Supplemental Appendix for "Out-of-Sample Comparisons of Overfit Models"

Gray Calhoun*
Iowa State University

March 28, 2014

This appendix presents simulation results that accompany Calhoun (2014). The Monte Carlo here is exactly the same as presented in Calhoun (2014), except that a recursive window is used. Please read the original paper for further details.

- Figure 1 in this Appendix corresponds to Figure 3 in the original paper.
- Figure 2 corresponds to Figure 4 in the paper.
- Figure 3 corresponds to Figure 5 in the paper.
- Figure 4 corresponds to Figure 6 in the paper.
- Figure 5 corresponds to Figure 7 in the paper.

References

G. Calhoun. Out-of-sample comparisons of overfit models. Working Paper 11002, Iowa State University, 2014.

T. E. Clark and K. D. West. Using out-of-sample mean squared prediction errors to test the martingale difference hypothesis. *Journal of Econometrics*, 135(1):155–186, 2006.

T. E. Clark and K. D. West. Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics*, 138(1):291–311, May 2007.

F. X. Diebold and R. S. Mariano. Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13(3):253–263, 1995.

M. W. McCracken. Asymptotics for out of sample tests of Granger causality. *Journal of Econometrics*, 140(2):719–752, Oct. 2007.

*email: gcalhoun@iastate.edu

K. D. West. Asymptotic inference about predictive ability. *Econometrica*, 64(5):1067–1084, Sept. 1996.

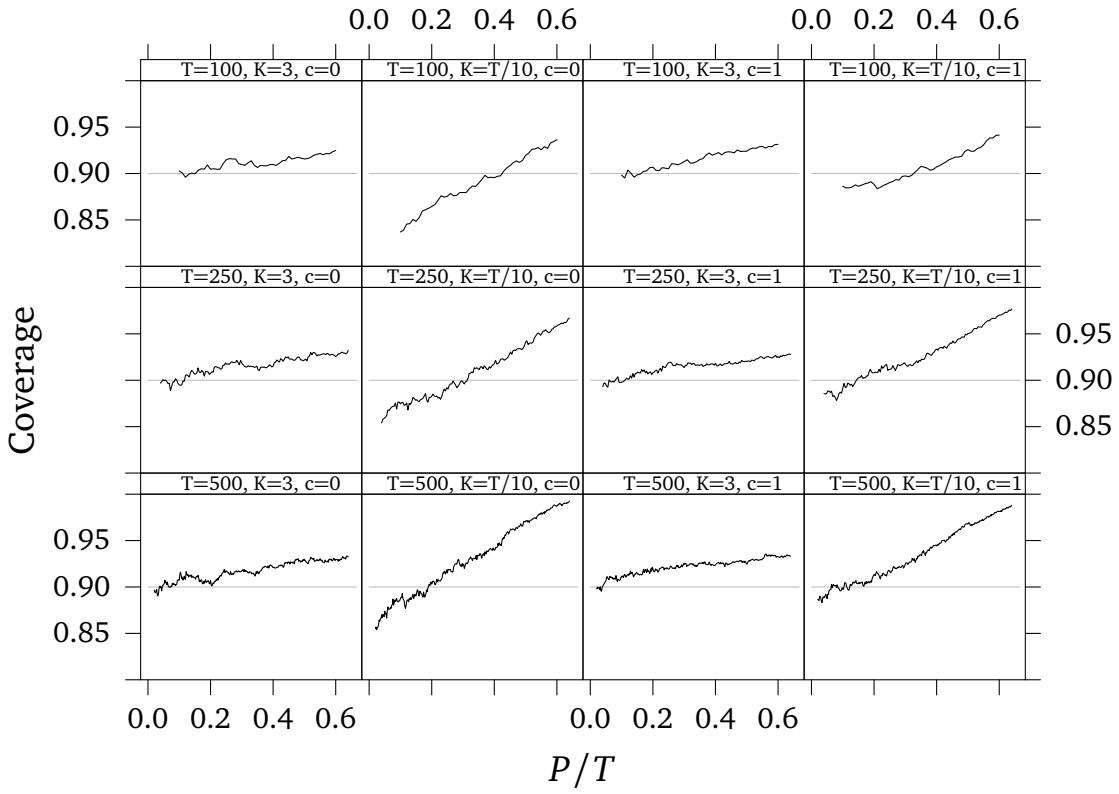

Figure 1: Simulated coverage of $\mathrm{E}_T\,\bar{D}_T$ at 90% confidence using a one-sided interval based on the recursive DMW OOS-t test, plotted as a function of the fraction of observations used in the test sample, P/T. The solid horizontal line denotes the intervals' nominal coverage.

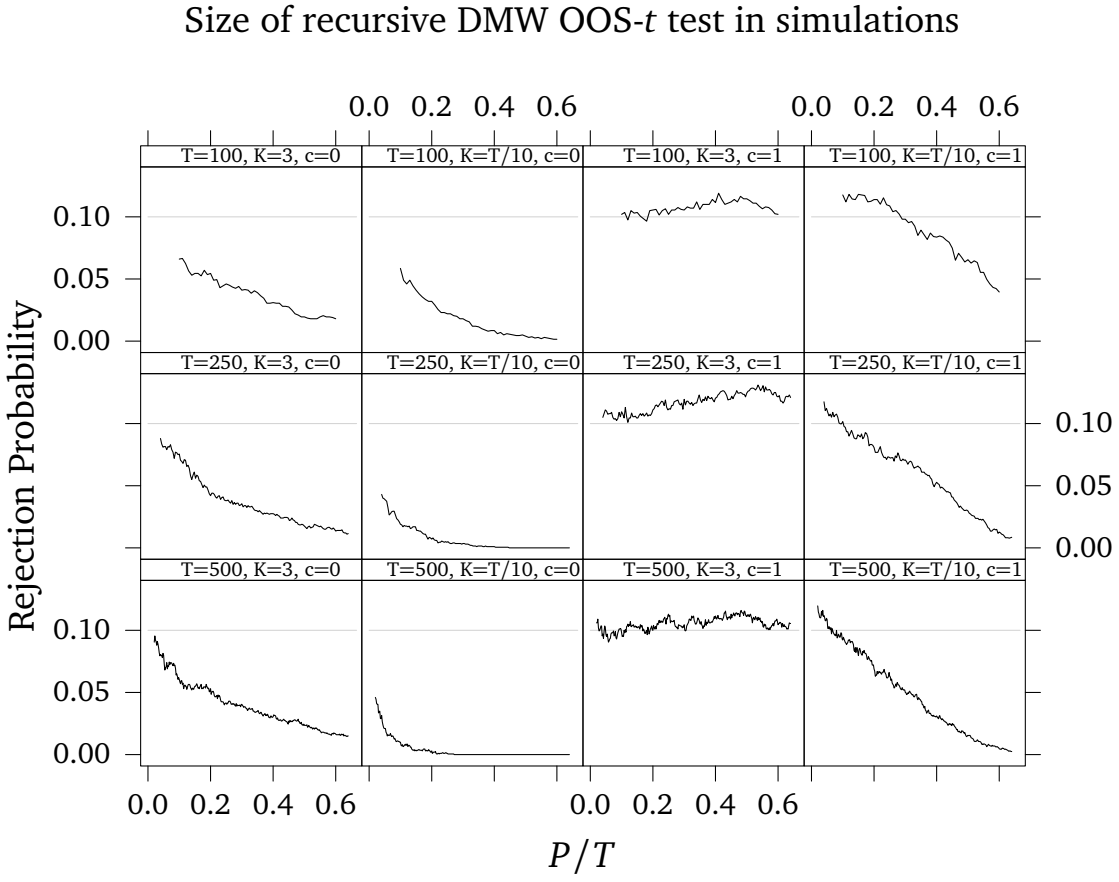

Figure 2: Simulated rejection probabilities for the recursive DMW OOS-t test under the null hypothesis that the benchmark model is more accurate, $\mathrm{E}_T \bar{D}_T \leq 0$. Nominal size is 10% and is marked with a solid horizontal line. Values greater than 10% indicate that the test rejects the benchmark model too often.

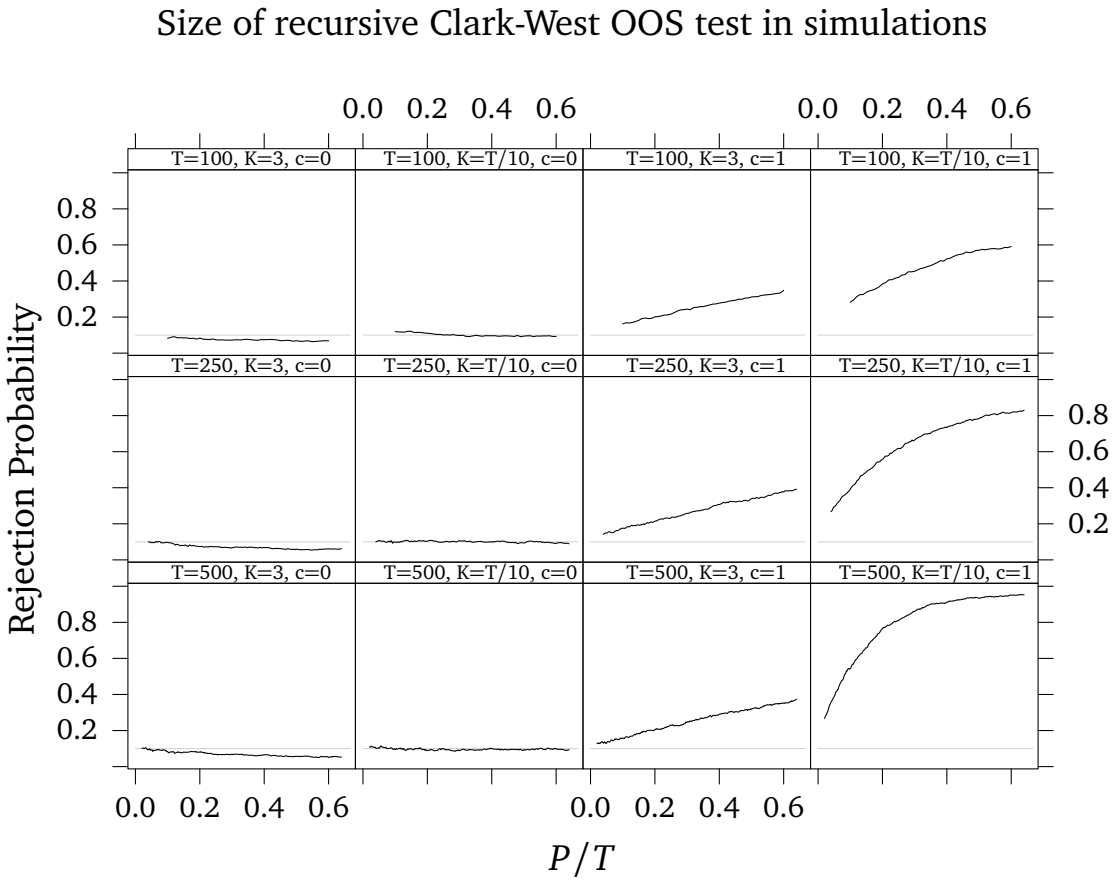

Figure 3: Simulated rejection probabilities for Clark and West's (2006, 2007) recursive OOS test statistic under the null hypothesis that the benchmark model is more accurate, $\mathrm{E}_T \bar{D}_T \leq 0$. Nominal size is 10% and is marked with a solid horizontal line. Values greater than 10% indicate that the test rejects the benchmark model too often.

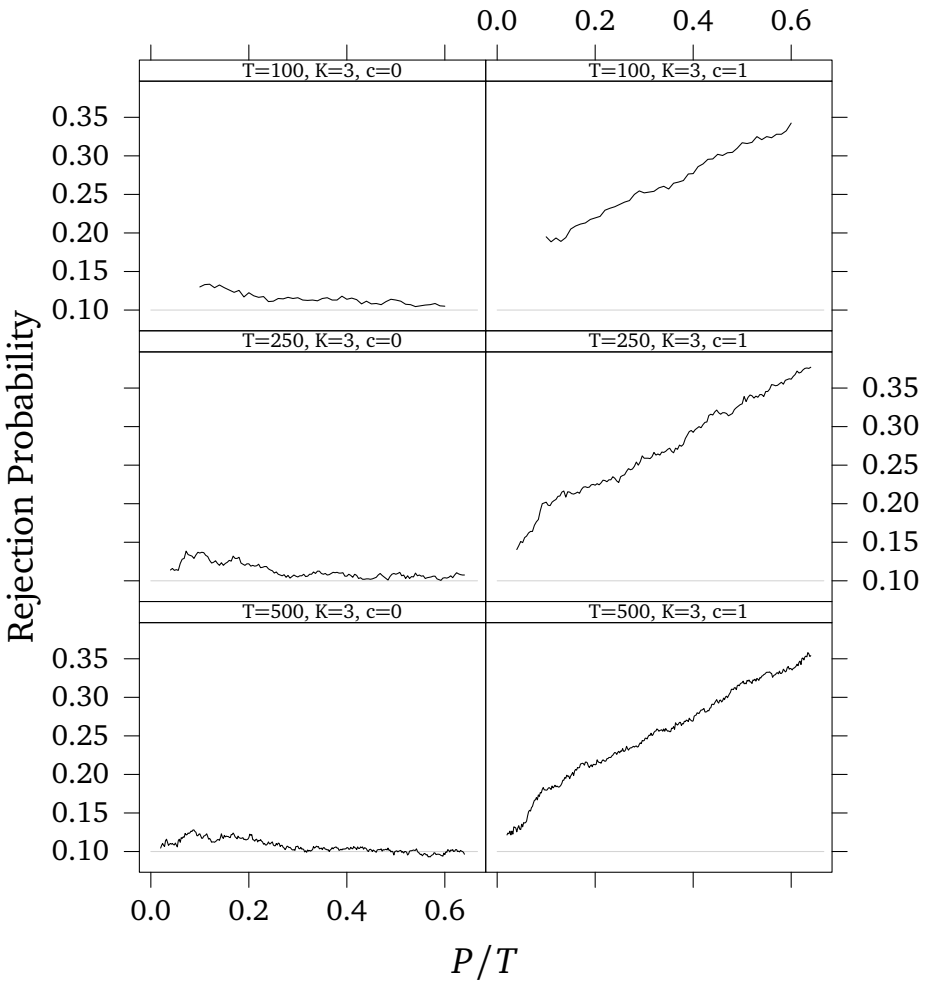

Figure 4: Simulated rejection probabilities for McCracken's (2007) recursive OOS-t test under the null hypothesis that the benchmark model is more accurate, $\mathrm{E}_T \bar{D}_T \leq 0$. Nominal size is 10% and is marked with a solid horizontal line. Values greater than 10% indicate that the test rejects the benchmark model too often.

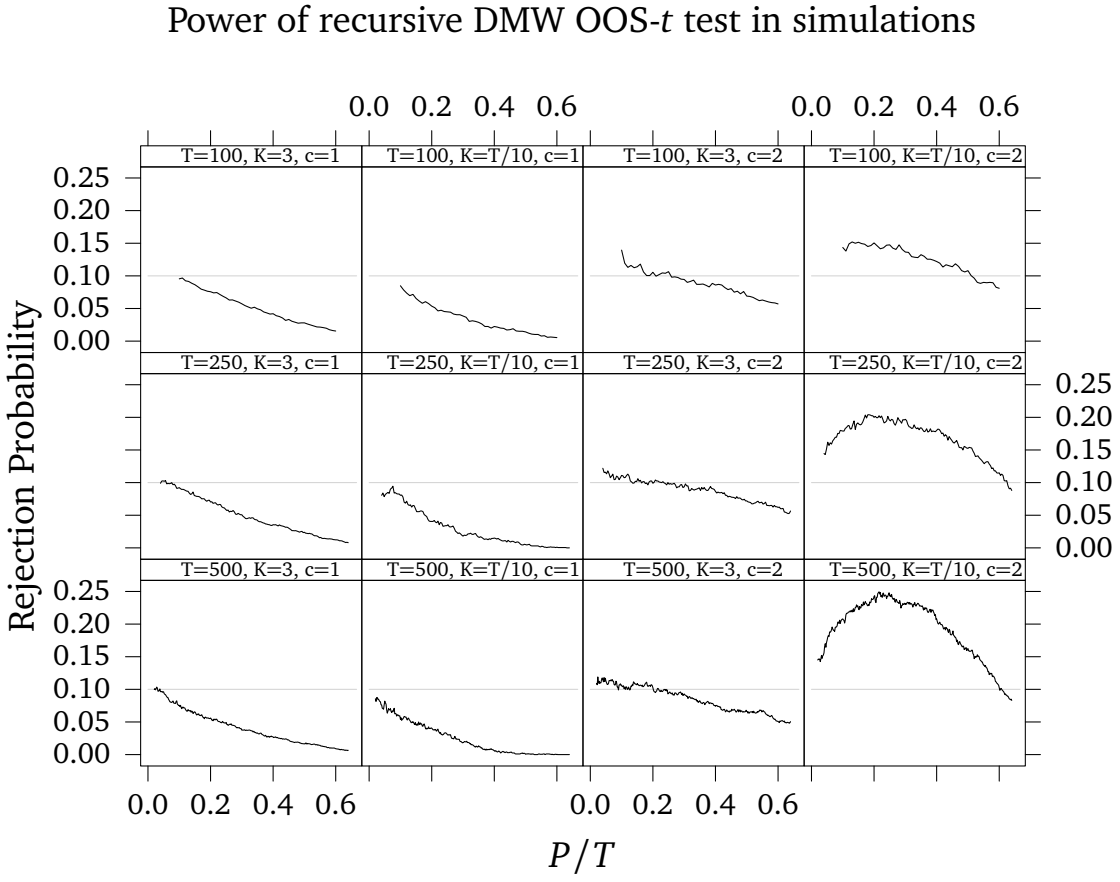

Figure 5: Simulated rejection probabilities for the recursive DMW OOS-t test under the alternative that the benchmark model is less accurate, $\mathrm{E}_T \bar{D}_T > 0$. Nominal size is 10% and is marked with a solid horizontal line. Values greater than 10% indicate that the test rejects the benchmark model too often.