Electronic Supplementary Material for

Vijayakumar, R., & Cheung, M. W.-L. (2018). Replicability of Machine Learning Models in the Social Sciences. *Zeitschrift für Psychologie*. https://doi.org/10.1027/2151-2604/a000344

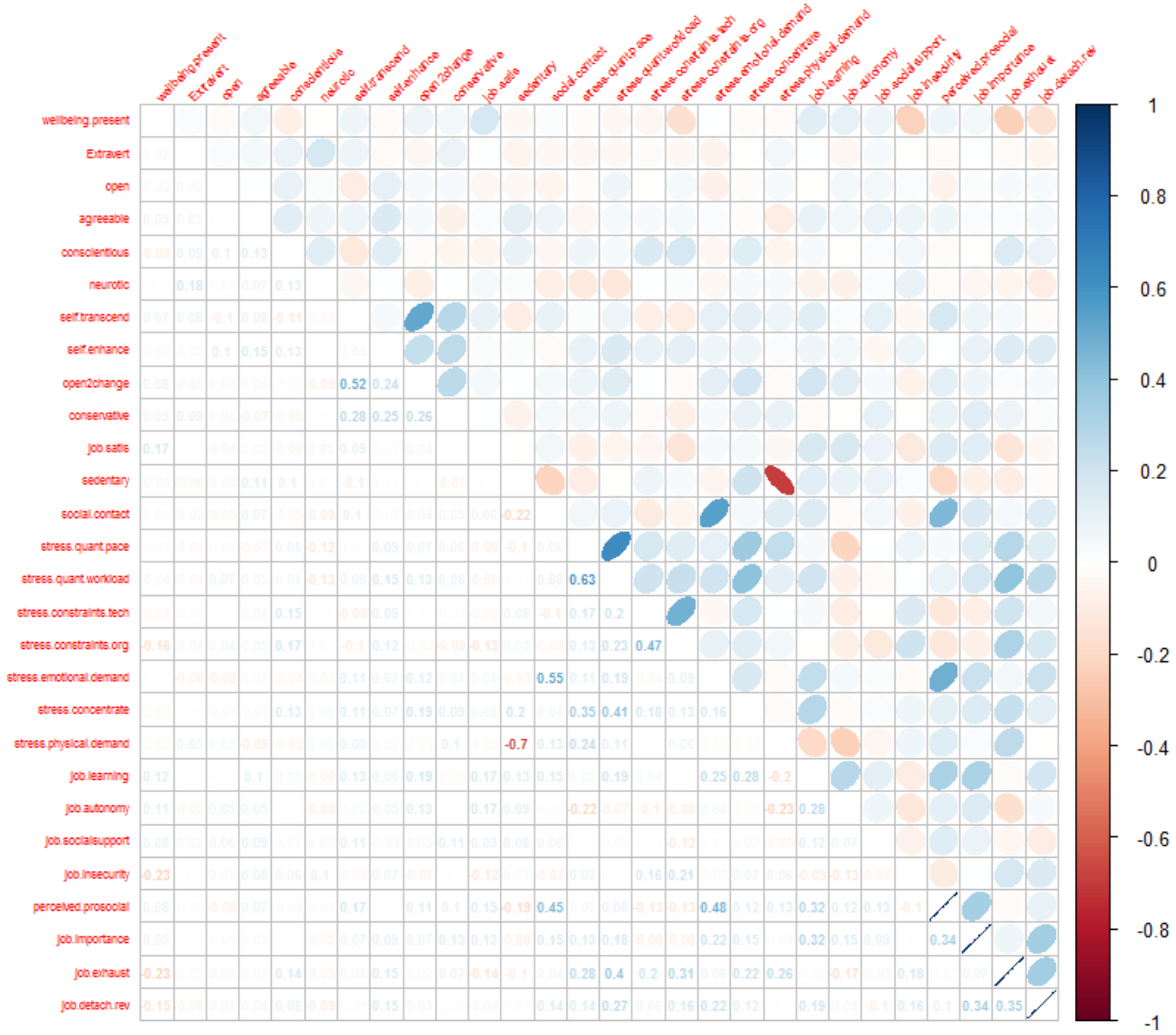

Figure 1. Correlation matrix for variables in the GESIS dataset. *Well-being.present* is the dependent variable in our study. The inter-predictor correlations form this dataset are used to generate the simulated data.

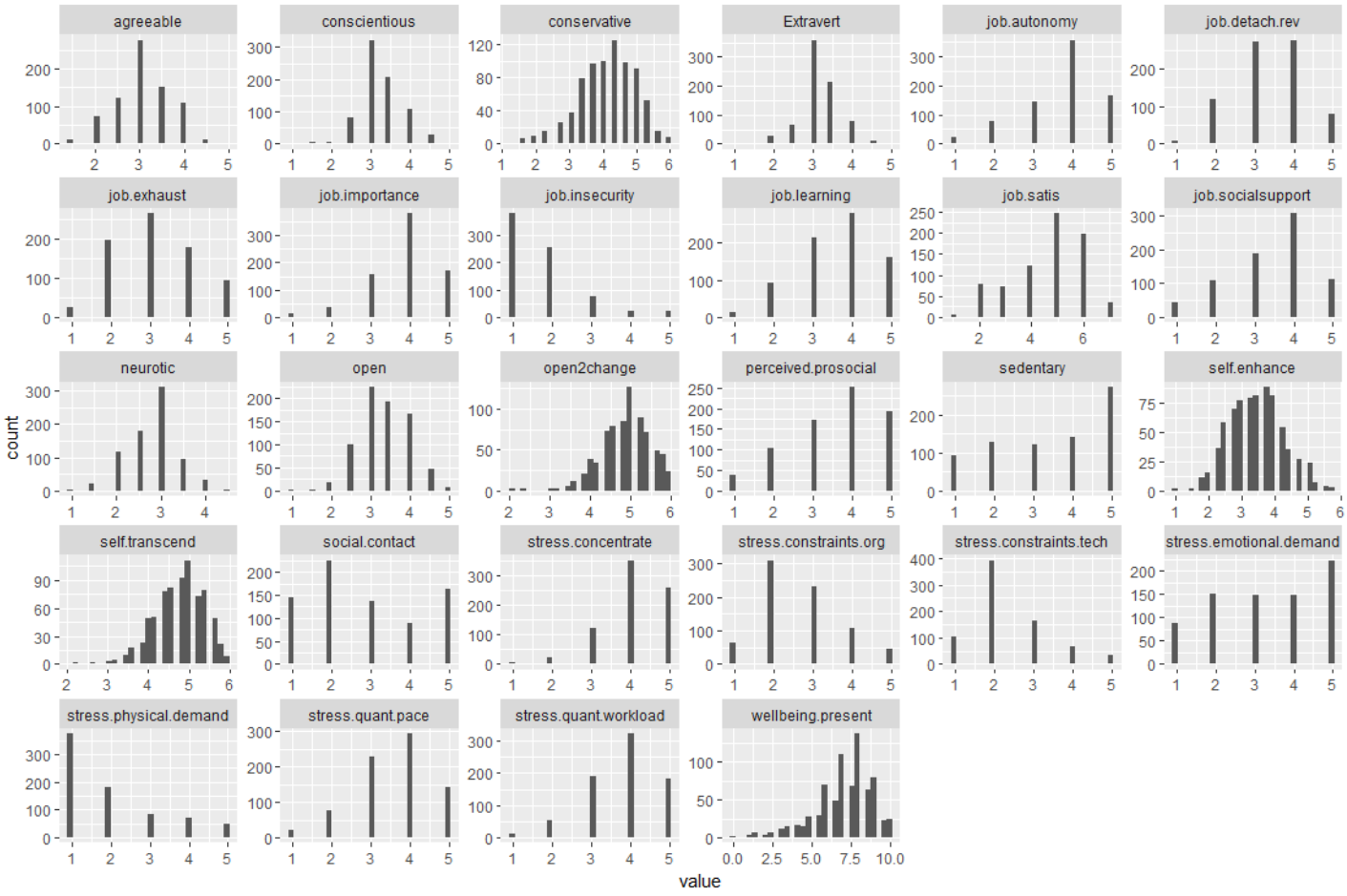

Figure 2. Histogram of all variables in the GESIS dataset. *Well-being.present* is the dependent variable in our study.

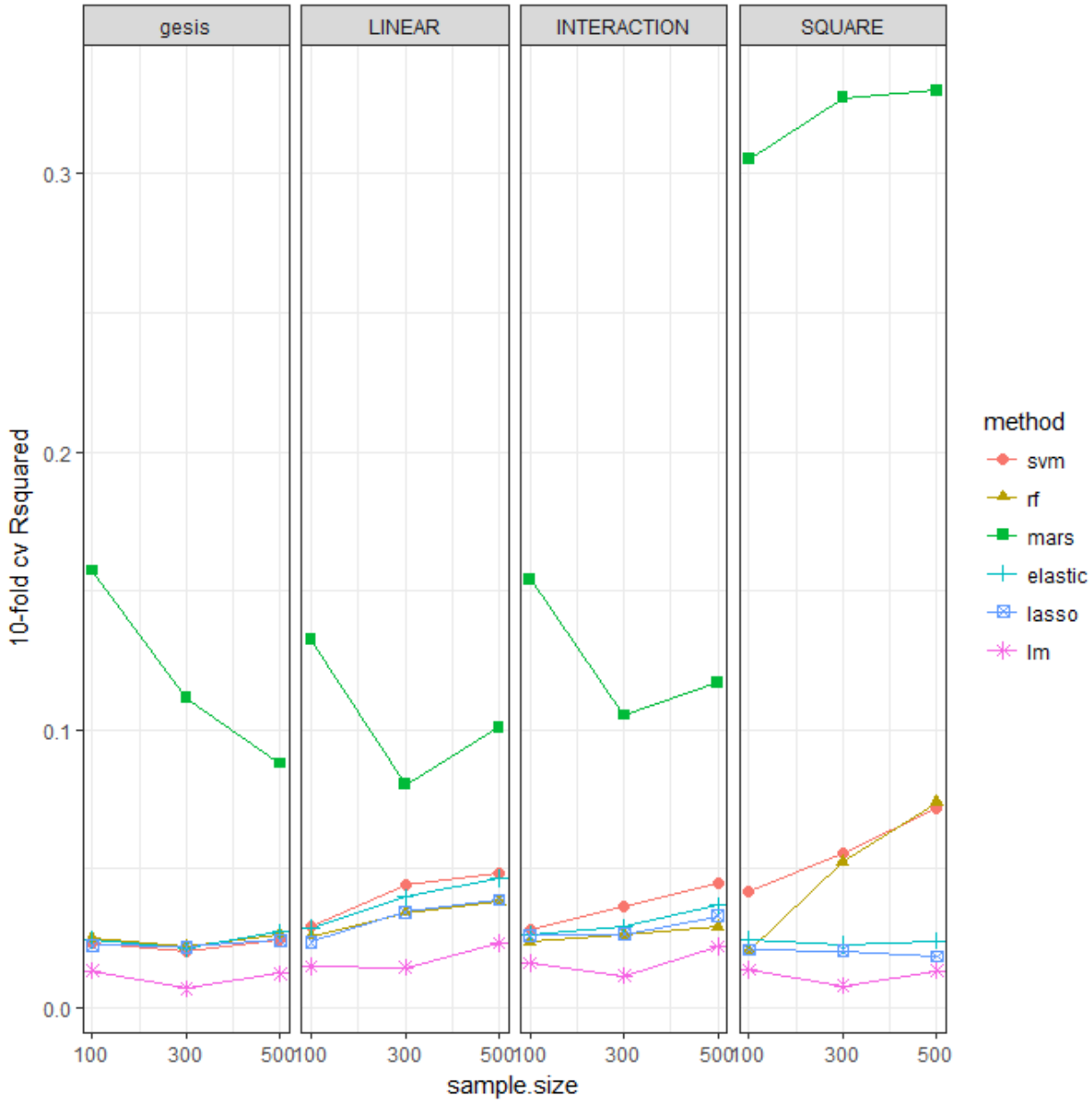

Figure 3. 10 fold-CV R^2 across sample sizes for empirical (GESIS) and simulated (linear, interaction and square) datasets. The labels 'rf' and 'lm' stand for *random forest* and *linear regression model* respectively. MARS is superior to all other methods, with SVM and RF performing better than regression-based methods for square data. Regression performs worst in all scenarios.

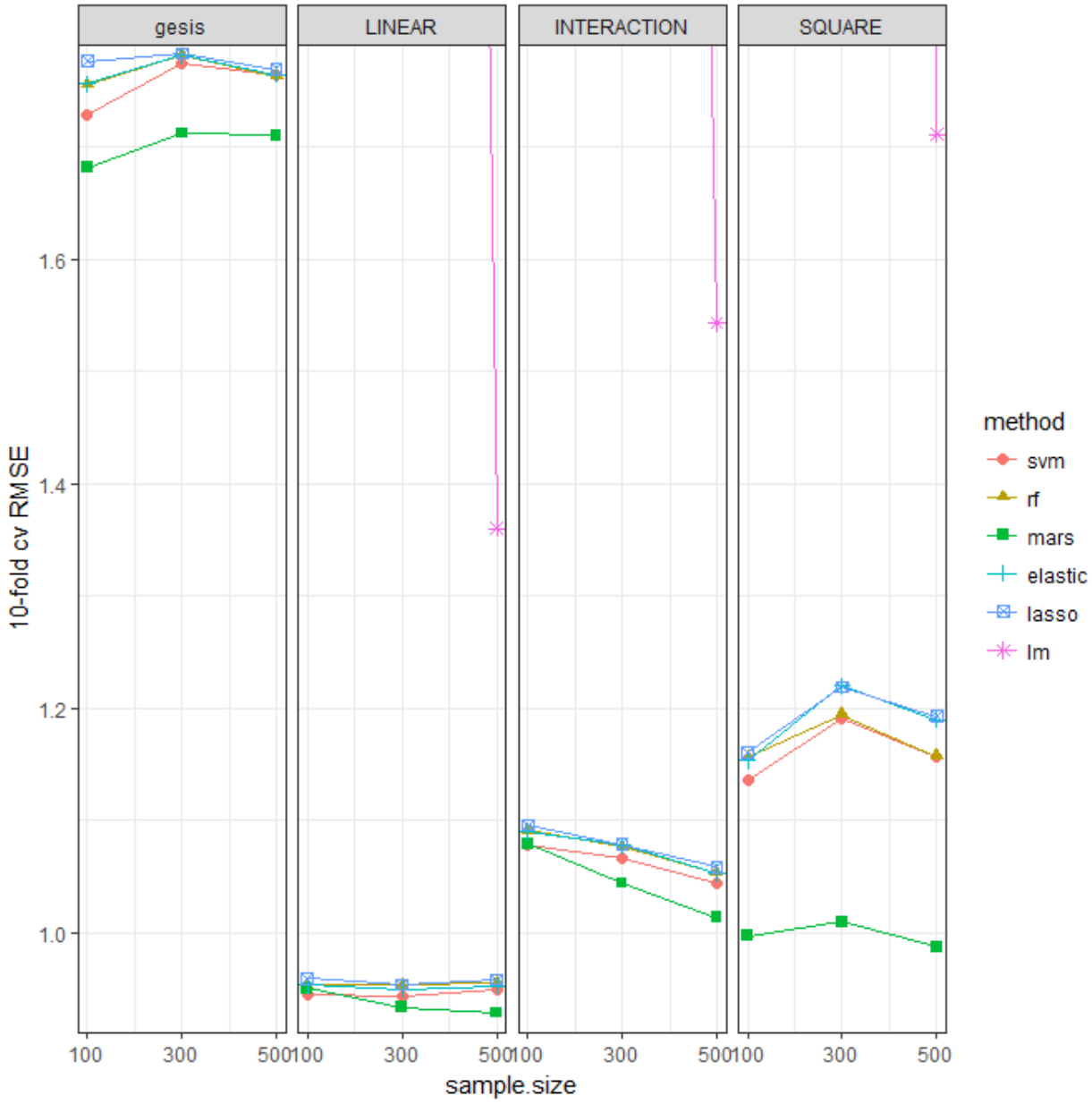

Figure 4. 10 fold-CV RMSE across sample sizes for empirical (GESIS) and simulated (linear, interaction and square) datasets. The labels 'rf' and 'lm' stand for *random forest* and *linear regression model* respectively. MARS is superior to all other methods, with SVM and RF performing better than regression-based methods for square data. Regression performs worst in all scenarios, with many RMSE values so large they could not be included in the graph (all sample sizes in GESIS data, and sample sizes less than 500 in simulated data).