Online Appendix - A Note on the Structural Change Test in Finite Samples: Using a Permutation Approach to Estimate the Sampling Distribution

Huth, K.[1,2,3], Waldorp, L. J.[1], Luigjes, J.[2], Goudriaan, A. E. [2,4], van Holst, R. J. [2,3], & Marsman, M.[1]

1 Department of Psychology, University of Amsterdam
2 Department of Psychiatry, Amsterdam University Medical Center
3 Centre for Urban Mental Health, University of Amsterdam
4 Arkin Mental Health Institute

This is the online appendix for the paper "A Note on the Structural Change Test in Finite Samples: Using a Permutation Approach to Estimate the Sampling Distribution". For details on the project please refer to the original article (INCLUDE LINK TO PREPRINT).

Small Sample Behavior of the Structural Change Test using the Double Maximum and Crámer-von Mises statistics

In this section, we elaborate on the results of the p-value and sampling distribution for structural change tests (SCTs) when applying the double maximum (DM; see Eq. (2) in the original paper) and the Cramér-von Mises (CvM; see Eq. (3) in the original paper) test statistic. Data were simulated for a linear regression model with two, four, and eight covariates k, for n is 50, 200, and 1000 cases. We sampled 5,000 datasets for each combination of k, and n.

The simulated p-value distributions are shown in Figure 1. The simulated p-value distributions in the top row are for the DM statistic, and those in the bottom row are for the CvM statistic. Observe that the p-values for the DM statistic do not follow the expected uniform distribution —shown as a black, dashed line— in the smaller sample sizes but converge to it as the sample size grows. The simulated p-value distributions for the CvM statistic also do not resemble a uniform distribution in smaller sample sizes. However, contrary to what we find for the other statistics, the p-value distribution also does not resemble a uniform distribution for larger sample sizes. The CvM p-value distributions appear to converge much slower than those of the other two statistics.

The simulated sampling distributions are shown in Figure 2 for the DM statistic and in Figure 3 for the CvM statistic. The asymptotic sampling distributions are indicated

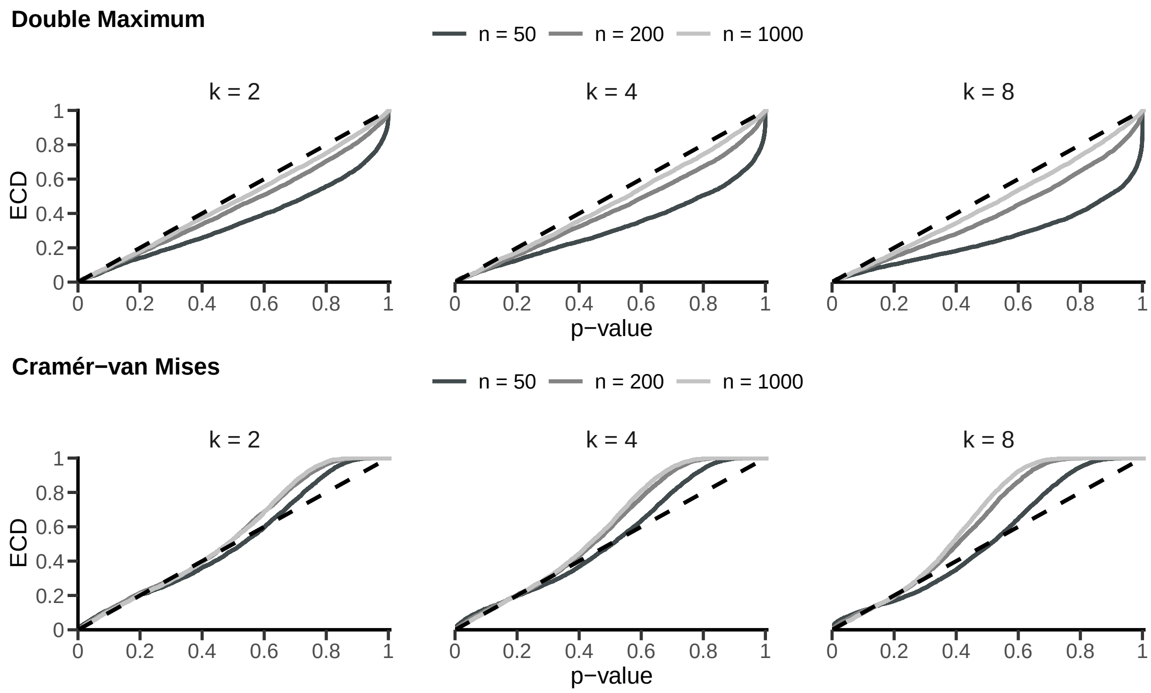

Figure 1. Empirical cumulative distributions for the p-value under the null-hypothesis for a linear regression model for different test statistics and simulation settings. The top row shows the DM results and the bottom row the CvM results. In each plot, the black, dashed line represents the expected uniform distribution.

with a black solid line in these graphs. They were generated by repeatedly simulating values from a Brownian bridge and then computing the statistic on the generated data (e.g., see Andrews, 1993; Zeileis, 2006). For both statistics, the theoretical sampling distributions do not match the empirical sampling distributions for smaller sample sizes, but their fit improves for the larger sample sizes.

To conclude, the p-value and sampling distributions for the SCT appear to be misspecified for the three test statistics investigated in this paper. This means that the problems are not isolated to the use of one of the test statistics.

References

Andrews, D. (1993). Tests for parameter instability and structural change with unknown change point. *Econometrica*, *61*, 821–856.

Zeileis, A. (2006). Implementing a class of structural change tests: An econometric computing approach. *Computational Statistics & Data Analysis*, *50*, 2987–3008. doi: 10.1016/j.csda.2005.07.001

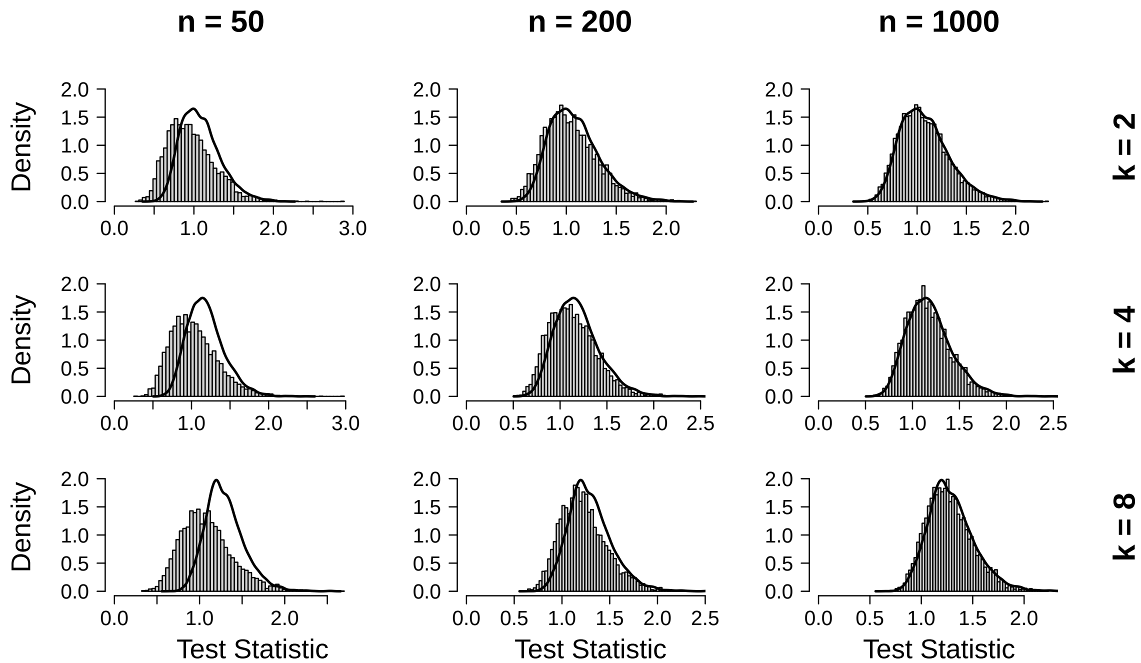

Figure 2. Distributions of the DM statistic under the null hypothesis for the linear regression model. The expected sampling distribution is depicted as a black line and was obtained by simulating observations from a Brownian bridge and applying the DM statistic to them (e.g., see Zeileis (2006)).

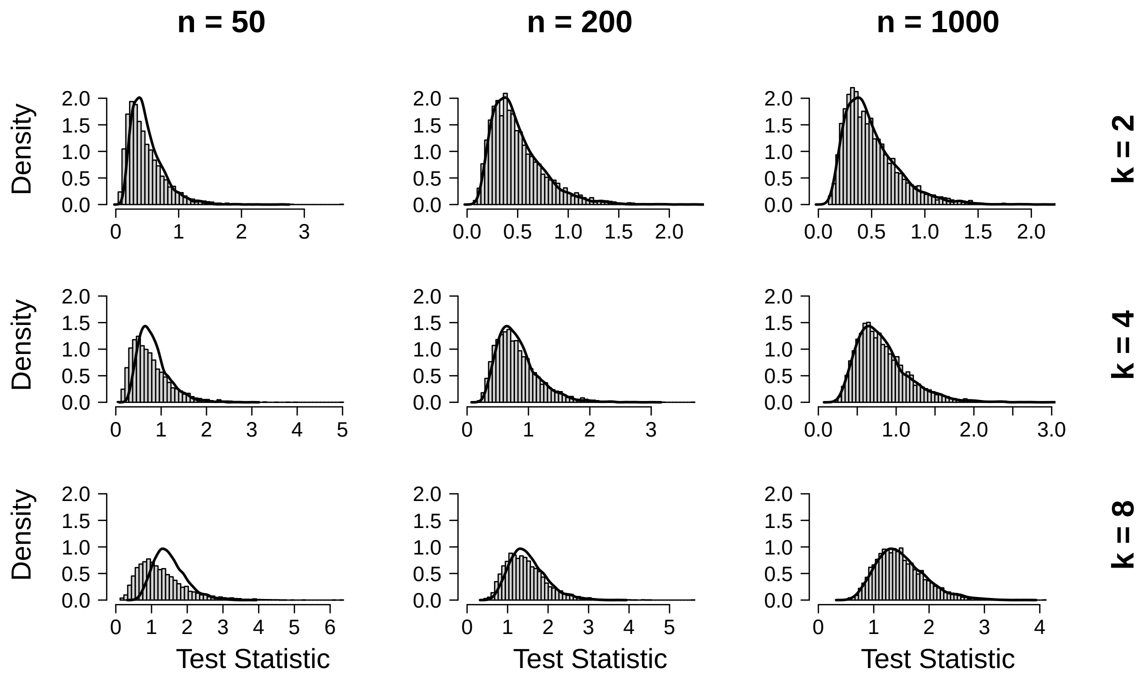

Figure 3. Distributions of the CvM statistic under the null hypothesis for the linear regression model. The expected sampling distribution is depicted as a black line and was obtained by simulating observations from a Brownian bridge and applying the CvM statistic to them (e.g., see Zeileis (2006)).