Supplement to "A nondegenerate Vuong test"

Xiaoxia Shi
University of Wisconsin–Madison

Appendix C: Additional simulation example

In this section, I present the third simulation example. This example—borrowed from Schennach and Wilhelm (2011)—proves to be a very clever design that allows us to neatly demonstrate the power against both the $n^{-1/2}$-local alternatives defined in Assumption 5.1 and the n^{-1}-local alternatives defined in Assumption 5.2.

Example 3 (Normal Mean and Variance). Let the two models compared be

$$\mathcal{F} : \big\{ N(\theta, 1) : \theta \in \Theta \subset R \big\},$$
$$\mathcal{G} : \big\{ N(0, \beta) : \beta \in \mathcal{B} \subset (0, \infty) \big\}.$$

Let Y be generated from $\sim N(\mu, v^2)$, where $\mu = \sqrt{e^{2 \cdot lr - 1 + v^2} - v^2}$, where $lr \in \{x \in R : e^{2 \cdot lr - 1 + v^2} - v^2 \geq 0\}$. Under DGPs of this form, $E[\Lambda_i(\phi^*)] = lr$. Thus, varying lr controls how far the deviation is from H_0. On the other hand, when $lr = 0$, varying the parameter v^2 controls how large ω^2 is. Setting $v^2 = 1$ makes $\omega^2 = 0$, and setting v^2 far from 1 makes ω^2 large.

First, I fix $lr = 0$ and study the null rejection probabilities of the different tests. The simulation results are reported in the top three subplots of Figure 4. The figure shows that my nondegenerate test has remarkable size control at all three sample sizes. On the other hand, the one-step and the two-step Vuong tests, as well as the SW tests have large size distortion at $n = 100$, and still some noticeable size distortion at $n = 250$.

Second, I fix $v^2 = 5$ and study the power of the different tests as lr varies from 0 to $1.6\sqrt{250/n}$. The $n^{-1/2}$-local power is considered because $v^2 = 5$ represents the nondegenerate case $\omega_{P_0}^2 > 0$, and local alternatives around this null DGP should have $\omega_{P_n}^2 \nrightarrow 0$. The results are reported in the middle three subplots of Figure 4. The plots show that the power figures of all four tests stay constant as the sample size increases with $\sqrt{n}lr$ kept constant. My nondegenerate test has power similar to that of the one-step and the two-step Vuong tests, and higher than that of the SW test. The power disadvantage of the SW test perhaps is due to the loss of efficiency from the sample splitting.
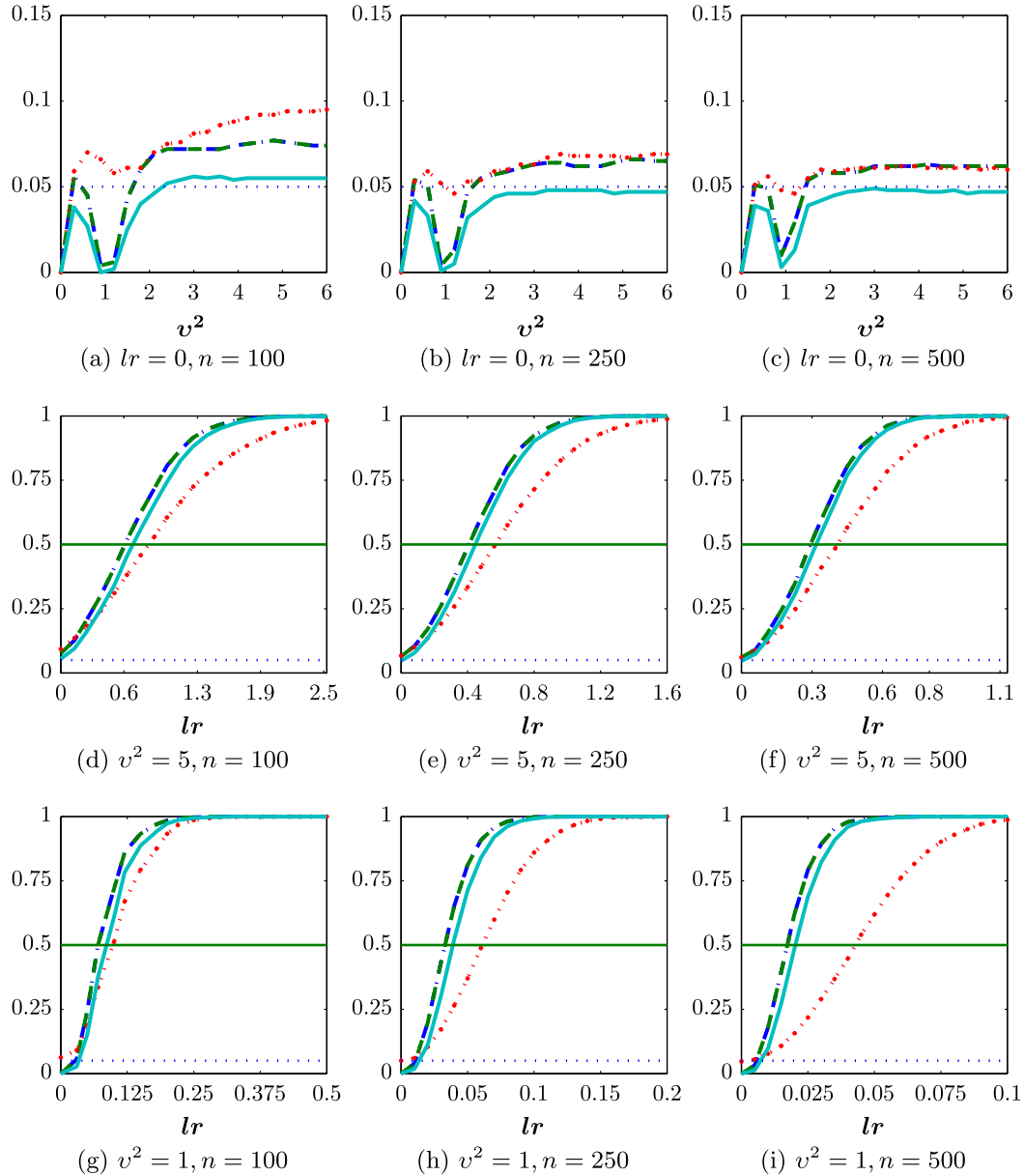
Xiaoxia Shi: xshi@ssc.wisc.edu

FIGURE 4. Rejection probabilities for the one-step Vuong test (dash-dotted line), two-step Vuong test (dashed line), the SW test (dotted line) and my new nondegenerate test (solid line) for Example 3. The horizontal dotted line indicates the nominal level 5%.

Last, I fix $v^2 = 1$ and study the power of the four different tests as lr varies from 0 to $0.2 \cdot (250/n)$. The n^{-1}-local power is considered because $v^2 = 1$ represents the degenerate case: $\omega_{P_0}^2 = 0$. The results are reported in the middle three subplots of Figure 4. The plots show that the power of my nondegenerate test and that of the classical Vuong tests are similar, and all stay approximately constant as n increases with $n \times lr$ kept constant.

On the other hand, the SW test appears to have lower power and its power decreases as n increases with $n \times lr$ kept constant. Thus, the power gap between my new nondegenerate test and the SW test increases as the sample size increases.

APPENDIX D: SCHENNACH AND WILHELM (2011) TEST

Here I briefly describe Schennach and Wilhelm's (2011) split sample test (SW test) in my notation. To construct the SW test, first, split the full sample $\{X_i\}_{i=1}^{n}$ into two equal-sized samples $\{X_{(1)i}\}_{i=1}^{n/2}$ and $\{X_{(2)i}\}_{i=1}^{n/2}$ (for example, split into two halves according to the natural ordering); second, let the split-sample log-likelihood ratio estimator be

$$\widehat{\mathrm{LR}}_n^{\mathrm{splt}} = \frac{2}{n(2+\varepsilon_n)} \sum_{i=1}^{n/2} \left[\left(\log f(X_{(1)i}, \hat{\theta}_n) - \log g(X_{(2)i}, \hat{\beta}_n) \right) \right.$$
$$\left. + (1+\varepsilon_n)\left(\log f(X_{(2)i}, \hat{\theta}_n) - \log g(X_{(1)i}, \hat{\beta}_n) \right) \right], \tag{D.1}$$

where $\varepsilon_n \in R \setminus \{0, -2\}$ is a user-chosen weighting. Let the variance estimator be

$$\left(\hat{\omega}_n^{\mathrm{splt}} \right)^2 = \frac{2}{n(2+\varepsilon_n)^2} \sum_{i=1}^{n/2} \left[\left(\log f(X_{(1)i}, \hat{\theta}_n) - \log g(X_{(2)i}, \hat{\beta}_n) \right) \right.$$
$$\left. + (1+\varepsilon_n)\left(\log f(X_{(2)i}, \hat{\theta}_n) - \log g(X_{(1)i}, \hat{\beta}_n) \right) \right]^2 - \left(\widehat{\mathrm{LR}}_n^{\mathrm{splt}} \right)^2; \tag{D.2}$$

third, let

$$\hat{T}_n^{\mathrm{splt}} = (n/2)^{1/2} \widehat{\mathrm{LR}}_n^{\mathrm{splt}} / \hat{\omega}_n^{\mathrm{splt}}. \tag{D.3}$$

Finally, reject H_0 if $|\hat{T}_n^{\mathrm{splt}}| > z_{\alpha/2}$. When H_0 is rejected, pick model $\mathcal{F}$ if $\widehat{\mathrm{LR}}_n^{\mathrm{splt}} > 0$ and pick model $\mathcal{G}$ if $\widehat{\mathrm{LR}}_n^{\mathrm{splt}} < 0$.[1]

In the 2011 version of their paper, they suggest a robust choice for the weighting parameter ε_n,

$$\varepsilon_n^* = \max\left\{ \frac{\mathrm{Cov}_n(\log f_i(\hat{\theta}_n), \log g_i(\hat{\beta}_n))}{\mathrm{Var}_n(\log f_i(\hat{\theta}_n)) + \mathrm{Var}_n(\log g_i(\hat{\beta}_n))}, 0 \right\} - 1, \tag{D.4}$$

where Cov_n stands for sample covariance and Var_n stands for sample variance.

REFERENCE

Schennach, S. M. and D. Wilhelm (2011), "A simple parametric model selection test." Unpublished manuscript, University of Chicago. [1, 3]

---

[1]Schennach and Wilhelm (2011) write their test as a Wald test from a GMM problem formed by the maximum likelihood estimator (MLE) f.o.c.s and the null hypothesis of the Vuong test. Some algebra shows that their GMM estimators of θ and β are exactly the MLEs because they have to satisfy the MLE f.o.c.s, and their regularized Wald statistic is exactly $(\hat{T}_n^{\mathrm{splt}})^2$.