Online Supplement to "Asymptotically Uniform Tests After Consistent Model Selection in the Linear Regression Model"

Adam McCloskey*
University of Colorado at Boulder

February 2019

Details for Size Distortions when Using Properly Studentized Test Statistics: This online supplement proves the claim made in Section 3 of the main text of the paper that although a properly studentized test statistic does not have a discontinuity in its asymptotic distribution under fixed parameters, a test based upon such a statistic and standard normal critical values has incorrect asymptotic size.

For simplicity, we consider the special case described in the running example for which $v = 1$ and we have i.i.d. and conditionally homoskedastic data. We consider the one-sided version of the testing problem to simplify notation though completely analogous results hold for the two-sided problem. Consider the post-selection t-statistic for testing $H_0 : \theta = \theta_0$ against the one-sided alternative $H_a : \theta > \theta_0$:

$$T_n(\theta_0) = \frac{\sqrt{n}(\hat{\theta}_{n,1} - \theta_0)}{\hat{\sigma}_{n,1}} \mathbf{1}\left(|t_n| \leq c_n\right) + \frac{\sqrt{n}(\hat{\theta}_{n,f} - \theta_0)}{\hat{\sigma}_{n,f}} \mathbf{1}\left(|t_n| > c_n\right),$$

where $\hat{\sigma}^2_{n,1} = \hat{\sigma}^2_e(n^{-1}\sum_{i=1}^n x_i^2)^{-1}$ and $\hat{\sigma}^2_{n,f} = \hat{\sigma}^2_e(\hat{Q}_n^{-1})_{1,1}$ with $\hat{\sigma}^2_e = n^{-1}\sum_{i=1}^n (y_i - h_i'\hat{\alpha}_{n,f})^2$. Hence, we are examining the centered and scaled OLS estimator resulting from a pretest and are studentizing also according to the pretest. The statistic $T_n(\theta_0)$ does not exhibit a discontinuity in its asymptotic distribution as $T_n(\theta_0)$ converges to a standard normal random variable under H_0 for any *fixed* DGP (under the specified conditions). However, this convergence is not uniform in the DGP.

*Department of Economics, University of Colorado at Boulder, 256 UCB, Boulder, CO, 80309 (adam.mccloskey@colorado.edu)

Theorem. *Under H_0, (i)-(viii) in the restricted parameter space context of this section and $\{\gamma_{n,h}\}$, $T_n(\theta_0)$ converges in distribution to a random variable with distribution function*

$$\widetilde{J}_h(x) = \Phi(-g_2)\Phi(g_1)\Phi\left(x + \frac{\widetilde{h}_1}{(1-\rho^2_{h_{2,2}})^{1/2}}\right) + [1-\Phi(g_2)\Phi(-g_1)]\Phi(x)$$
$$+ \int_{-\infty}^{x} \phi(u)\left[\Phi\left(\frac{g_2+\rho_{h_{2,2}}u}{(1-\rho^2_{h_{2,2}})^{1/2}}\right) - \Phi\left(\frac{g_1+\rho_{h_{2,2}}u}{(1-\rho^2_{h_{2,2}})^{1/2}}\right)\right]du$$

where $\widetilde{h}_1 \equiv -h_{1,1}/[(\Omega^{1/2}_{h_{2,1}}Q^{-1/2}_{h_{2,2}})_{1,1}(Q^{-1/2}_{h_{2,2}})_{1,1}]$, $\rho_{h_{2,2}} \equiv -(Q_{h_{2,2}})_{1,2}/\sqrt{(Q_{h_{2,2}})_{1,1}(Q_{h_{2,2}})_{2,2}}$, $\phi(\cdot)$ *denotes the density function and* $\Phi(\cdot)$ *denotes the distribution function of a standard normal random variable defined over the extended real line.*

Proof: To simplify notation, let $\sigma_{n,\xi} = (Q^{-1}_{\gamma_{n,h,2,2}}\Omega^{1/2}_{\gamma_{n,h,2,1}})_{2,2}$ and $\sigma_{\theta,h_2} = (Q^{-1}_{h_{2,2}}\Omega^{1/2}_{h_{2,1}})_{2,2}$. The proof makes heavy use of Proposition A.2 of Leeb and Pötscher (2005) and its proof. For brevity, we will subsequently refer to this proposition simply as "Proposition A.2". To use these results, we divide the proof into four cases: (A) $\lim_{n\to\infty}\sqrt{n}\xi_n/(\sigma_{n,\xi}c_n) \in (-1,1)$, (B) $\lim_{n\to\infty}\sqrt{n}\xi_n/(\sigma_{n,\xi}c_n) \in \mathbb{R}_\infty \setminus [-1,1]$, (C) $\lim_{n\to\infty}\sqrt{n}\xi_n/(\sigma_{n,\xi}c_n) = 1$, (D) $\lim_{n\to\infty}\sqrt{n}\xi_n/(\sigma_{n,\xi}c_n) = -1$.

Case (A). In this case, $g_1 = -g_2 = \infty$ so that $\widetilde{J}_h(x) = \Phi(x + \widetilde{h}_1/(1-\rho^2_{h_{2,2}})^{1/2})$, which is the limiting distribution function of $T_n(\theta_0)$ according to 1.(i) of Proposition A.2 given that $\hat{\sigma}^2_{n,1} \xrightarrow{p} \sigma^2_{\theta,h_2}(1-\rho^2_{h_{2,2}})$.

Case (B). In this case, 2.(i) of Proposition A.2 implies that the limiting distribution function of $T_n(\theta_0)$ is $\Phi(x)$. We now examine three subcases: (I) $\lim_{n\to\infty}\sqrt{n}\xi_n/(\sigma_{n,\xi}c_n) = -\infty$, (II) $\lim_{n\to\infty}\sqrt{n}\xi_n/(\sigma_{n,\xi}c_n) \in (1,\infty]$, (III) $\lim_{n\to\infty}\sqrt{n}\xi_n/(\sigma_{n,\xi}c_n) \in (-\infty,-1)$. In subcase (I), $g_1 = -g_2 = -\infty$ so that $\widetilde{J}_h(x) = \Phi(x)$. In subcase (II), $g_1 = g_2 = \infty$ so that $\widetilde{J}_h(x) = \Phi(x)$. In subcase (III), $g_1 = g_2 = -\infty$ so that $\widetilde{J}_h(x) = \Phi(x)$.

Case (C). In this case $g_1 = \infty$. We now examine three subcases: (I) $g_2 = -\infty$, (II) $g_2 = \infty$, (III) $g_2 \in \mathbb{R}$. In subcase (I), $\widetilde{J}_h(x) = \Phi(x+\widetilde{h}_1/(1-\rho^2_{h_{2,2}})^{1/2})$, which is the limiting distribution of function $T_n(\theta_0)$ according to 1.(ii) of Proposition A.2 given that $\hat{\sigma}^2_{n,1} \xrightarrow{p} \sigma^2_{\theta,h_2}(1-\rho^2_{h_{2,2}})$. In subcase (II), $\widetilde{J}_h(x) = \Phi(x)$, which is the limiting distribution function of $T_n(\theta_0)$ according to 2.(ii) of Proposition A.2. In subcase (III),

$$\widetilde{J}_h(x) = \Phi(-g_2)\Phi\left(x + \frac{\widetilde{h}_1}{(1-\rho^2_{h_{2,2}})^{1/2}}\right) + \int_{-\infty}^{x} \phi(u)\Phi\left(\frac{g_2+\rho_{h_{2,2}}u}{(1-h^2_{2,2)})^{1/2}}\right)du,$$

which is the limiting distribution function of $T_n(\theta_0)$ implied by the proof of 3. of Proposition A.2.

Case (D). The proof for this case is very similar to that for Case (C), relying on the fact that $g_2 = -\infty$, 1.(iii) and 2.(iii) of Proposition A.2 and the proof of 4. of Proposition A.2. ■

The above theorem directly provides us with the least-favorable parameter sequences: those for which the null rejection probability of the t-test is highest. This will correspond to the h such that $\widetilde{J}_h(x) = 0$ for all $x \in \mathbb{R}$, i.e., $T_n(\theta_0)$ diverges under the corresponding drifting sequences. These sequences are characterized by $\widetilde{h}_1 = -\infty$ and $g_1 = -g_2 = \infty$. Since $T_n(\theta_0)$ diverges under these sequences while the usual critical values, quantiles of a standard normal distribution, remain fixed in the sample size, the asymptotic size of the usual test is equal to one.

References

Leeb, H., Pötscher, B. M., 2005. Model selection and inference: facts and fiction. Econometric Theory 21, 21–59.