Supplement: Appendix S2

To give a comprehensive assessment of the R^2s, this supplement discusses and plots a complete set of simulations, from which a subset was presented in the main text. The different models (LMM, PGLS, GLMM, and PLOG) are discussed in turn.

LMM

Simulations were performed with only fixed effects (Fig. S1a-c), only random effects (Fig. S1d-f), and both fixed and random effects (Fig. S1g-i). For all 20 simulations per panel, the simulation model parameters were the same, so variation was caused by different random outcomes of the simulations. All methods performed fairly well in distinguishing the source of variation in the models: for simulations with only fixed effects, the partial R^2s for the fixed effect x were substantially higher than the partial R^2s for the random effects (Fig. S1a vs. S1b), and the converse held for simulations with only random effects (Fig. S1d vs. S1e). All five R^2s gave similar values when all of the variation was in the fixed effects (Fig. S1a-c). However, when there was variation caused by random effects, R^2_{lr} was lower than R^2_{ce} which was lower than R^2_{ls} (Fig. S1e, f, h, i). Because R^2_{lr} is a function of likelihoods, the total R^2_{lr} is necessarily a monotonically increasing function of LLR. Total R^2_{ls} and R^2_{ce} are also strongly related to LLR, indicating that they closely correspond to the results of a χ^2 significance test for the inclusion of both fixed and random effects in the LMM fitted to the data.

R^2_{ls} gave similar results to R^2_{glmm} and R^2_{ols}, except the partial R^2s in simulations with both fixed and random effects (Fig. S1g,h) in which R^2_{glmm} gave lower values. These lower values of R^2_{glmm} arise because I have compared the marginal $R^2_{glmm(m)}$ and $R^2_{glmm(v)}$ derived from Nakagawa and Schielzeth (2013) with the partial R^2s for the other formulas (see main text). For total R^2s, R^2_{ls} and R^2_{glmm} were almost identical (Fig. S3).

Plotting mean R^2s versus sample size for simulations which contain both fixed and random effects shows that R^2_{ls}, R^2_{glmm}, and R^2_{ols} are independent of sample size, whereas both R^2_{lr} and R^2_{ce} decrease with smaller samples (Fig. S3).

PGLS

For simulations under the same parameter values, R^2_{ls}, R^2_{lr} and R^2_{ce} showed similar behaviors as found for the LMM simulations (compare Figs. S4, S5 with Figs. S1, S2): all three

R^2s could distinguish between variation caused by the predictor variable x and the phylogeny (Fig. S4a vs. S4b, and S4d vs. S4e). The total R^2s were not very sensitive to sample size, and R^2_{lr} was slightly lower than R^2_{ls} and R^2_{ce} (Fig. S6). The partial R^2s changed with sample size, but in different directions. At least in part, this is caused by the effect of sample size on estimating phylogenetic signal λ; for small sample sizes, the maximum likelihood value of λ is more likely to be zero, leading to lower partial R^2s for the phylogeny than for x. Finally, the partial R^2 with respect to the predictor variable was lowest for R^2_{ce}, but the partial R^2 with respect to the phylogeny was lowest for R^2_{lr}, although it is not clear (to me) why.

GLMM

The results for the GLMM simulations were similar to those for the LMMs, although the differences among R^2s were greater (Figs. S7, S9). In particular, the tight relationships among R^2_{ls} and R^2_{glmm} found in the LMM (Figs. S1, S3) were not observed in the GLMM, although these three measures were closer to each other than R^2_{lr} or R^2_{ce}, which often showed similar values (Fig. S8). I also used an R^2_{ols} for GLMMs derived from equation 4 with $\mathbf{V}(\theta) = \mathbf{I}$. Although the partial R^2s could in general identify the source of variation in the simulations (Fig. S7a,b,d,e), $R^2_{ls,}$ R^2_{glmm} and especially R^2_{ols} could show relatively high values of the partial R^2 for the random effect u when in fact u was zero in the simulations (Fig. S7b). All five R^2s were sensitive to sample sizes, but in different directions: R^2_{ls} and R^2_{ols} tended to increase with decreasing sample sizes, whereas the other R^2s decreased (Fig. S9).

The behaviors of the R^2s at small sample sizes were in general not very good and were vulnerable to poor estimates of the random effect $\hat{\sigma}_l^2$ (eqn 8). When the estimate of the standard deviation of the random effect in the full model was large, R^2_{ls} could take values of either 1 or less than –1. Because R^2_{glmm} depends directly on $\hat{\sigma}_l^2$, it is also sensitive to the estimate of the random effect. The partial R^2_{ce} for the random effect was particularly prone to negative values for small sample sizes (Fig. S9b). These issues involving the estimate of the random effect are less severe for R^2_{lr} whose calculation does not use $\hat{\sigma}_l^2$. Thus, R^2_{lr} showed relatively low variation at small sample sizes.

PLOG

For the binary phylogenetic model, R^2_{ls} tended to be higher than R^2_{lr} and R^2_{ce} (Fig. S10, S11), and all three appeared to be independent of sample size (Fig. S12). Although the standard deviation of R^2_{ls} among simulations was larger than R^2_{lr} and R^2_{ce}, this was mainly due to the mean values of R^2_{ls} being larger.

R^2_{ls} showed a seemingly odd pattern of not tending to zero as the LLR approached zero in simulations with phylogenetic signal $\lambda > 0$ (Fig. S10e,f). This is an artifact of different methods used to estimate the phylogenetic models: binaryPGLMM was used for R^2_{ls} and R^2_{ce}, and phyloglm was used for R^2_{lr} and the likelihood ratio LLR. In the simulations, binaryPGLMM would often estimate a positive phylogenetic signal when phyloglm would not, leading to $R^2_{ls} > 0$ even when LLR = 0. This underscores the dependence of R^2s on the specific model and estimation method used to fit the data.

FIGURES

Figure S1: Simulation results for a Linear Mixed Model (LMM) giving R^2_{ls} (eqn 2), R^2_{lr} (eqn 6), R^2_{ce} (eqn 7), R^2_{glmm} (eqn 8), and R^2_{ols} versus the log likelihood ratio (LLR) between full and reduced models. The simulation model (eqn 9) contained both a fixed effect β for a continuous variable x and a random effect b for a categorical variable u. For (a), (b) and (c), data were simulated without the random effect ($\beta = 1$, $\sigma = 0$), and for (d), (e) and (f), data were simulated without the fixed effect ($\beta = 0$, $\sigma = 1.5$). Simulations for (g), (h), and (i) contained both fixed and random effects. Columns give different partial R^2s for each method. Specifically, (a), (d), and (g) give the partial R^2s in which the reduced model removes the fixed effect for x: therefore, these give partial R^2s for the fixed effect. Panels (b), (e), and (h) give the partial R^2s in which the reduced model removes the random effect for u: therefore, these give partial R^2s for the random effect. In panels (c), (f) and (i), the reduced model removes both fixed and random effects, giving the total R^2s. Each data set consisted of 100 simulated points, x was simulated as a normal (0, 1) random variable, and u had 10 levels with b is simulated as a normal (0, σ). All analyses were performed with the function lmer().

Figure S2: Simulation results for a Linear Mixed Model (LMM) giving associations between R^2_{ls}, R^2_{lr}, R^2_{ce}, R^2_{glmm}, and R^2_{ols}. Data are the same as presented in figure S1.

Figure S3: Simulation results for a Linear Mixed Model (LMM) showing means and standard deviations of R^2_{ls} (eqn 2), R^2_{lr} (eqn 6), R^2_{ce} (eqn 7), R^2_{glmm} (eqn 8), and R^2_{ols} versus sample size. The simulation model (eqn 9) contained both a fixed effect β for a continuous variable *x* and a random effect *b* for a categorical variable *u*. For each level of *u*, from 4 to 16 replicates were simulated. (a), (b), and (c) give means of each R^2 were calculated for 500 simulations at each sample, and (d), (e), and (f) give standard deviations. Columns give different partial R^2s, with (a) and (d) giving the partial R^2s for the fixed effect, (b) and (e) giving the partial R^2s for the random effect, and (c) and (f) giving the total R^2s. In the simulations, *x* is simulated as a normal (0, 1) random variable with $\beta = 1$; *u* has 10 levels and *b* is simulated as a normal ($0, \sigma = 1.5$) random variable; and residuals *e* are independent (0, 1) random variables. All analyses were performed with the function lmer().

Figure S4: Simulation results for the phylogenetic model with a continuous predictor variable *x* giving R^2_{ls} (eqn 2), R^2_{lr} (eqn 6), and R^2_{ce} (eqn 7) versus the log likelihood ratio (LLR) between full and reduced models. For each simulation, a phylogenetic tree was first simulated, and the values of *x* were simulated up the phylogeny assuming Brownian Motion evolution. Data were simulated using equation 9 with $b = 0$, and residuals e_i were simulated from a multivariate normal distribution with mean zero and covariance matrix $\mathbf{\Sigma}(\lambda) = (1 - \lambda)\mathbf{I} + \lambda\mathbf{\Sigma}_{\mathrm{BM}}$. For (a), (b) and (c), data were simulated without phylogenetic signal ($\lambda = 0$, $\beta = 1.5$), and for (d), (e) and (f), data were simulated without the fixed effect ($\lambda = 0.7$, $\beta = 0$). Simulations for (g), (h), and (i) contained both fixed and phylogenetic effects ($\beta = 1.5$, $\lambda = 0.7$). (a), (d), and (g) give the partial R^2s for the fixed effect. Panels (b), (e), and (h) give the partial R^2s for the phylogenetic effect. In panels (c), (f) and (i), the reduced model removes both fixed and phylogenetic effects, giving the total R^2s. All analyses were performed with the function phylolm().

Figure S5: Simulation results for a PGLS model giving associations between R^2_{ls}, R^2_{lr}, and R^2_{ce}. Data are the same as presented in figure S3.

Figure S6: Simulation results for the phylogenetic model with a continuous response variable showing means and standard deviations of R^2_{ls} (eqn 2), R^2_{lr} (eqn 6), and R^2_{ce} (eqn 7) versus sample size. For each simulation, a phylogenetic tree was first simulated, and the values of the predictor variable x were simulated up the phylogeny assuming Brownian Motion evolution. In equation 9, residuals e_i were simulated from a multivariate normal distribution with mean zero and covariance matrix $\mathbf{\Sigma}(\lambda) = (1 - \lambda)\mathbf{I} + \lambda\mathbf{\Sigma}_{\mathrm{BM}}$, and the parameter values were $\lambda = 0.7$, $\beta = 1.5$, and $b = 0$. (a), (b), and (c) give means of each R^2 were calculated for 500 simulations at each sample, and (d), (e), and (f) give standard deviations. Columns give different partial R^2s, with (a) and (d) giving the partial R^2s for x, (b) and (e) giving the partial R^2s for phylogenetic signal λ, and (c) and (f) giving the total R^2s. All analyses were performed with the function phylolm().

Figure S7: Simulation results for a binary Generalized Linear Mixed Model (GLMM) giving R^2_{ls} (eqn 2), R^2_{lr} (eqn 6), R^2_{ce} (eqn 7), R^2_{glmm} (eqn 8), and R^2_{ols} versus the log likelihood ratio (LLR) between full and reduced models. The simulation model (eqn 9) contained both a fixed effect β for a continuous variable x and a random effect b for a categorical variable u. For (a), (b) and (c), data were simulated without the random effect ($\beta = 1.8$, $\sigma = 0$), and for (d), (e) and (f), data were simulated without the fixed effect ($\beta = 0$, $\sigma = 1.8$). Simulations for (g), (h), and (i) contained both fixed and random effects. (a), (d), and (g) give the partial R^2s for the fixed effect, and panels (b), (e), and (h) give the partial R^2s for the random effect. In panels (c), (f) and (i) give total R^2s. In the simulations, x is simulated as a normal (0, 1) random variable and u has 10 levels and b is simulated as a normal (0, σ). All analyses were performed with the function glmer().

Figure S8: Simulation results for a Generalized Linear Mixed Model (GLMM) giving associations between R^2_{ls}, R^2_{lr}, R^2_{ce}, R^2_{glmm}, and R^2_{ols}. Data are the same as presented in figure S7.

Figure S9: Simulation results for a binary Generalized Linear Mixed Model (GLMM) showing means and standard deviations of R^2_{ls} (eqn 2), R^2_{lr} (eqn 6), R^2_{ce} (eqn 7), R^2_{glmm} (eqn 8), and R^2_{ols} versus sample size. The simulation model (eqn 9) contained both a fixed effect β for a continuous variable x and a random effect b for a categorical variable u. For each level of u, from 4 to 16 replicates were simulated. (a), (b), and (c) give means of each R^2 were calculated for 1000 simulations at each sample, and (d), (e), and (f) give standard deviations. Columns give different partial R^2s, with (a) and (d) giving the partial R^2s for the fixed effect, (b) and (e) giving the partial R^2s for the random effect, and (c) and (f) giving the total R^2s. In the simulations, x is simulated as a normal (0, 1) random variable with $\beta = 1.8$ and u has 10 levels and b is simulated as a normal (0, $\sigma = 1.8$) random variable. All analyses were performed with the function glmer().

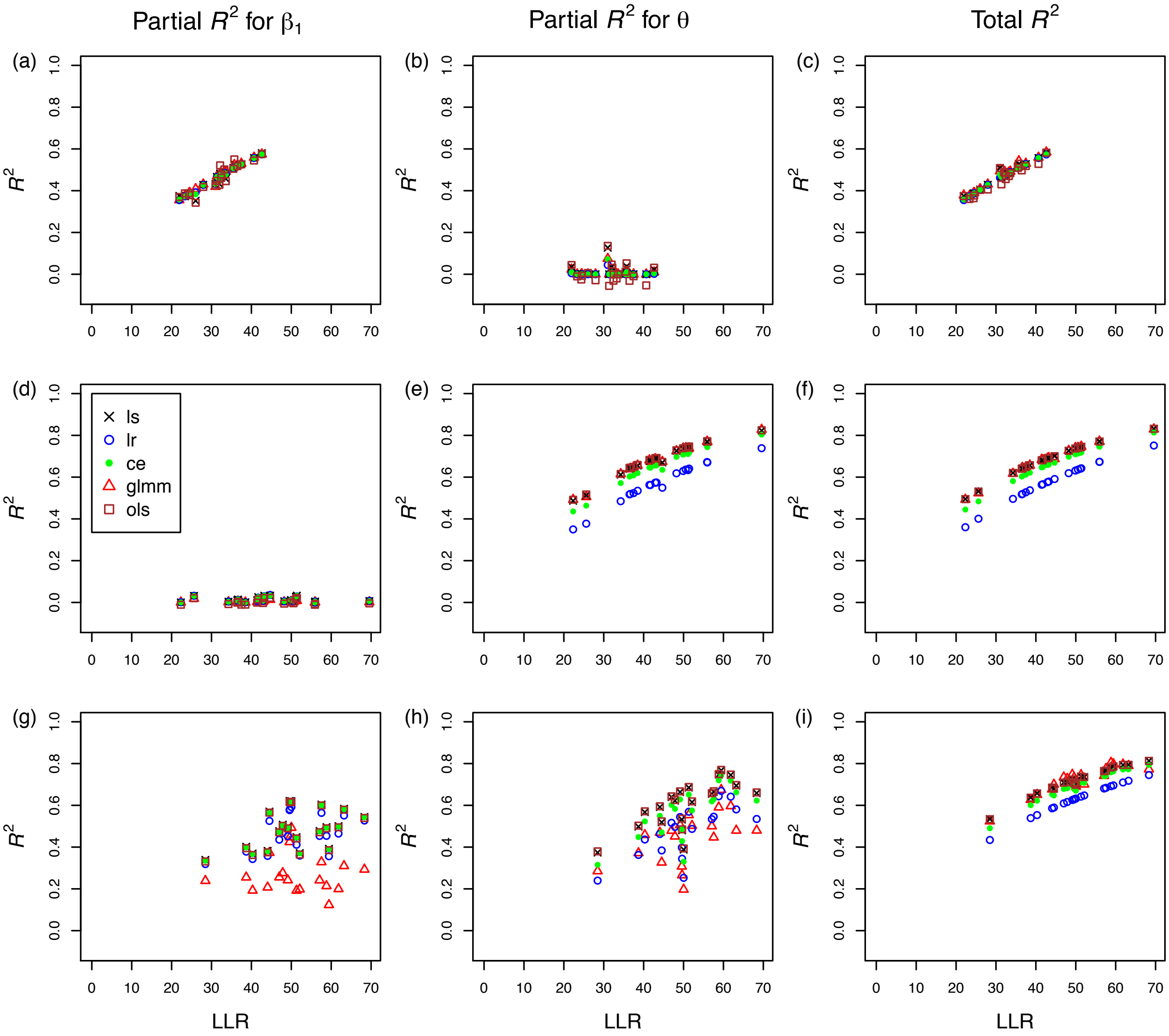
Figure S10: Simulation results for the phylogenetic model with a continuous predictor variable x giving R^2_{ls} (eqn 2), R^2_{lr} (eqn 6), and R^2_{ce} (eqn 7) versus the log likelihood ratio (LLR) between full and reduced models. For each simulation, a phylogenetic tree was first simulated, and the values of x were simulated up the phylogeny assuming Brownian Motion evolution. Data were simulated using equation 9 with $b = 0$, and residuals e_i were simulated from a multivariate normal distribution with mean zero and covariance matrix $\mathbf{\Sigma}(\lambda) = (1 - \lambda)\mathbf{I} + \lambda\mathbf{\Sigma}_{\mathrm{BM}}$. For (a), (b) and (c), data were simulated without phylogenetic signal ($\lambda = 0$, $\beta = 1.5$), and for (d), (e) and (f), data were simulated without the fixed effect ($\lambda = 2$, $\beta = 0$). Simulations for (g), (h), and (i) contained both fixed and phylogenetic effects ($\lambda = 2$, $\beta = 1.5$). (a), (d), and (g) give the partial R^2s for the fixed effect. Panels (b), (e), and (h) give the partial R^2s for the phylogenetic effect. In panels (c), (f) and (i), the reduced model removes both fixed and phylogenetic effects, giving the total R^2s. All analyses were performed with the function phylolm().

Figure S11: Simulation results for a phylogenetic logistic regression model giving associations between R^2_{ls}, R^2_{lr}, and R^2_{ce}. Data are the same as presented in figure S10.

Figure S12: Simulation results for the phylogenetic model with a binary response variable showing means and standard deviations of R^2_{ls} (eqn 2), R^2_{lr} (eqn 6), and R^2_{ce} (eqn 7) versus sample size. For each simulation, a phylogenetic tree was first simulated, and residuals e_i (eqn 9) were simulated from a multivariate normal distribution with mean zero and covariance matrix $\mathbf{\Sigma}(\lambda) = \lambda\mathbf{\Sigma}_{\mathrm{BM}}$. Values of the predictor variable x were assumed to be independently distributed by a (0,1) normal distribution, and the parameter values were $\lambda = 2$, $\beta = 1.5$, and $b = 0$. (a), (b), and (c) give means of each R^2 were calculated for 500 simulations at each sample, and (d), (e), and (f) give standard deviations. Columns give different partial R^2s, with (a) and (d) giving the partial R^2s for x, (b) and (e) giving the partial R^2s for phylogenetic signal λ, and (c) and (f) giving the total R^2s. Calculations of R^2_{lr} were performed with a modified version of the function phyloglm()

and the function glm(). Calculations of R^2_{ls} and R^2_{ce} were performed with the function binaryPGLMM().

Fig. S1



Partial R^2 for β_1
Partial R^2 for θ
Total R^2
(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
R^2
0.0
0.2
0.4
0.6
0.8
1.0
0
10
20
30
40
50
60
70
ls
lr
ce
glmm
ols
LLR

Fig. S2

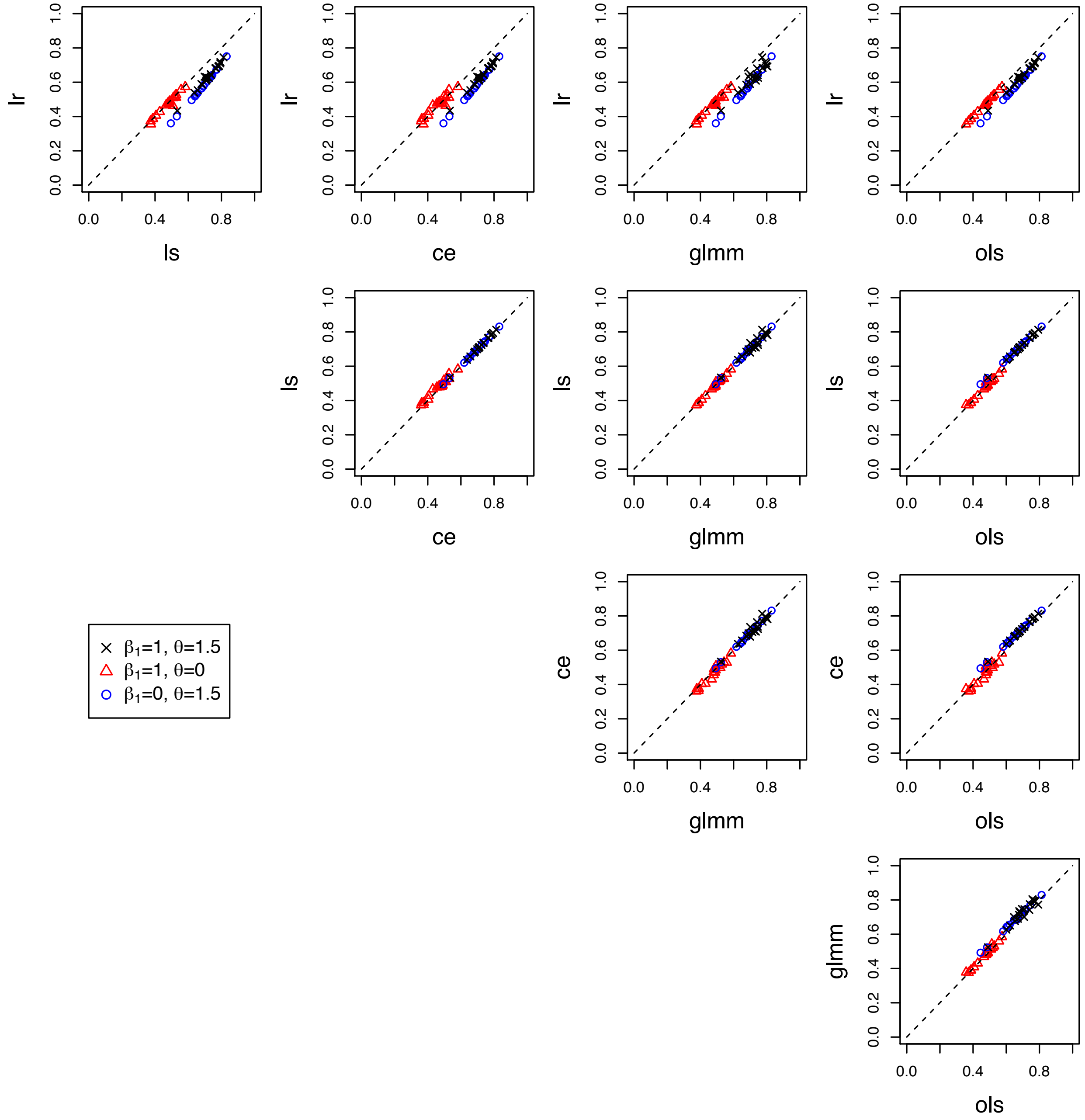

lr
ls
ce
glmm
ols
0.0
0.2
0.4
0.6
0.8
1.0
× β_1=1, θ=1.5
△ β_1=1, θ=0
○ β_1=0, θ=1.5

Fig. S3

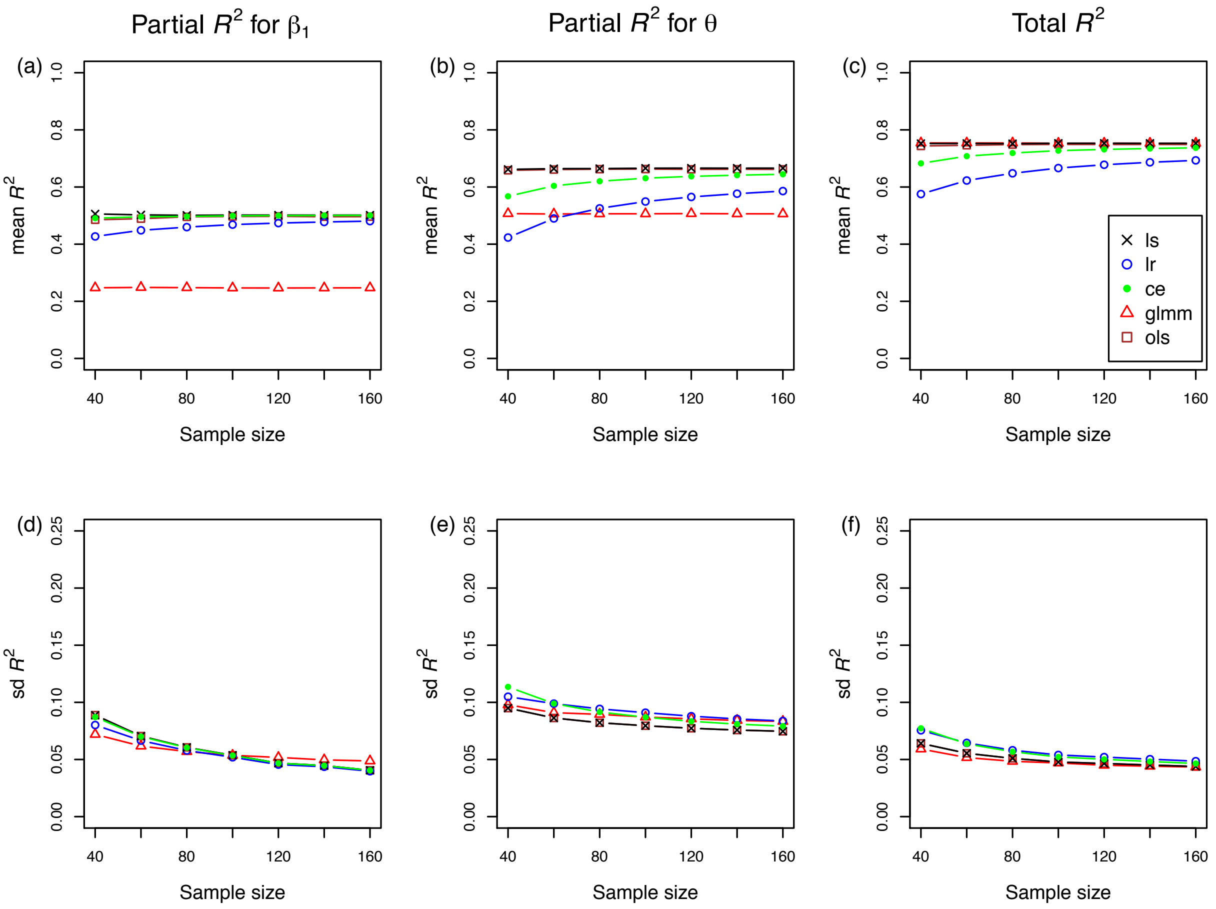
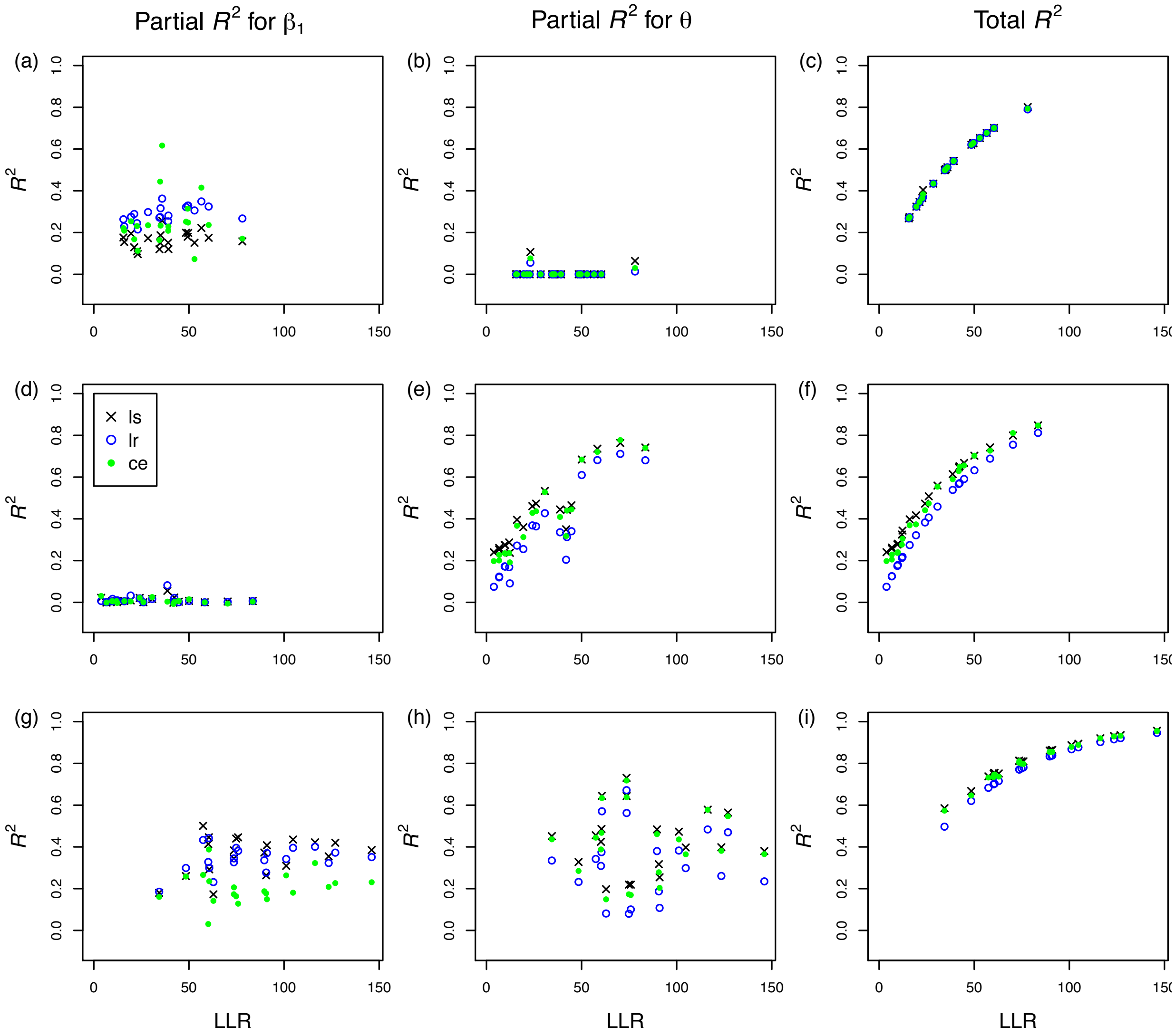

Partial R^2 for β_1
Partial R^2 for θ
Total R^2
(a)
(b)
(c)
(d)
(e)
(f)
mean R^2
sd R^2
0.0
0.2
0.4
0.6
0.8
1.0
0.00
0.05
0.10
0.15
0.20
0.25
40
80
120
160
Sample size
ls
lr
ce
glmm
ols

Fig. S4



Partial R^2 for β_1
Partial R^2 for θ
Total R^2
(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
R^2
0.0
0.2
0.4
0.6
0.8
1.0
0
50
100
150
ls
lr
ce
LLR

Fig. S5

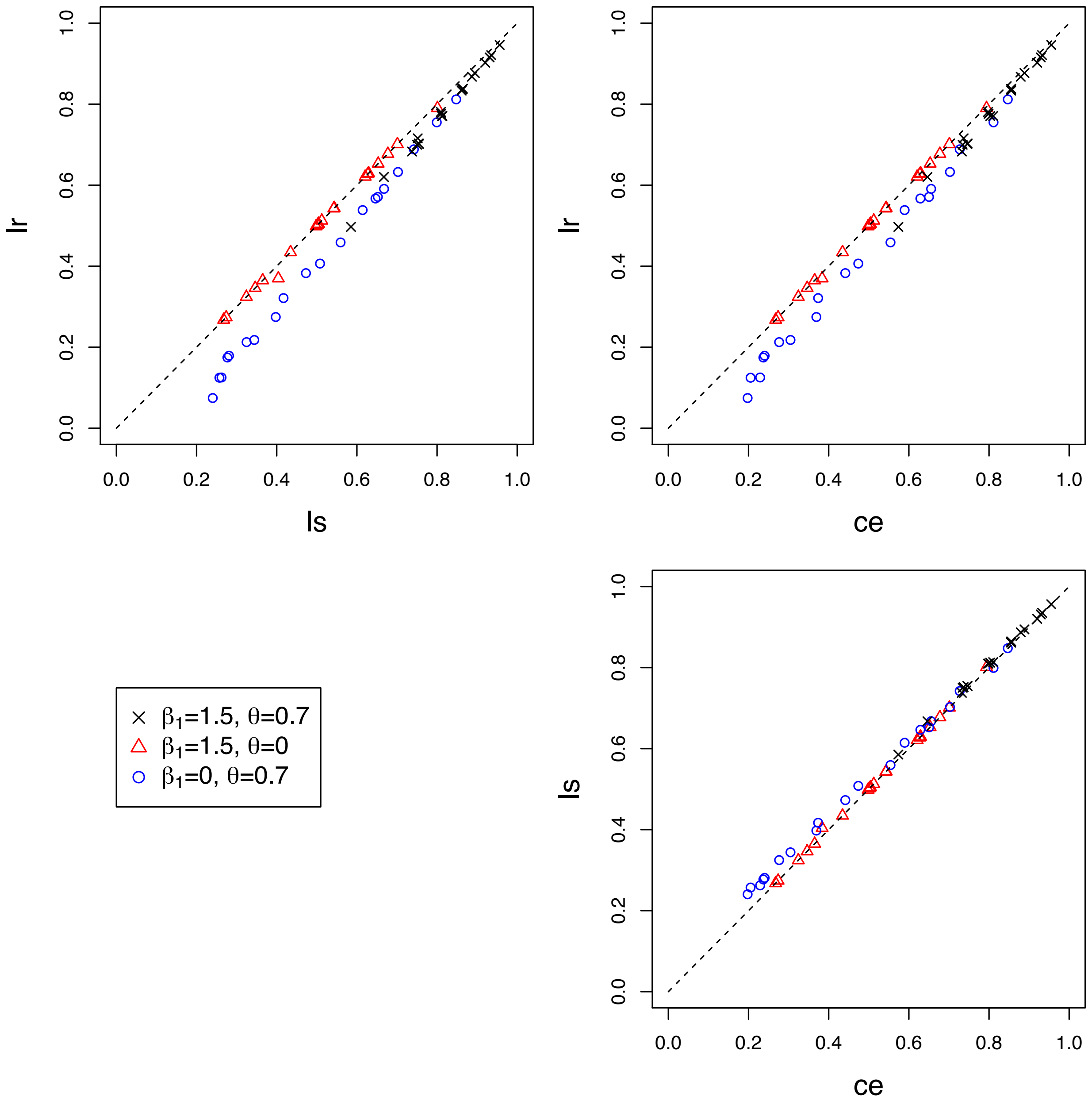


Fig. S6

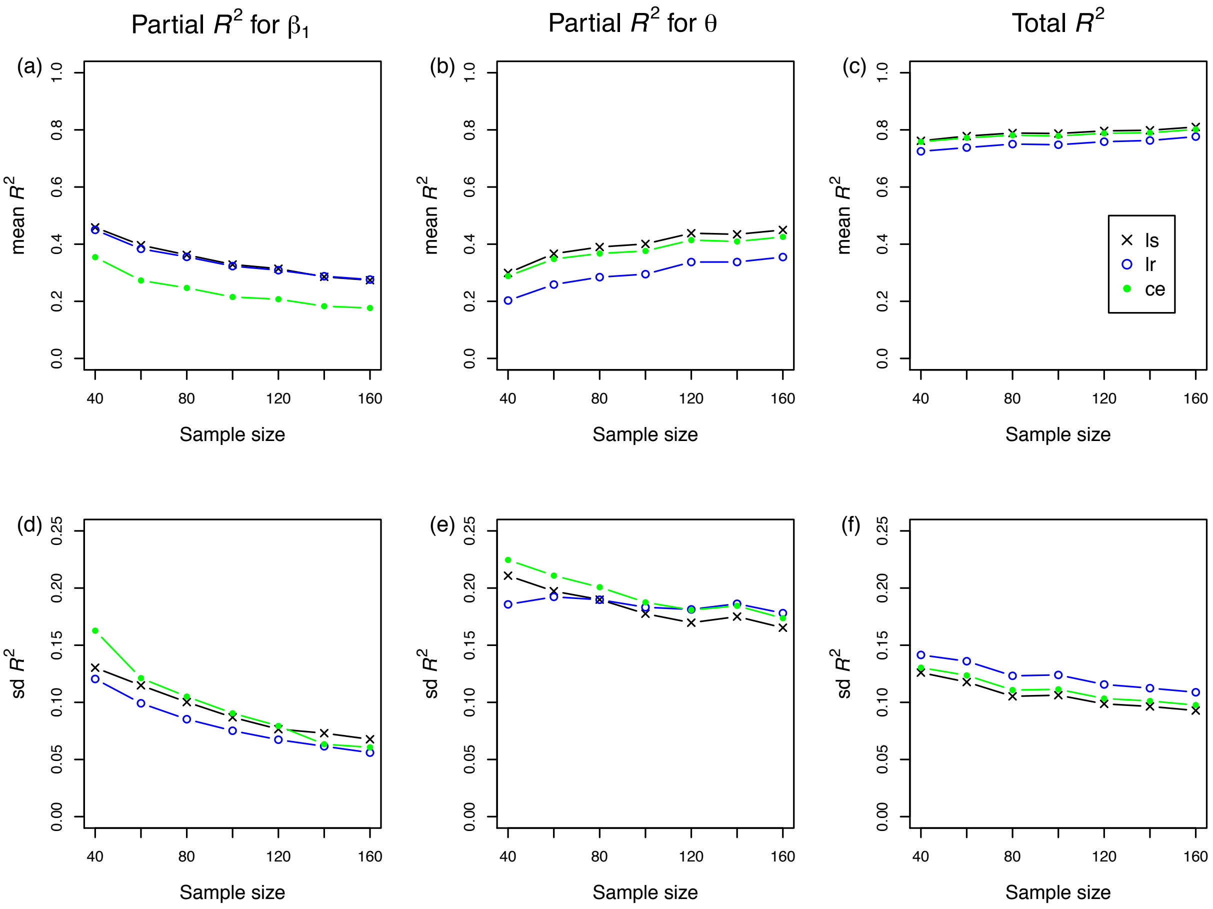
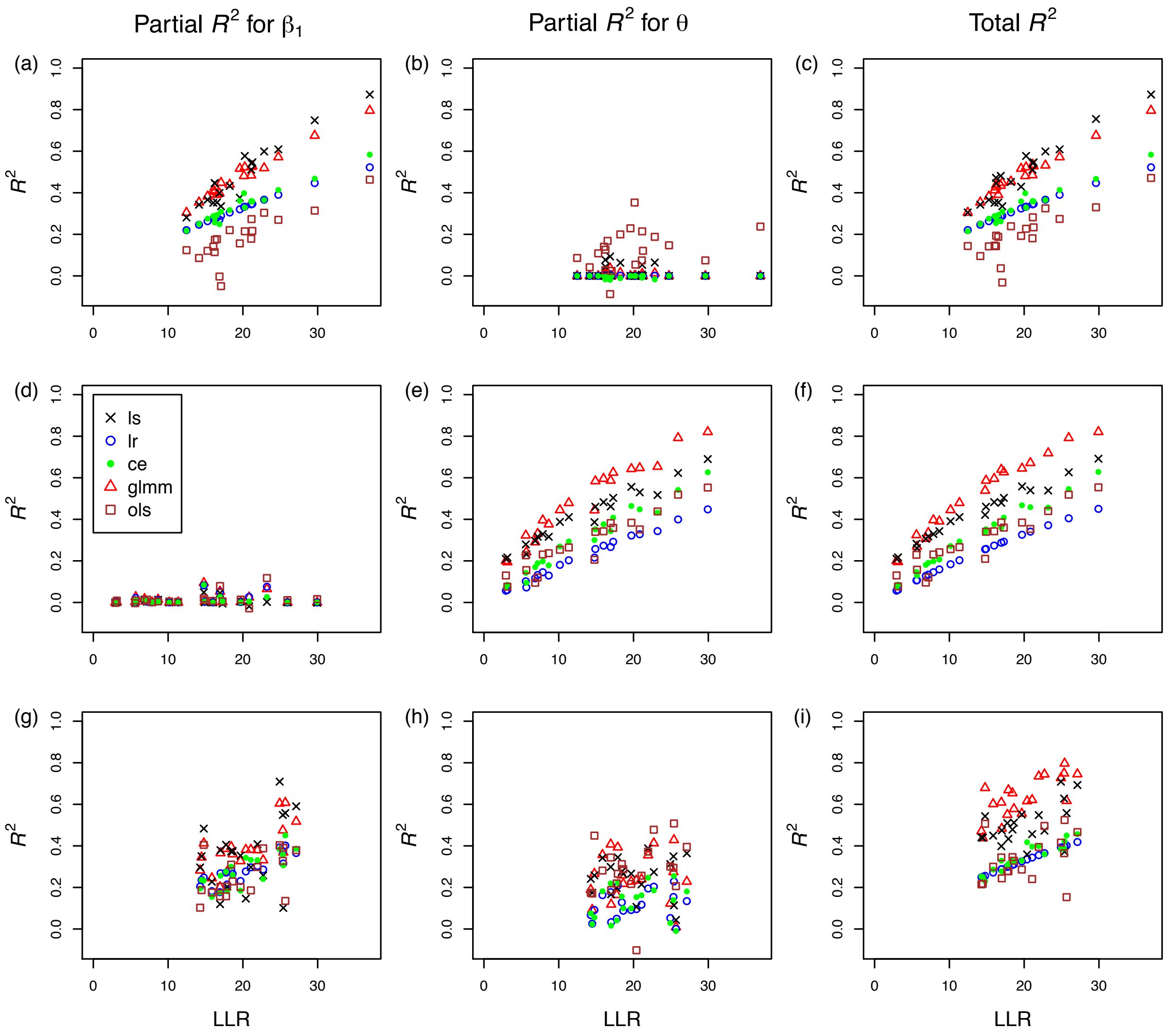

Partial R^2 for β_1
Partial R^2 for θ
Total R^2
(a)
(b)
(c)
(d)
(e)
(f)
mean R^2
sd R^2
Sample size
ls
lr
ce
0.0
0.2
0.4
0.6
0.8
1.0
0.00
0.05
0.10
0.15
0.20
0.25
40
80
120
160

Fig. S7



Partial R^2 for β_1
Partial R^2 for θ
Total R^2
(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
R^2
0.0
0.2
0.4
0.6
0.8
1.0
0
10
20
30
LLR
ls
lr
ce
glmm
ols

Fig. S8

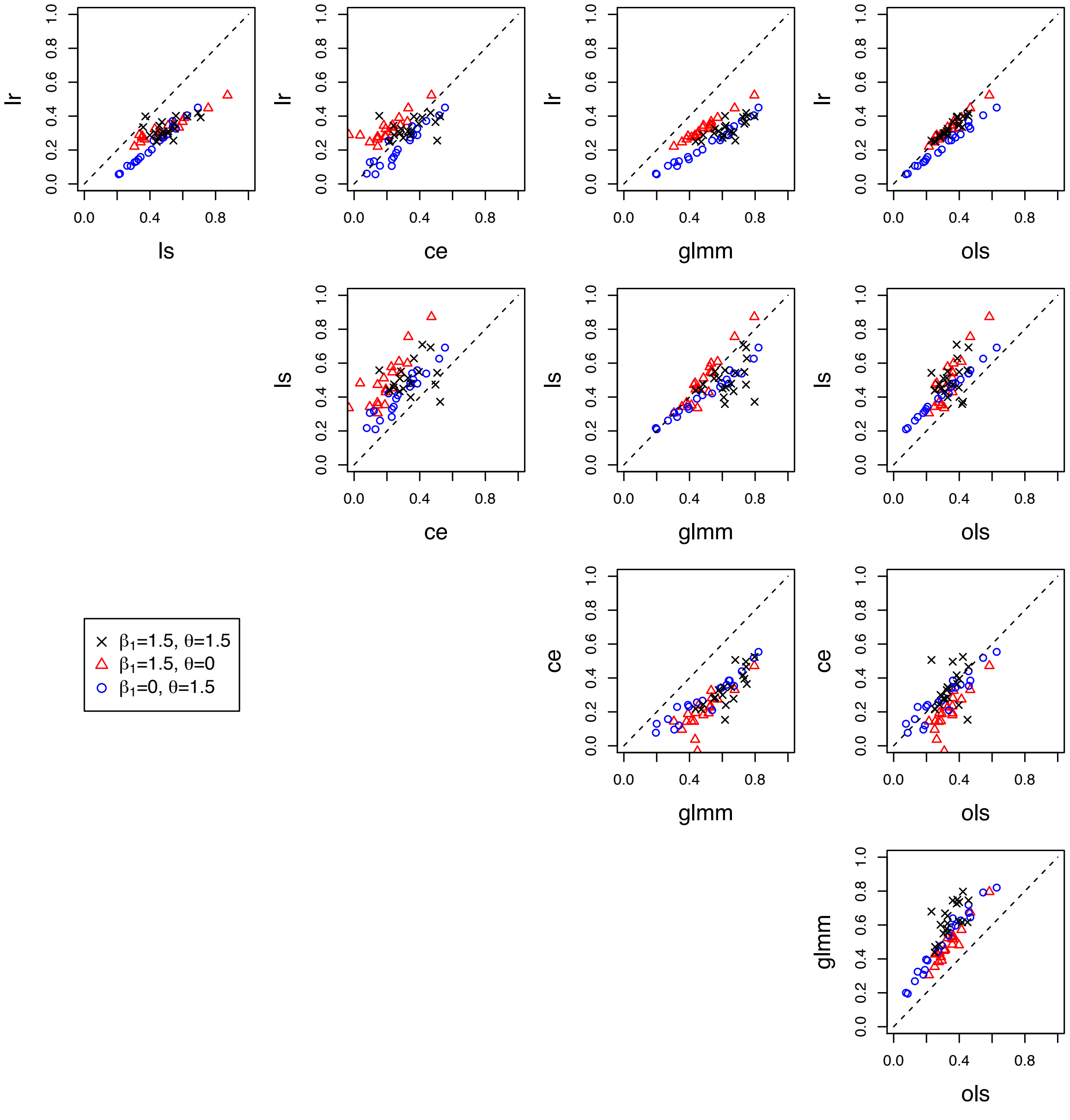
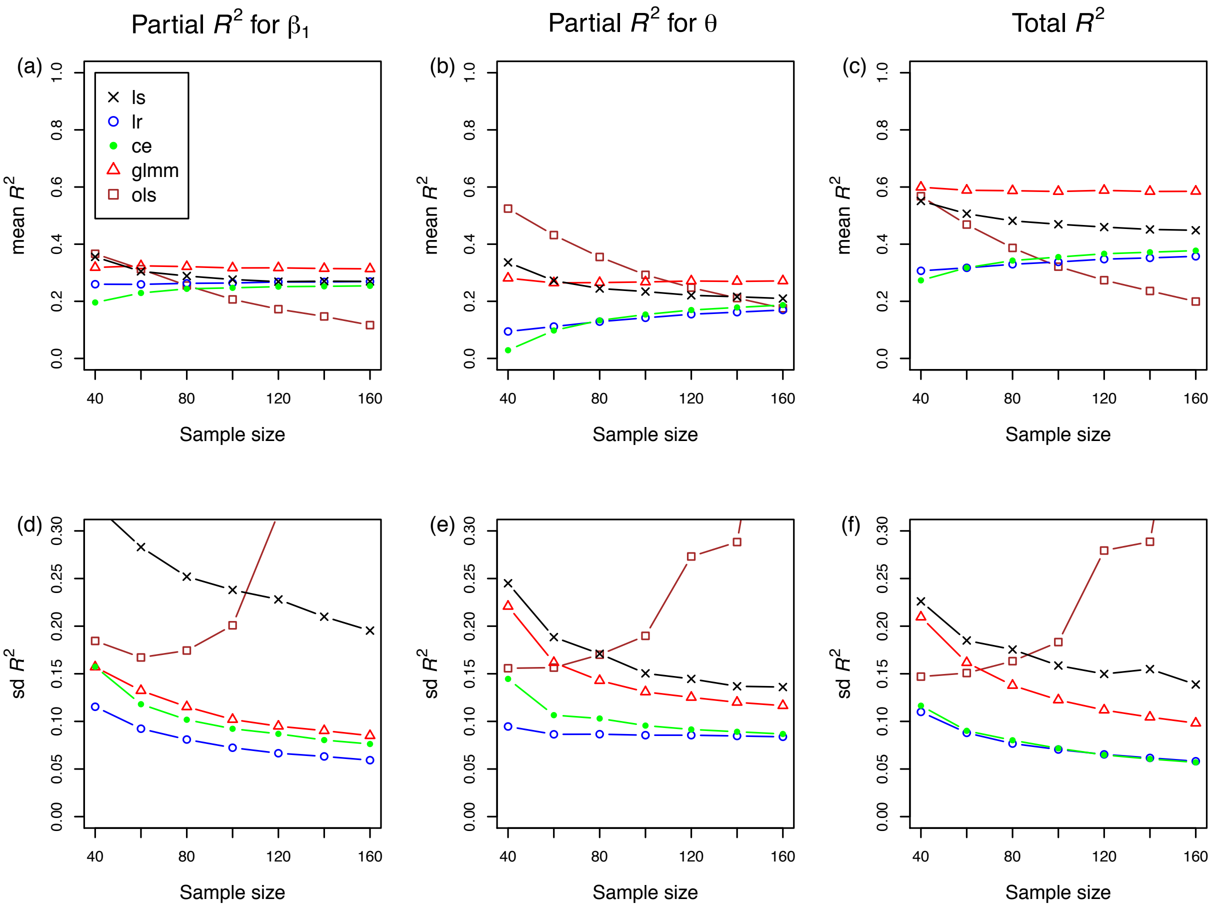


Fig. S9



Partial R^2 for β_1
Partial R^2 for θ
Total R^2
(a)
(b)
(c)
(d)
(e)
(f)
ls
lr
ce
glmm
ols
mean R^2
sd R^2
0.0
0.2
0.4
0.6
0.8
1.0
0.00
0.05
0.10
0.15
0.20
0.25
0.30
40
80
120
160
Sample size

Fig. S10

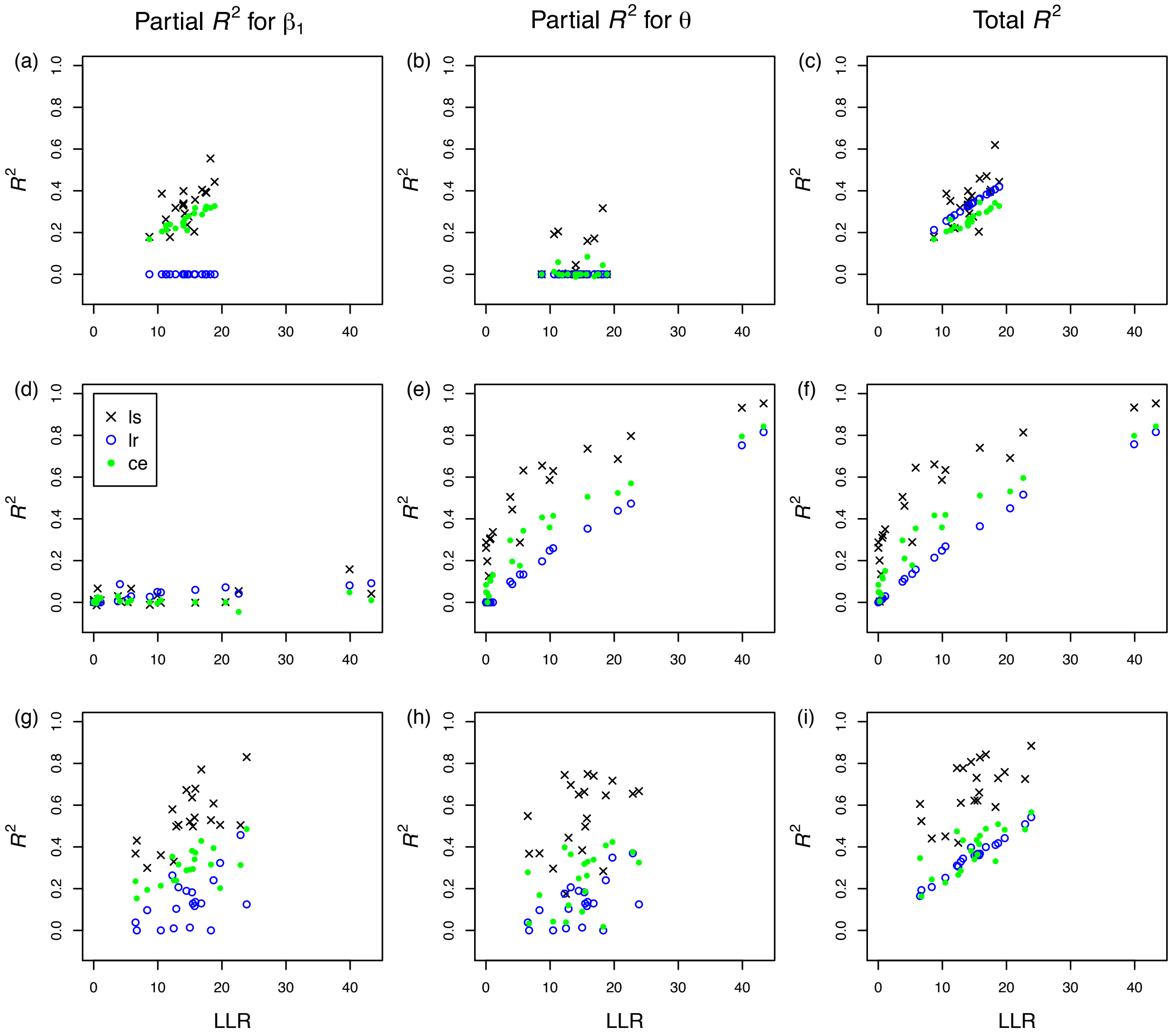

Partial R^2 for β_1
Partial R^2 for θ
Total R^2
(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
R^2
0.0
0.2
0.4
0.6
0.8
1.0
0
10
20
30
40
ls
lr
ce
LLR

Fig. S11

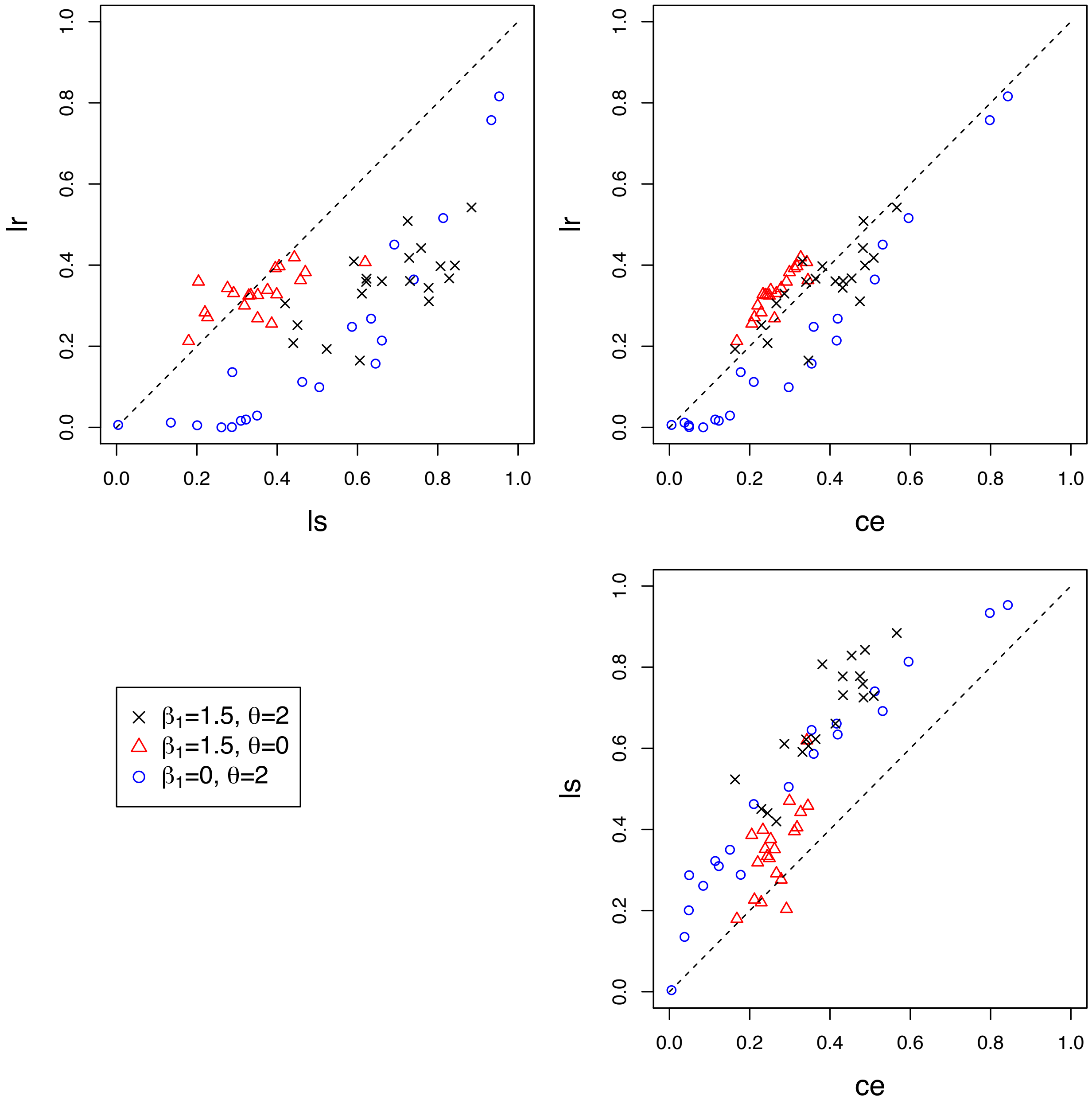
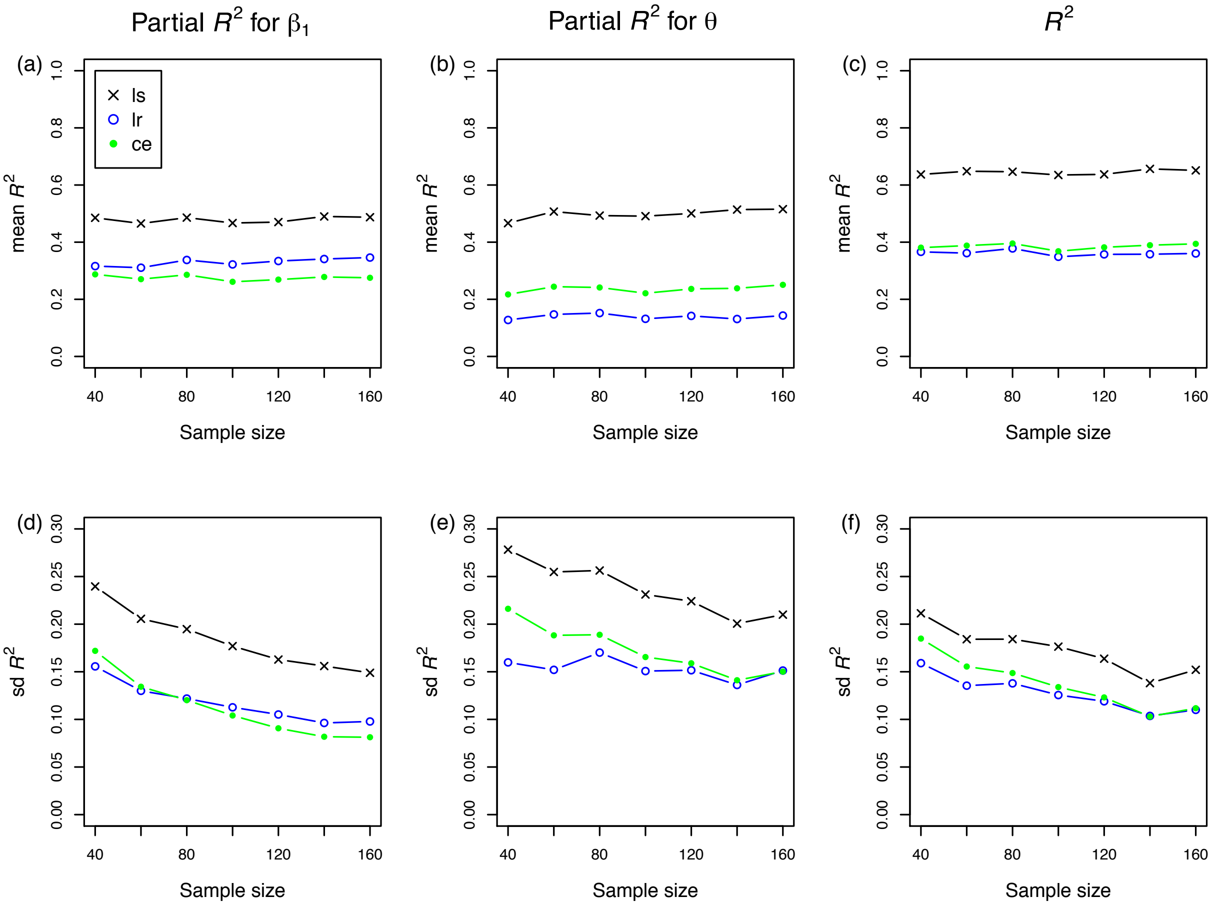

lr
ls
lr
ce
ls
ce
0.0
0.2
0.4
0.6
0.8
1.0
× β_1=1.5, θ=2
△ β_1=1.5, θ=0
○ β_1=0, θ=2

Fig. S12



Partial R^2 for β_1
Partial R^2 for θ
R^2
(a)
(b)
(c)
(d)
(e)
(f)
ls
lr
ce
mean R^2
sd R^2
0.0
0.2
0.4
0.6
0.8
1.0
0.00
0.05
0.10
0.15
0.20
0.25
0.30
40
80
120
160
Sample size