## Robust model selection: New developments in the *R* package robust HD

## Andreas Alfons<sup>\*</sup>

Erasmus School of Economics, Erasmus Universiteit Rotterdam

\*Contact author: alfons@ese.eur.nl

**Keywords:** Outliers, Robust groupwise LARS, Sparse S-regression, Sparse MM-regression, C++

Variable selection is a common task in regression analysis to improve prediction performance by variance reduction, and to increase interpretability of the resulting models due to the smaller number of variables. In the presence of outliers, robust methods are necessary to prevent unreliable results. The R package **robustHD** [2] provides functionality for robust linear model selection with a focus on methods for high-dimensional data. New developments include robust groupwise least angle regression, sparse S-regression and sparse MM-regression. The package implements an object-oriented design, while large parts of the code are written in C++ to reduce computing time. Cross-validation functionality to select the final model is implemented via package **perry** [1] such that taking advantage of parallel computing is easy. In addition, diagnostic plots to evaluate the model selection procedures are available in **robustHD**.

## References

- [1] Alfons, A. (2013). **perry**: Resampling-based prediction error estimation for regression models. R package.
- [2] Alfons, A. (2014). robustHD: Robust methods for high-dimensional data. R package.