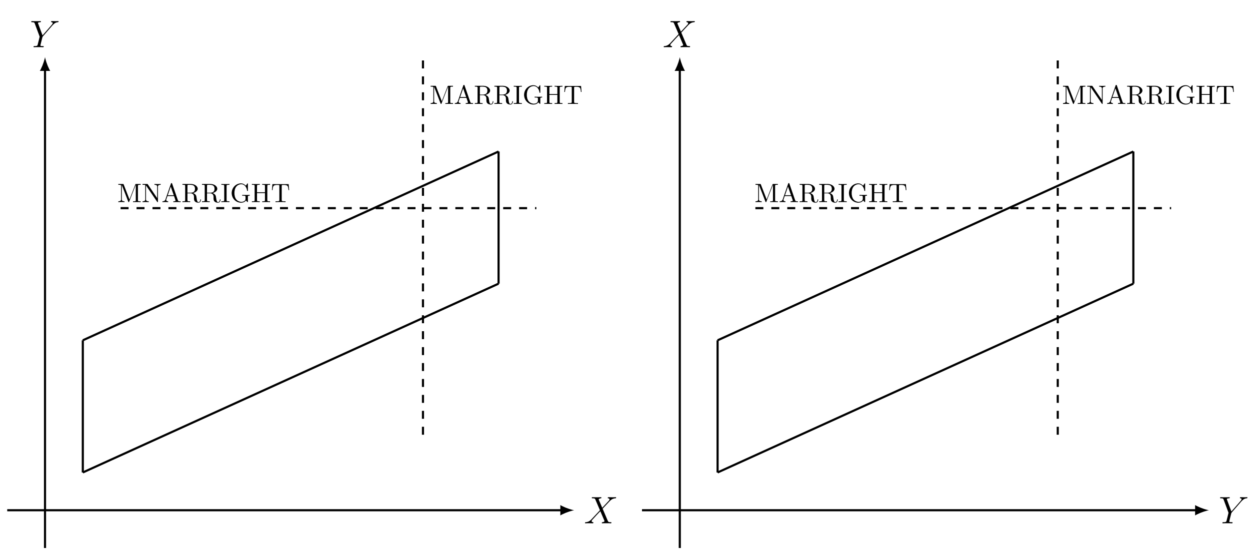
The dance of the mechanisms: β_X and β_Y

Method 1: We perform a model-based simulation by repeatedly drawing $N = 1,000$ cases from a bivariate normal distribution with `mvrnorm` in the package `MASS` with mean vector $\{5, 5\}$ and we vary the correlation between X and Y with ρ in $\{0.1, 0.2, \ldots, 0.9\}$. We induce missingness in variable Y. Variable Y is regressed on X ($Y \sim X$). Number of replications is 100. Evaluation of β_X.
Results 1: Figures 2 to 4. Bias in β_X increases when data correlations increase from $\rho = 0.1$ to $\rho = 0.6$. Then, the bias decreases when data correlations increase more to $\rho = 0.9$. The coverage rate stabilizes when $\rho \geq 0.6$. With CCA, MAR mechanisms do not seem to effect β_X, but MNAR mechanisms do (see Figure 1a).

Method 2: Same data as above. We induce missingness in variable Y. **But!** Variable X is regressed on Y ($X \sim Y$). Evaluation of β_Y.
Results 2: Here, Figure 1b applies. Since the missingness is in Y, MNAR mechanisms do not effect β_Y with CCA. MAR mechanisms on the other hand will affect β_Y. This is visible in Figures 5 to 7. Note that although MNAR mechanisms do not give bias with CCA, they do give bias (and low coverage) after MI !! Apparently the wrong value are imputed?

Conclusion: To know the effect of a missingness mechanism on a β coefficient, it is important to know whether the missing values are in the dependent or independent variable.



(a) (b)

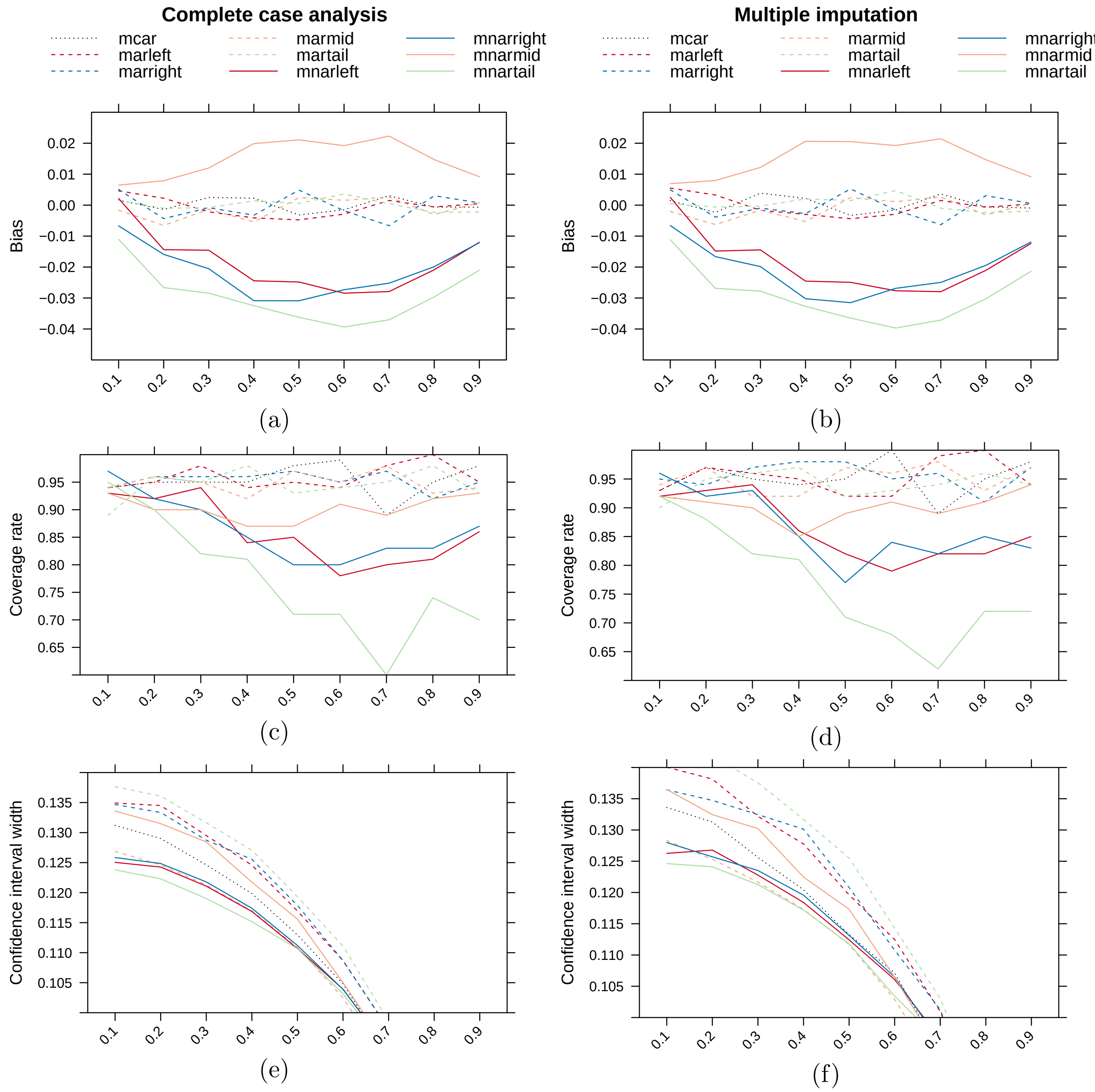


Figure 2: Coverage rate, average bias and average confidence interval width of β_X for complete case analysis and multiple imputation by Bayesian linear regression imputation. The x-axis displays the correlations between X and Y for all ρ in $\{0.1, 0.2, \ldots, 0.9\}$. Missingness proportion is 0.1.

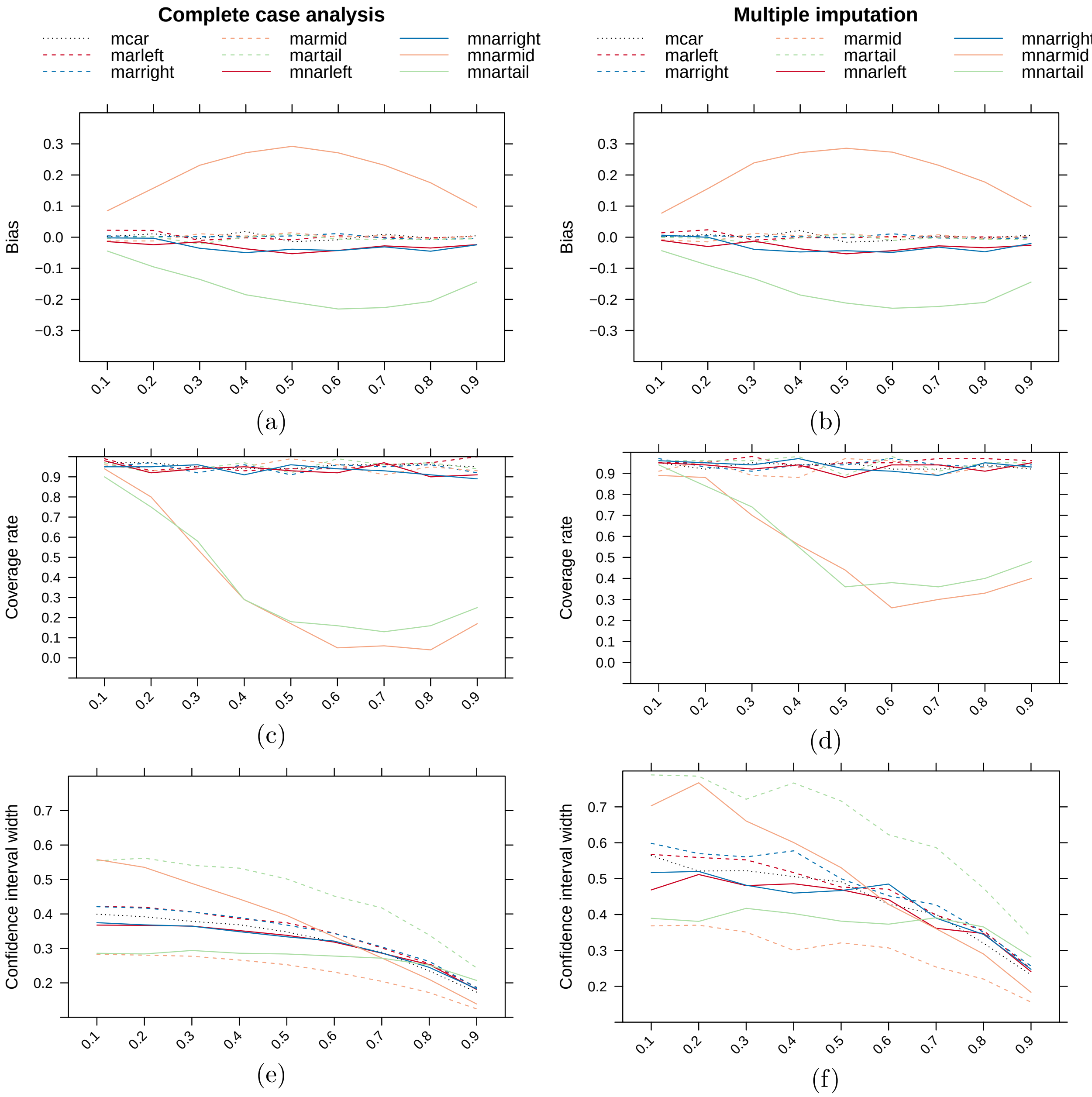


Figure 3: Coverage rate, average bias and average confidence interval width of β_X for complete case analysis and multiple imputation by Bayesian linear regression imputation. The x-axis displays the correlations between X and Y for all ρ in $\{0.1, 0.2, \ldots, 0.9\}$. Missingness proportion is 0.9.

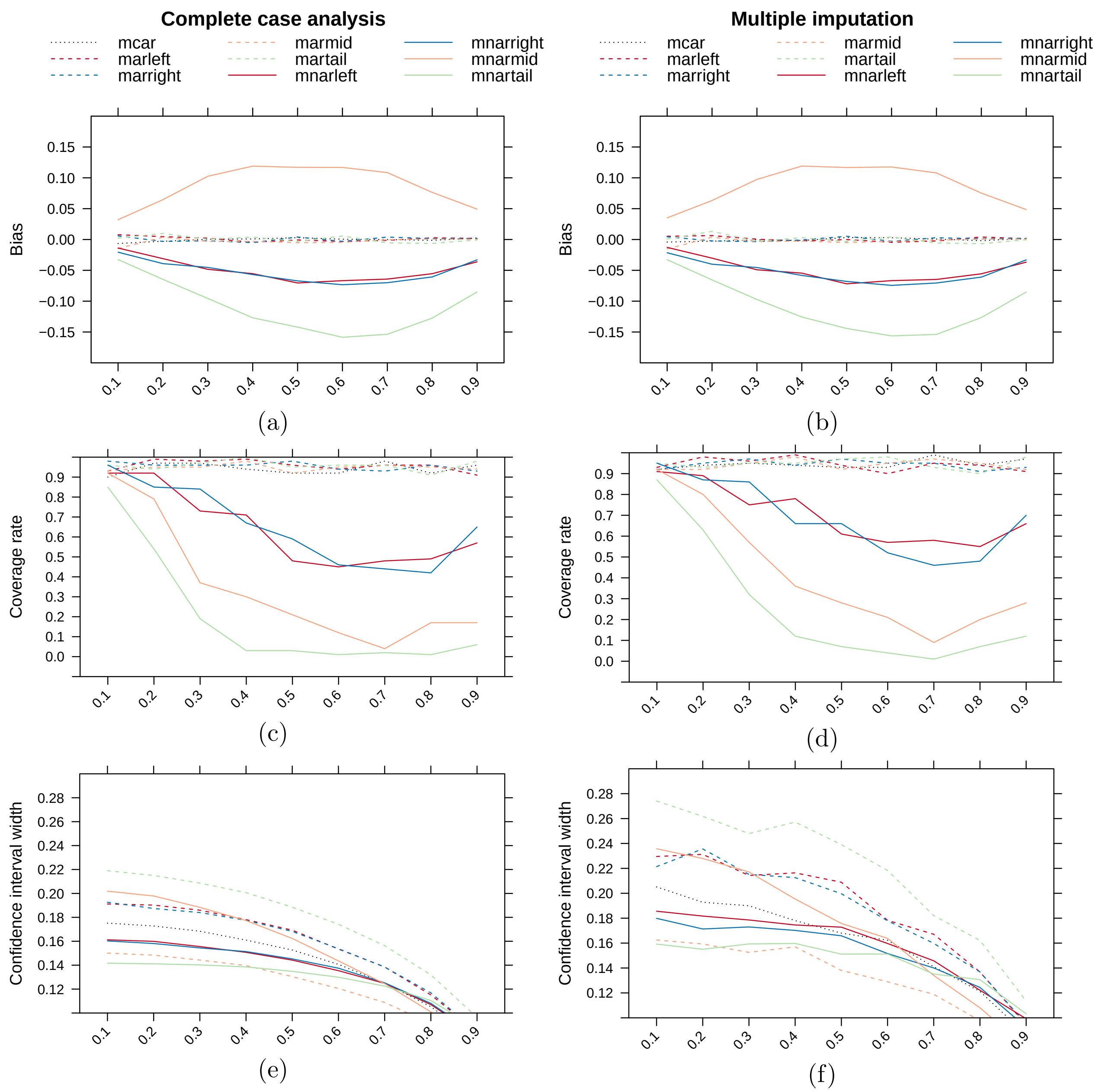


Figure 4: Coverage rate, average bias and average confidence interval width of β_X for complete case analysis and multiple imputation by Bayesian linear regression imputation. The x-axis displays the correlations between X and Y for all ρ in $\{0.1, 0.2, \ldots, 0.9\}$. Missingness proportion is 0.5.

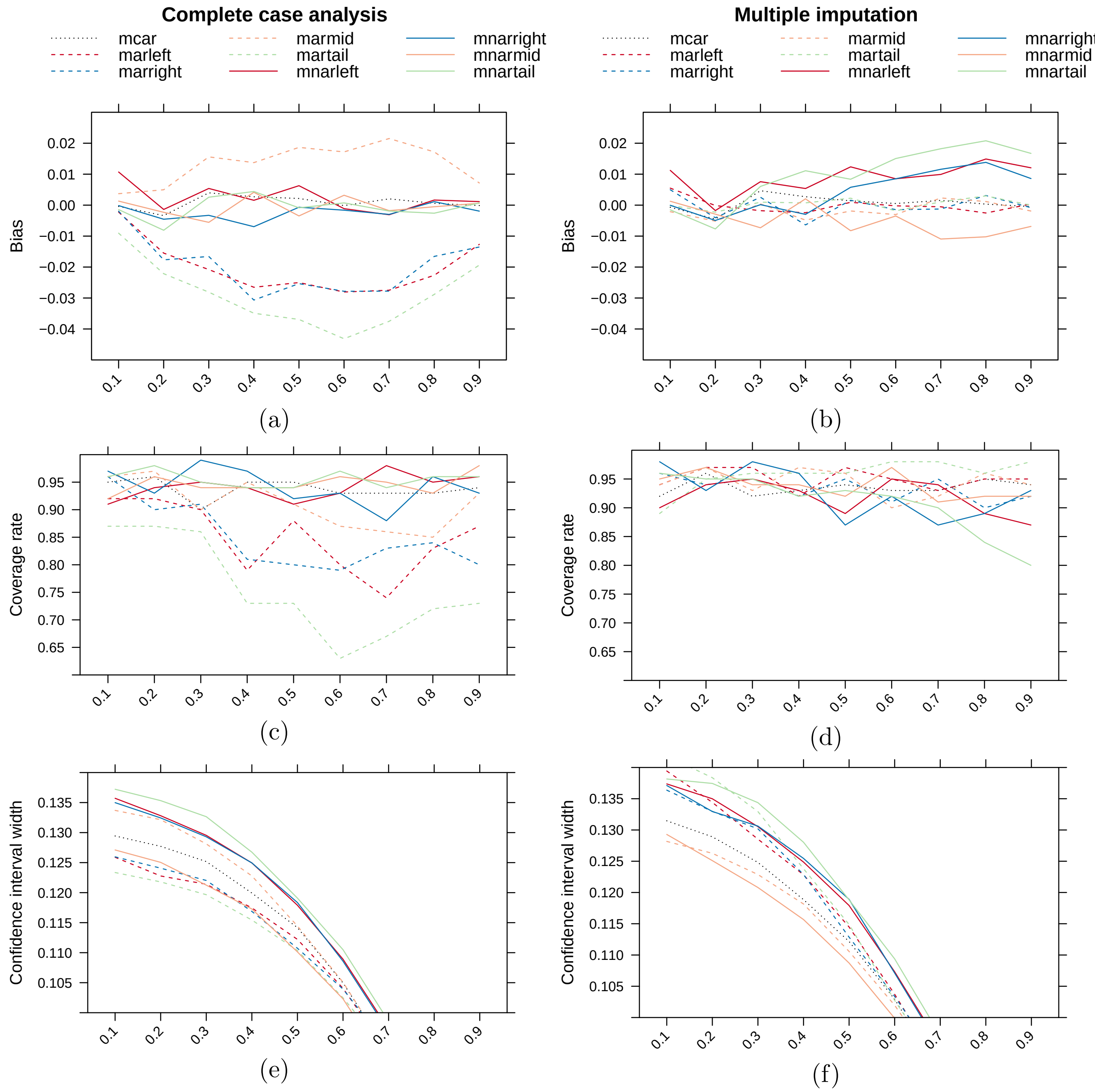


Figure 5: Coverage rate, average bias and average confidence interval width of β_Y for complete case analysis and multiple imputation by Bayesian linear regression imputation. The x-axis displays the correlations between X and Y for all ρ in $\{0.1, 0.2, \ldots, 0.9\}$. Missingness proportion is 0.1.

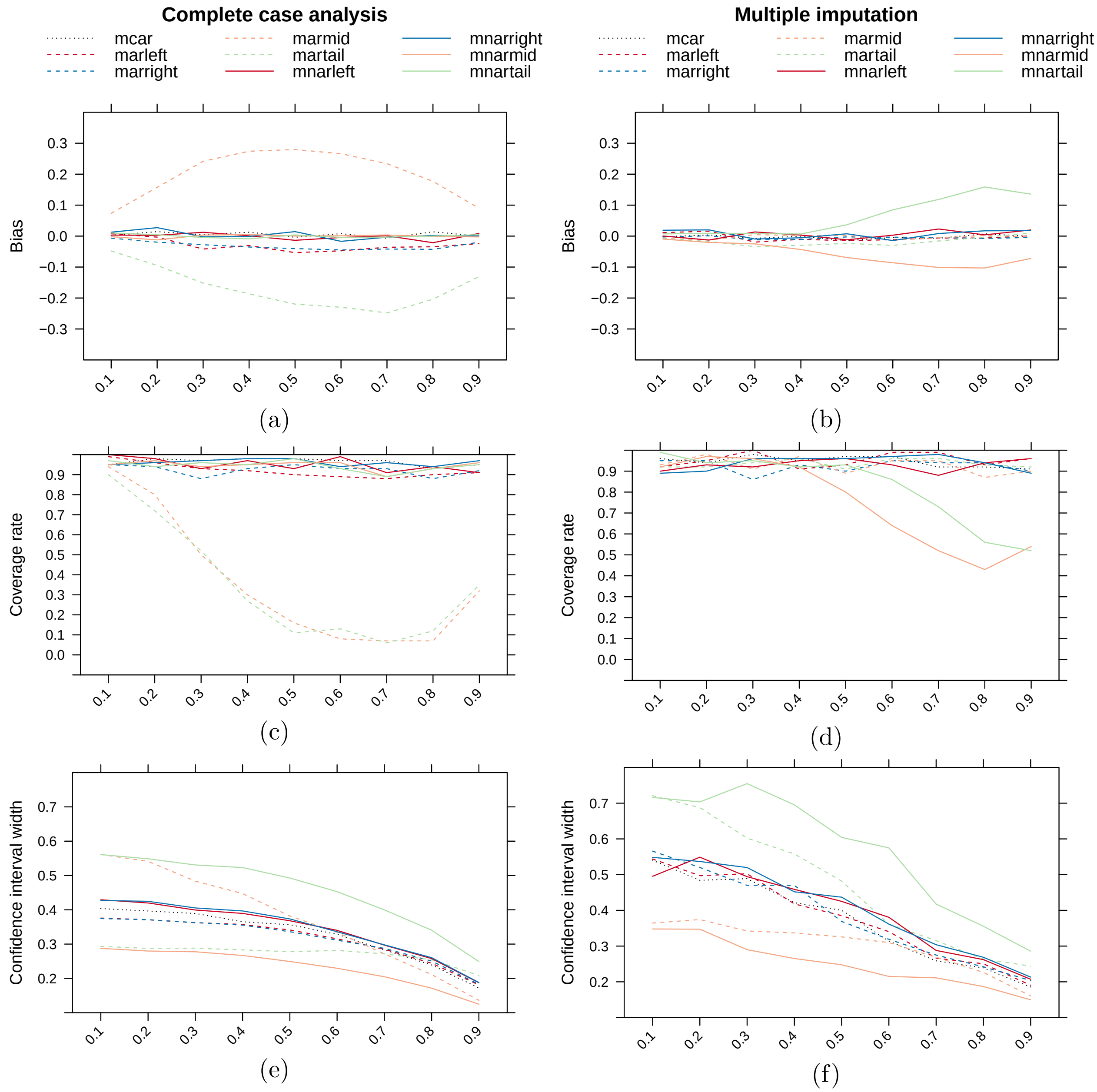


Figure 6: Coverage rate, average bias and average confidence interval width of β_Y for complete case analysis and multiple imputation by Bayesian linear regression imputation. The x-axis displays the correlations between X and Y for all ρ in $\{0.1, 0.2, \ldots, 0.9\}$. Missingness proportion is 0.9.

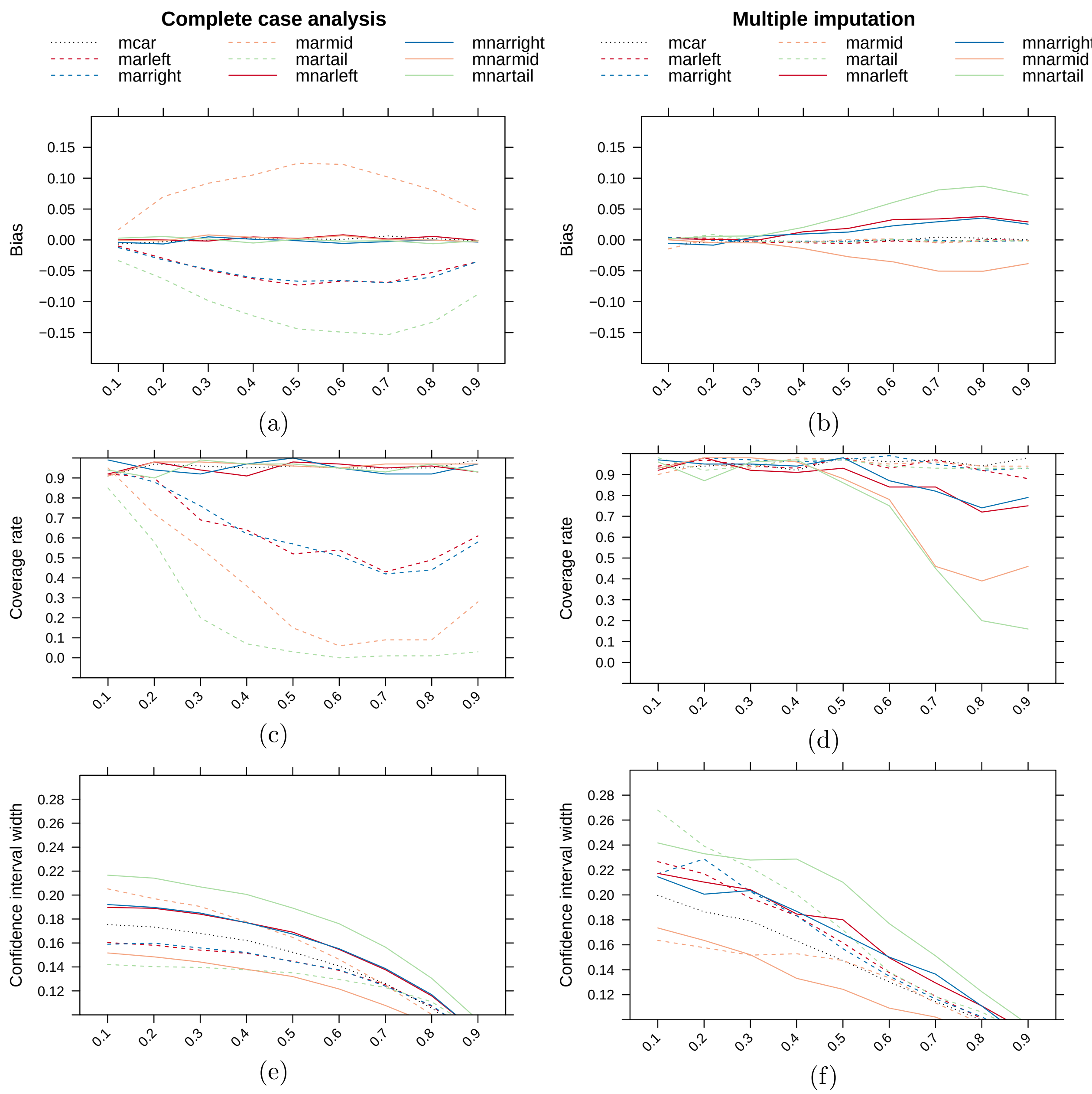


Figure 7: Coverage rate, average bias and average confidence interval width of β_Y for complete case analysis and multiple imputation by Bayesian linear regression imputation. The x-axis displays the correlations between X and Y for all ρ in $\{0.1, 0.2, \ldots, 0.9\}$. Missingness proportion is 0.5.