Preliminary Power Plots

Curtis Miller

April 7, 2019

The plots below are power simulations comparing the ability of different test statistics to detect structural change in linear regression models. Here, $\beta = (1, 2)^T$, so the regression model is $y_t = 1 + 2x_t + \epsilon_t$, where $\{\epsilon_t\}_{t \in \mathbb{N}}$ is a mean-zero stationary process and $x_t \sim \mathrm{N}(1, 1)$.

The test statistics displayed include the Rényi-type statistic with trimming parameter $t_T = T^{1/2}$ computed with the regression model coefficients directly (———) the Rényi-type statistic with trimming parameter $t_T = T^{1/2}$ computed on the residuals (– – – –), the CUSUM statistic (— — –), and the Hidalgo-Seo statistic (-·-·-·-·-). The Rényi-type statistic and CUSUM statistics are the univariate statistics computed on the estimated residuals of the regression model. The Hidalgo-Seo statistic was tailored to the regression context as described in their paper.

These statistics do not all use the same LRV procedure. The Hidalgo-Seo statistic uses the periodogram matrix for LRV estimation. The univariate statistics use the procedure described in [1], which uses the full sample but allows for changes in the mean of the univariate data. The regression-coefficient-based Rényi-type statistic uses either the long-run variance estimation procedure described in [2] with kernel methods (in the case of the ARMA(2, 2) process) or a procedure that does not use kernel methods (for all other cases). The kernel used in the LRV estimation procedures for all statistics using kernel methods (the Rényi-type and CUSUM statistics) is the quadratic spectral kernel, and the bandwidth was selected using the procedure advocated by [3] and implemented in [4].

Two different change point regimes are implemented. The first is an early change point context. For a sample size T, the change point occurs near $T^{3/5}$ if there is a change at all. If it does, the new model the data follows is $y_t = (1 + \delta) + (2 - \delta)x_t + \epsilon_t$, where $\delta \in [-2, 2]$. The second regime has the change occur near $\frac{T}{10}$.

In this document, ϵ_t follows one of three models: independent and identically distributed $N(0, 1)$ random variables; an ARMA(2, 2) process with $\phi_1 = 0.4$, $\phi_2 = -0.03$, $\theta_1 = 0.5$, $\theta_2 = -0.06$, and the variance of the noise parameter is chosen so that the long-run variance is one; and a GARCH(1, 1) process with $\omega = 0.2$, $\alpha = 0.3$, and $\beta = 0.5$ (again the long-run variance is one). The proportional change point is investigated only when the residuals are *iid* Normal.

For every combination of sample size, test statistic, δ, and data-generating process for ϵ_t, 20,000 independent test statistics were simulated, and their asymptotically-appropriate p-values were computed. H_0 is rejected for p-values less than $\alpha = 0.05$; the value of α is signified with a dotted line (··········). Values of δ increment by 0.1, and the sample sizes considered are $T \in \{50, 250, 1000\}$.

The best-performing test statistic in all plots is the Hidalgo-Seo statistic, but the statistic also always suffers from considerable size inflation even for large sample sizes. The univariate statistics (the CUSUM statistic and the univariate Rényi-type statistic) do not demonstrate good power properties; sometimes they experience size inflation, they have bad power for large sample sizes, and in the case of the ARMA(2, 2) residuals they even have power curves with the wrong shape. The CUSUM statistic is the worst performing while the univariate Rényi-type statistic sometimes has decent power for large sample sizes. On balance, though, the regression-based Rényi-type often has excellent power without suffering from size inflation; in fact, for the ARMA(2, 2) residuals, the statistic (using kernel methods for LRV estimation) has a smaller size than specified.

References

[1] G. Rice, C. Miller, and L. Horváth, "A new class of change point test of rényi type." in-press.

[2] L. Horváth, C. Miller, and G. Rice, "Detecting early or late changes in linear models." draft.

[3] D. W. K. Andrews, "Heteroskedasticity and autocorrelation consistent covariance matrix estimation," *Econometrica*, vol. 59, pp. 817–858, May 1991.

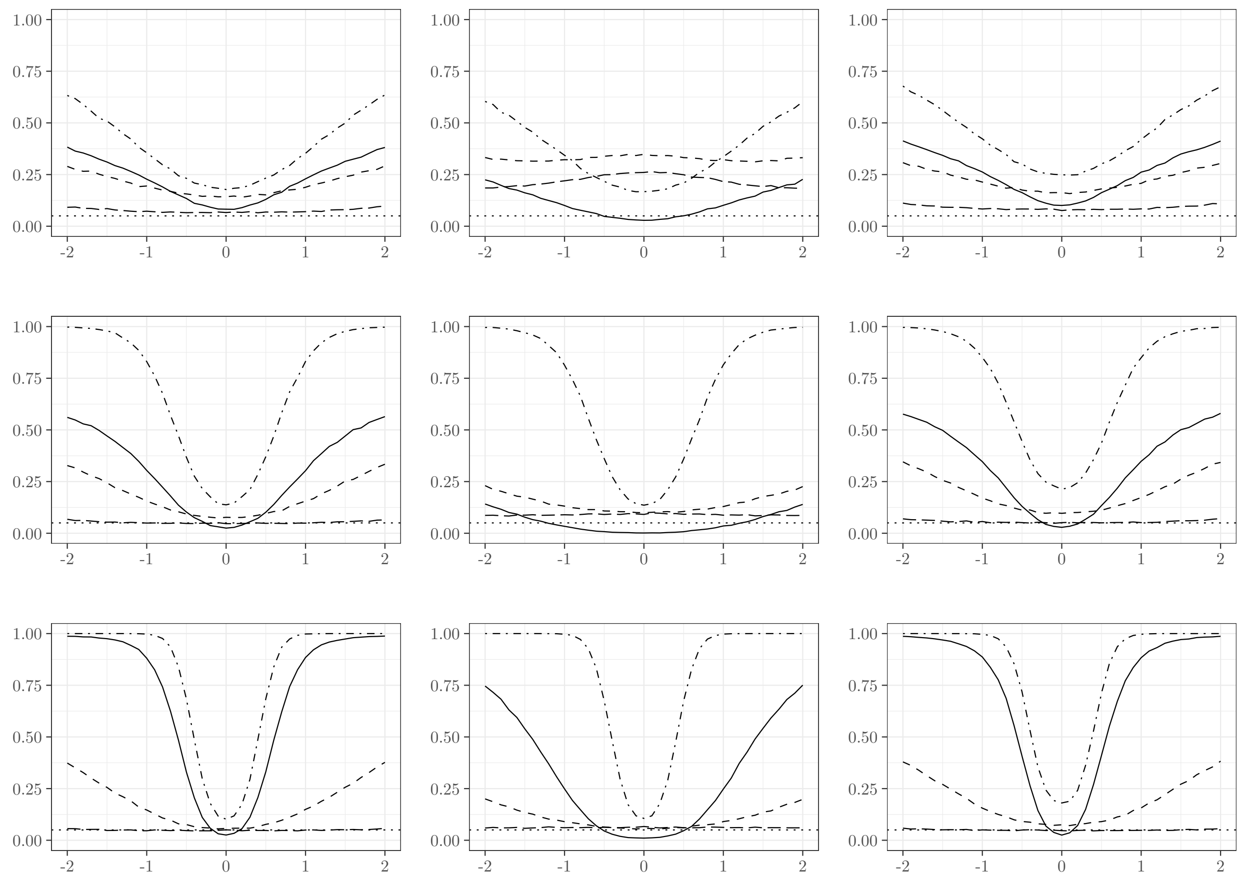

Figure 1: Power plots for when the change occurs near $T^{\frac{3}{5}}$; rows correspond to the sample sizes 50, 250, and 1000 respectively, the first column is for *iid* Normal residuals, the second column ARMA(2, 2) residuals, and the third column GARCH(1, 1) residuals. The x-axis tracks values of δ and the y-axis tracks observed power.

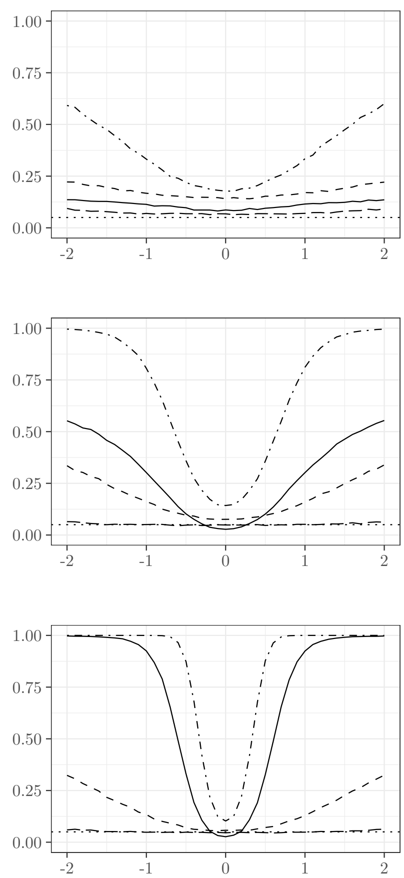

Figure 2: Power plots for when the change occurs near $\frac{T}{10}$; see Figure 1 for description.

[4] P. Aschersleben and M. Wagner, *cointReg: Parameter Estimation and Inference in a Cointegrating Regression*, 2016. R package version 0.2.0.