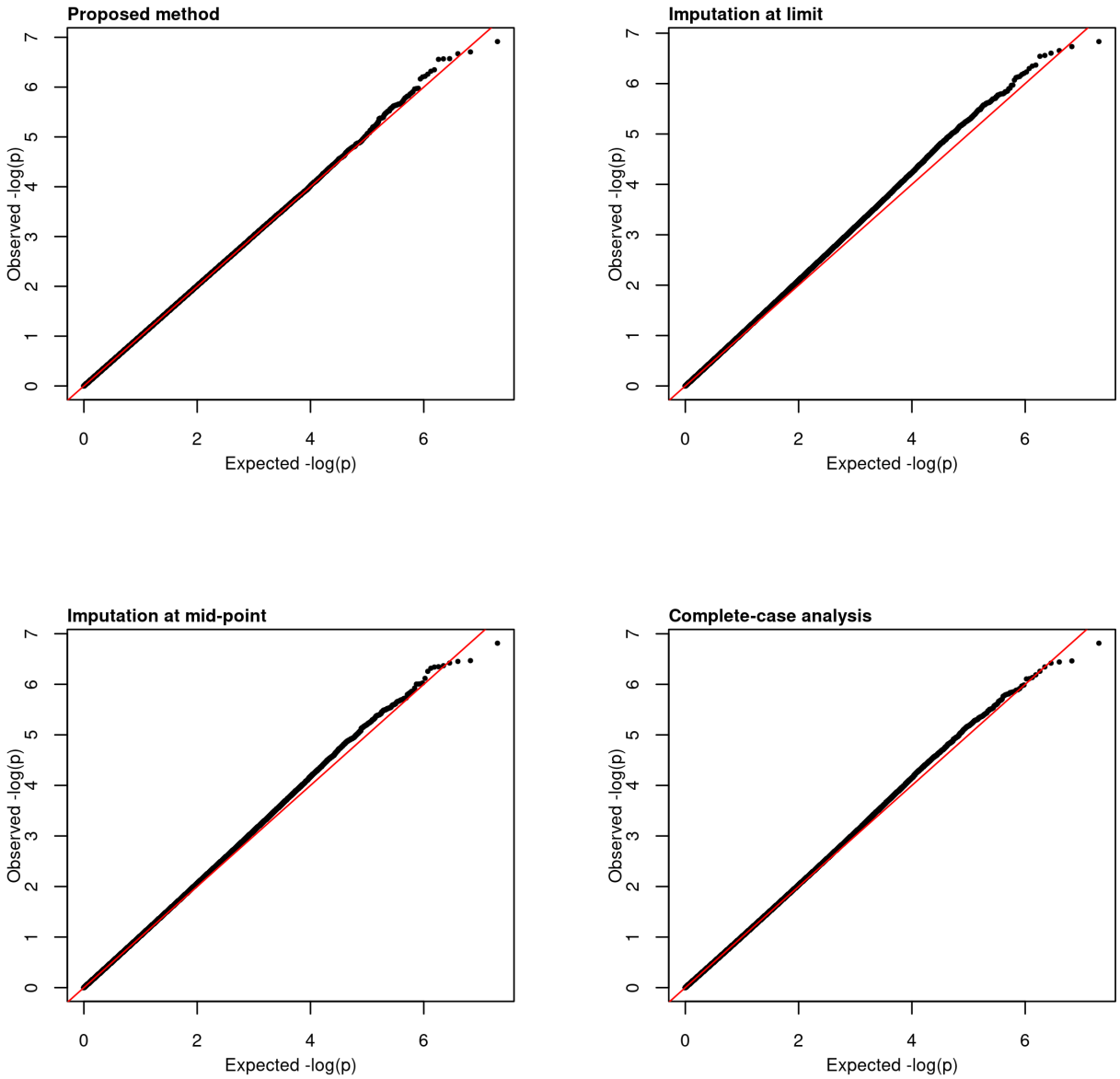
Proposed method
Imputation at limit
Imputation at mid-point
Complete-case analysis
Observed -log(p)
Expected -log(p)
0
1
2
3
4
5
6
7

Figure S1

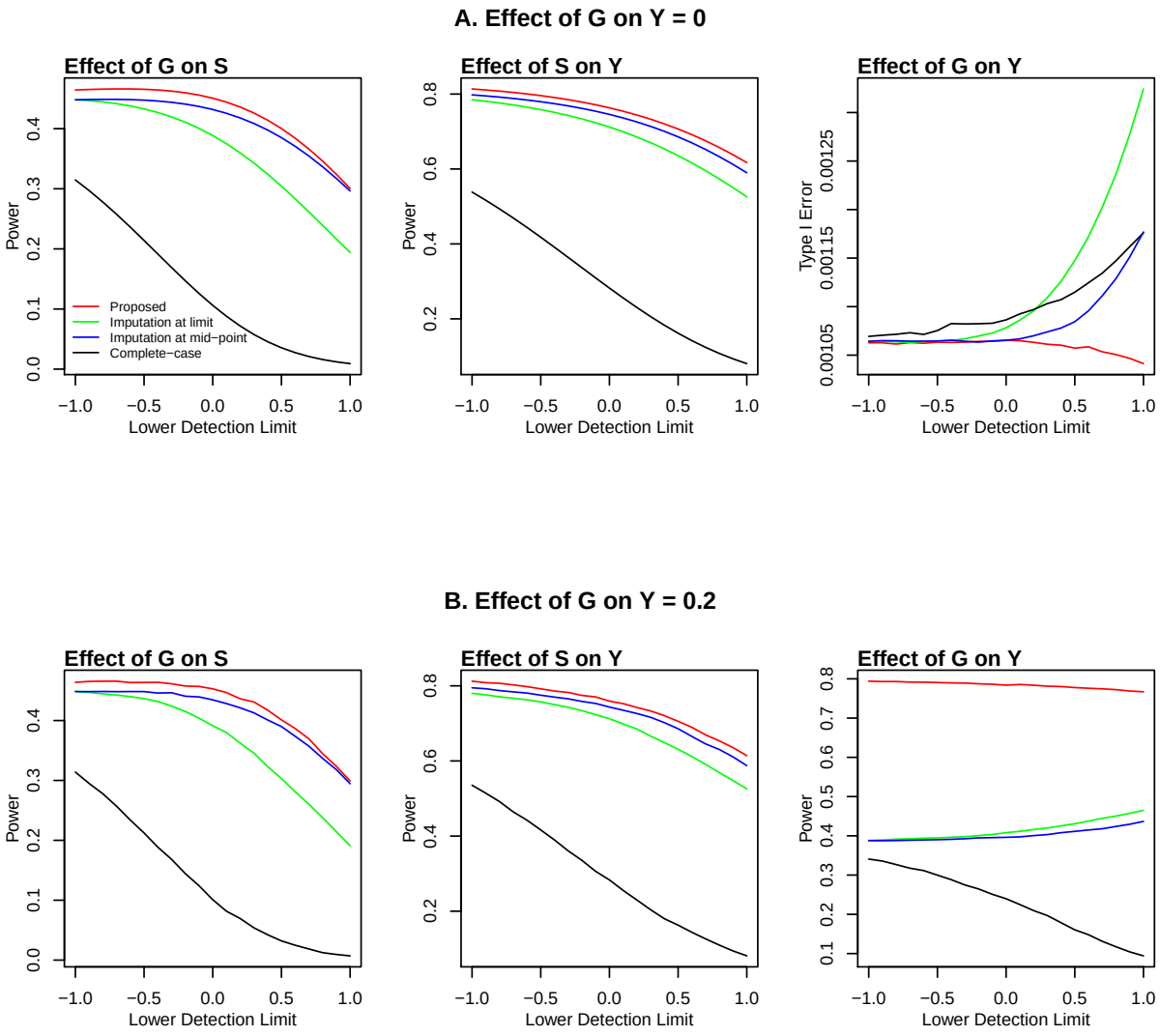
A. Effect of G on Y = 0
Effect of G on S
Power
Lower Detection Limit
Proposed
Imputation at limit
Imputation at mid−point
Complete−case
Effect of S on Y
Power
Lower Detection Limit
Effect of G on Y
Type I Error
Lower Detection Limit
B. Effect of G on Y = 0.2
Effect of G on S
Power
Lower Detection Limit
Effect of S on Y
Power
Lower Detection Limit
Effect of G on Y
Power
Lower Detection Limit

Figure S2

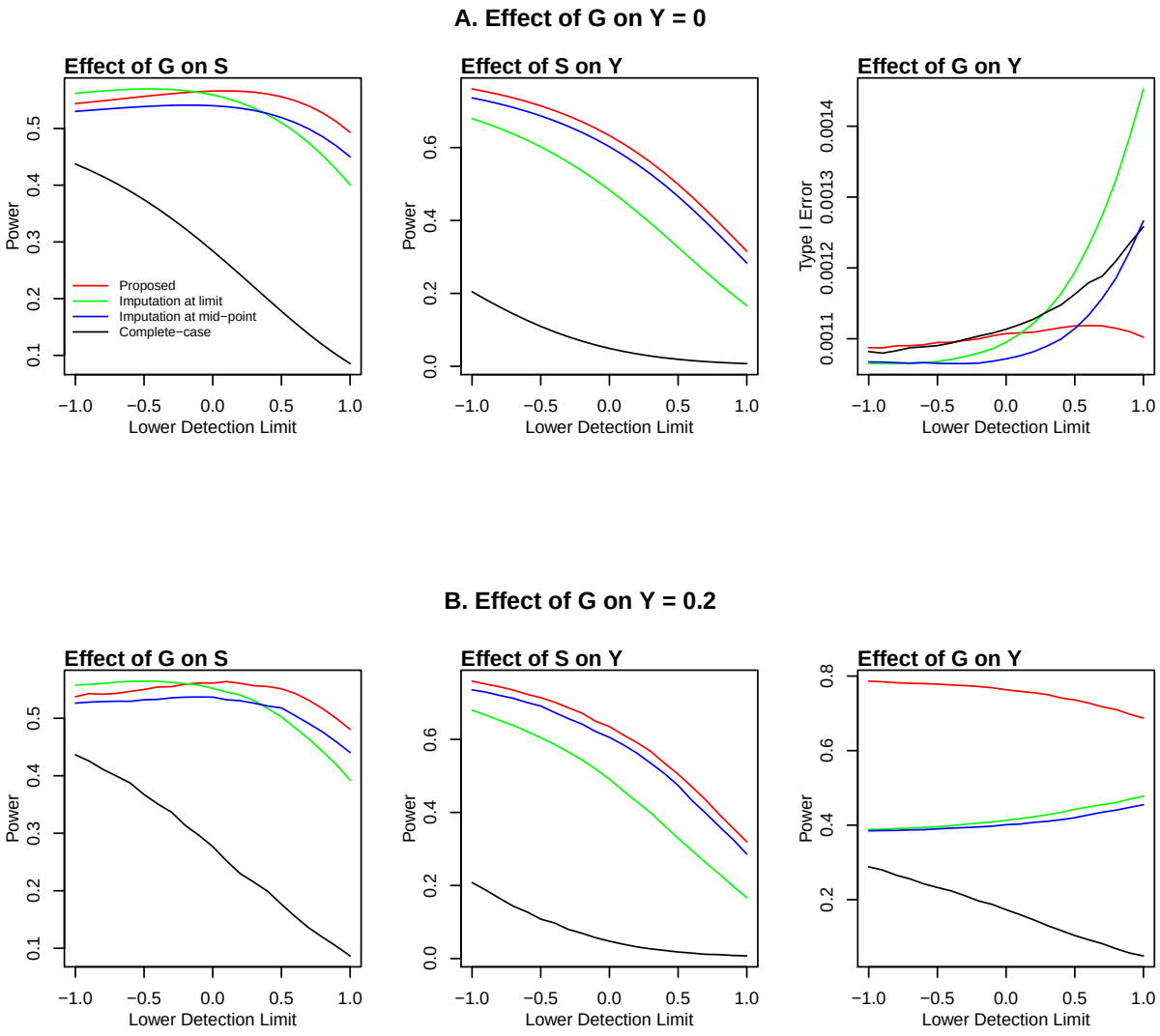
A. Effect of G on Y = 0
Effect of G on S
Power
Lower Detection Limit
Proposed
Imputation at limit
Imputation at mid−point
Complete−case
Effect of S on Y
Power
Lower Detection Limit
Effect of G on Y
Type I Error
Lower Detection Limit
B. Effect of G on Y = 0.2
Effect of G on S
Power
Lower Detection Limit
Effect of S on Y
Power
Lower Detection Limit
Effect of G on Y
Power
Lower Detection Limit

Figure S3

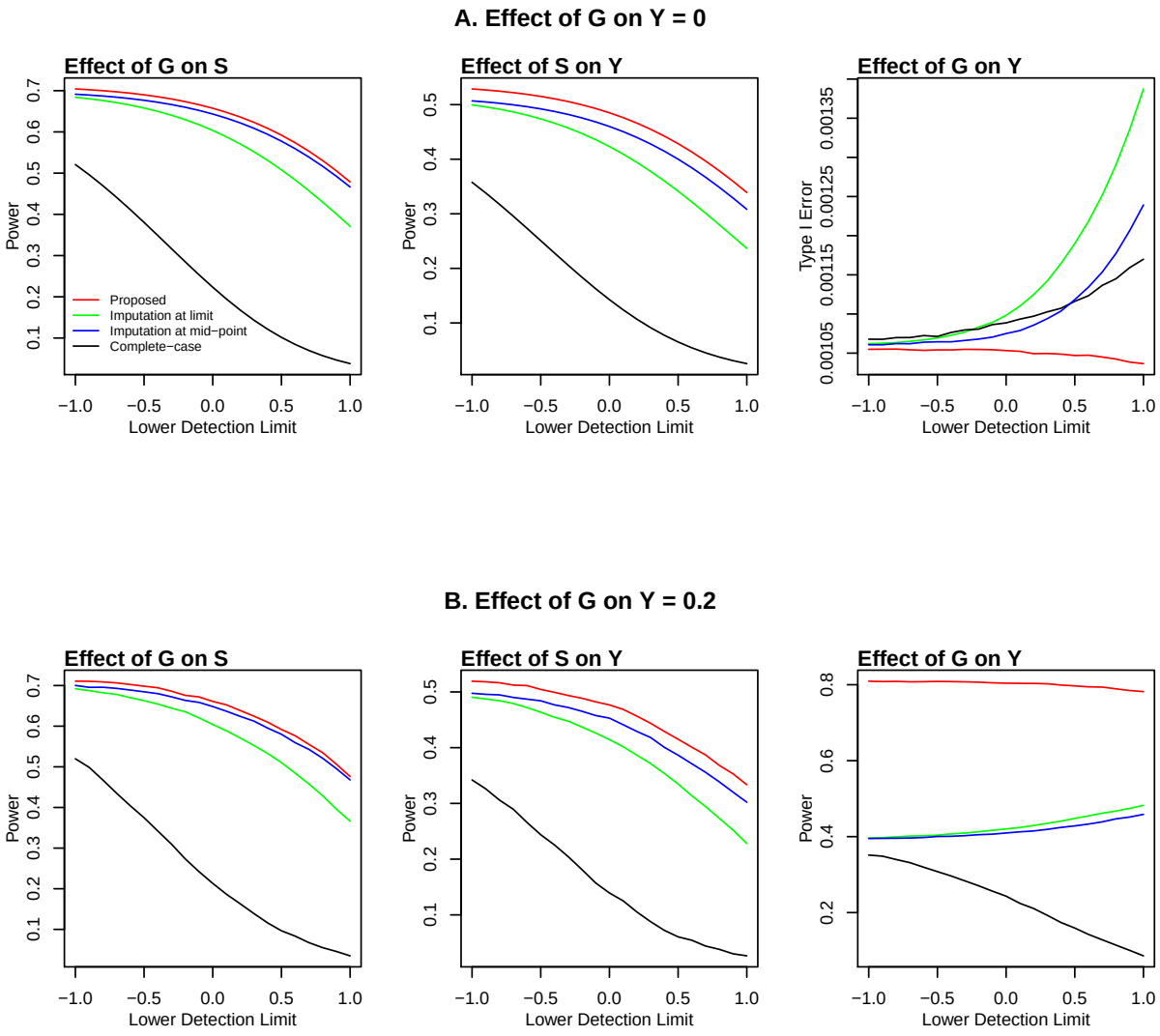
A. Effect of G on Y = 0
Effect of G on S
Power
Lower Detection Limit
Proposed
Imputation at limit
Imputation at mid−point
Complete−case
Effect of S on Y
Power
Lower Detection Limit
Effect of G on Y
Type I Error
Lower Detection Limit
B. Effect of G on Y = 0.2
Effect of G on S
Power
Lower Detection Limit
Effect of S on Y
Power
Lower Detection Limit
Effect of G on Y
Power
Lower Detection Limit

Figure S4

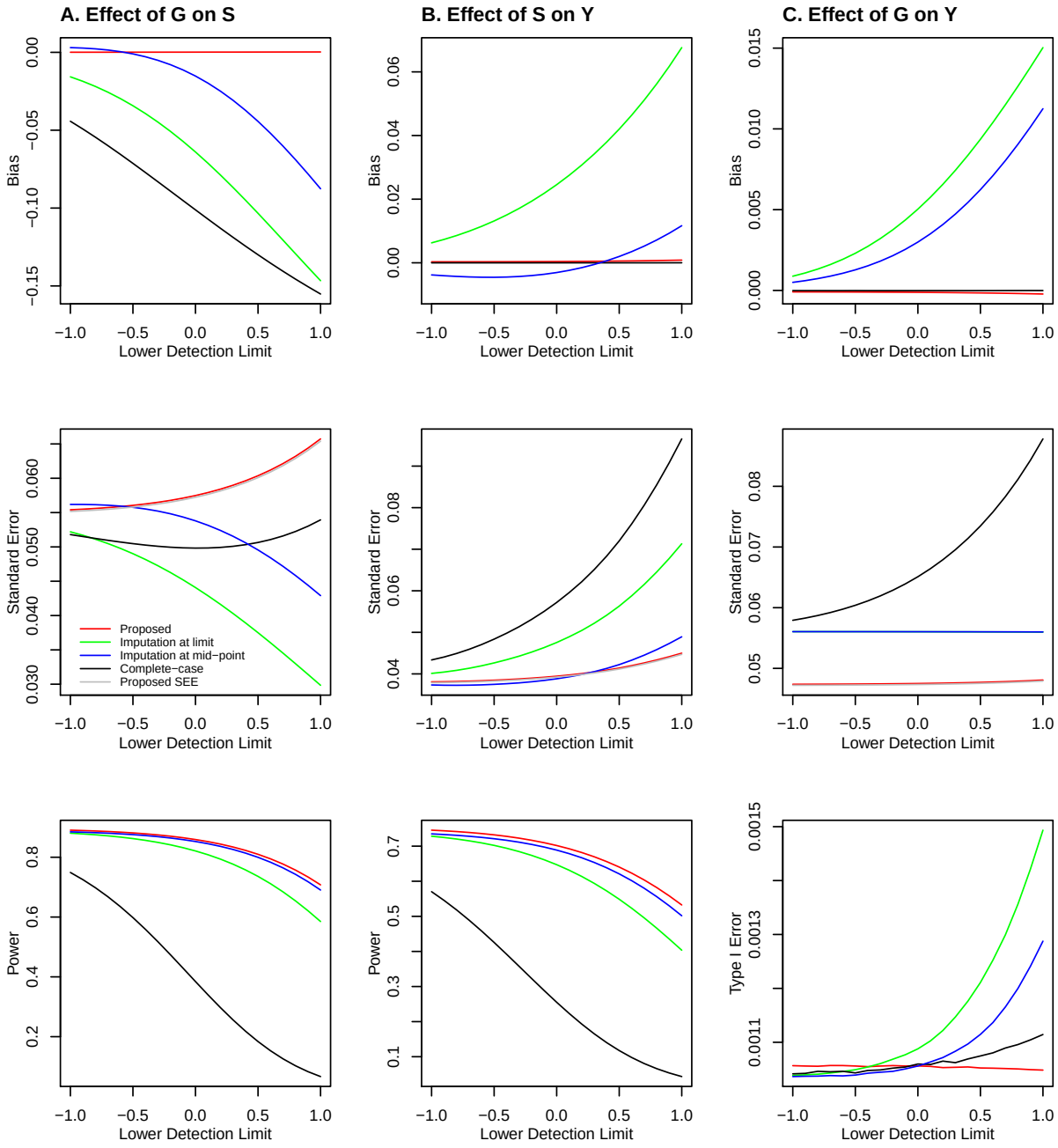
A. Effect of G on S
B. Effect of S on Y
C. Effect of G on Y
Bias
Standard Error
Power
Type I Error
Lower Detection Limit
Proposed
Imputation at limit
Imputation at mid−point
Complete−case
Proposed SEE

Figure S5

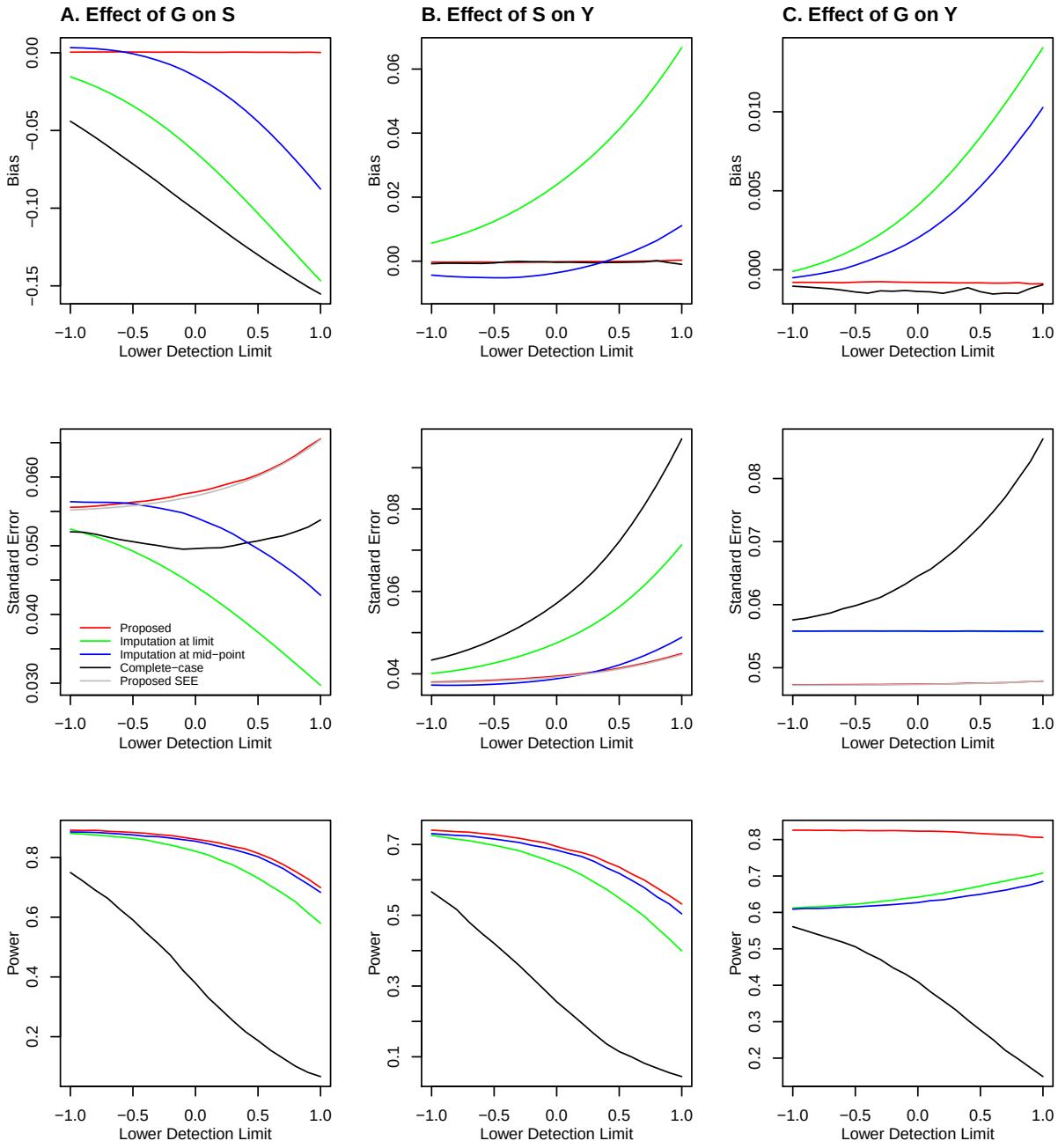
A. Effect of G on S
B. Effect of S on Y
C. Effect of G on Y
Bias
Standard Error
Power
Lower Detection Limit
Proposed
Imputation at limit
Imputation at mid−point
Complete−case
Proposed SEE

Figure S6

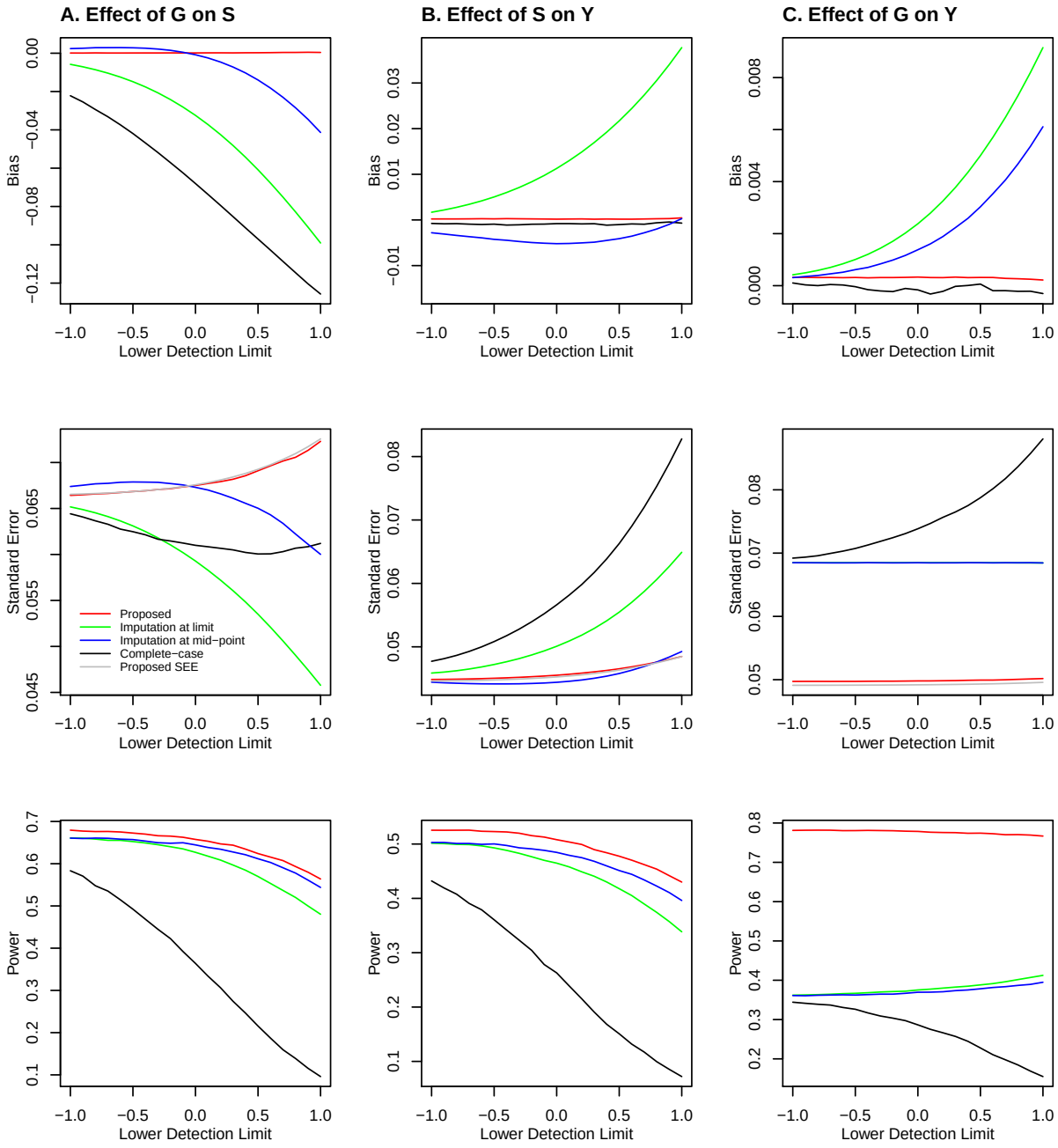
A. Effect of G on S
B. Effect of S on Y
C. Effect of G on Y
Bias
Standard Error
Power
Lower Detection Limit
Proposed
Imputation at limit
Imputation at mid−point
Complete−case
Proposed SEE

Figure S7