Size Distortion and Modification of Classical Vuong Tests

Xiaoxia Shi

University of Wisconsin at Madison

March 2011

Vuong Test (Vuong, 1989)

- Data $\{X_i\}_{i=1}^n$.
- Two competing parametric models:

$$f(x,\theta),\ \theta \in \Theta \quad \text{vs.} \quad g(x,\beta), \beta \in B.$$

- Evaluate the relative fit:

$$H_0 : LR \equiv \max_{\theta \in \Theta} E\left[\log f(X_i,\theta)\right] - \max_{\beta \in B} E\left[\log g(X_i,\beta)\right] = 0$$

- Likelihood ratio statistic:

$$LR_n = n^{-1} \sum_{i=1}^{n} \left[\log f\left(X_i, \hat{\theta}_n\right) - \log g\left(X_i, \hat{\beta}_n\right)\right].$$

Vuong Test (Vuong, 1989)

- If the two models are nonnested, under H_0:

$$\sqrt{n} LR_n \to_p N\left(0, \omega^2\right)$$

where $\omega^2 = E\left[\log f\left(X_i, \theta_*\right) - \log g\left(X_i, \beta_*\right)\right]^2$.

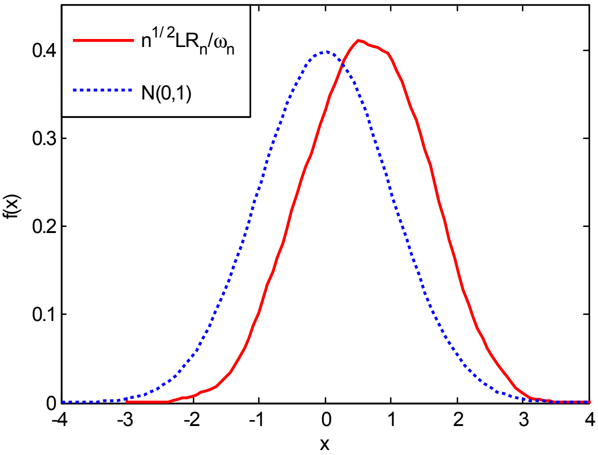
- One-Step Test: ($\hat{\omega}_n^2$: sample version of ω^2)

$$\text{Reject } H_0 \text{ if } \left|\frac{\sqrt{n} LR_n}{\hat{\omega}_n}\right| > z_{\alpha/2}.$$

- Two-step Test: reject H_0 if

$$n\hat{\omega}_n^2 > c_n\left(1-\alpha\right) \text{ and } \left|\frac{\sqrt{n} LR_n}{\hat{\omega}_n}\right| > z_{\alpha/2}.$$

Approximation Quality of Normal (n=1000)



About the Graph

- From the comparison of two normal regression models with 10 and 2 regressors respectively.
- Data generated under H_0.
- $\omega^2 > 0$, and the variance test $n\hat{\omega}_n^2$ rejects almost all the time.
- Rejection probability of a 5% test: 7.3%.

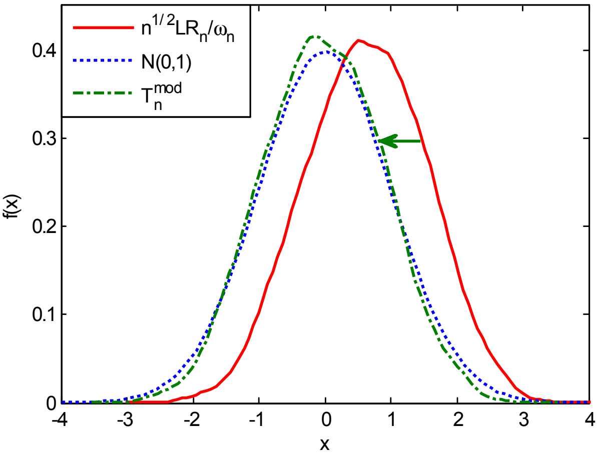
About the Graph

- From the comparison of two normal regression models with 10 and 2 regressors respectively.
- Data generated under H_0.
- $\omega^2 > 0$, and the variance test $n\hat{\omega}_n^2$ rejects almost all the time.
- Rejection probability of a 5% test: 7.3%.
- AIC, BIC corrections mentioned in Vuong (1989), but they do not move the red curve to the right place.

About the Graph

- From the comparison of two normal regression models with 10 and 2 regressors respectively.
- Data generated under H_0.
- $\omega^2 > 0$, and the variance test $n\hat{\omega}_n^2$ rejects almost all the time.
- Rejection probability of a 5% test: 7.3%.
- AIC, BIC corrections mentioned in Vuong (1989), but they do not move the red curve to the right place.
- I propose a new correction.

Approximation Quality of Normal (n=1000)



Outline

- Bias in LR_n
- Over-rejection of the Vuong tests
- Modified Test
- Examples
- Extensions to GMM Models

Bias in LRn

$$
\begin{aligned}
\sqrt{n} LR_n &= n^{-1/2} \sum_{i=1}^{n} \left[\log f\left(X_i, \hat{\theta}_n\right) - \log g\left(X_i, \hat{\beta}_n\right) \right] \\
&= n^{-1/2} \sum_{i=1}^{n} \left[\log f\left(X_i, \theta_*\right) - \log g\left(X_i, \beta_*\right) \right] - \\
&\quad \frac{1}{2\sqrt{n}} \cdot \sqrt{n}\left(\hat{\phi}_n - \phi_*\right)' A \sqrt{n}\left(\hat{\phi}_n - \phi_*\right) + o_p\left(n^{-1}\right) \\
&\equiv LR1_n - n^{-1/2} LR2_n + o_p\left(n^{-1}\right).
\end{aligned}
$$

Under H_0, $E\left[LR1_n\right] = 0$, but $E\left[LR2_n\right] \neq 0$

Bias in LRn

- $-n^{-1/2}E\left[LR2_n\right]$ is the higher-order bias in $\sqrt{n}LR_n$.
- How influential is the higher-order bias?

Bias in LRn

- $-n^{-1/2}E\left[LR2_n\right]$ is the higher-order bias in $\sqrt{n}LR_n$.
- How influential is the higher-order bias?
  - It depends on the relative magnitude of $LR1_n$ and $-n^{-1/2}LR2_n$

Bias in LRn

- $-n^{-1/2}E\left[LR2_n\right]$ is the higher-order bias in $\sqrt{n}LR_n$.
- How influential is the higher-order bias?
  - It depends on the relative magnitude of $LR1_n$ and $-n^{-1/2}LR2_n$
  - $-n^{-1/2}LR2_n$ is important if

Bias in LRn

- $-n^{-1/2}E\left[LR2_n\right]$ is the higher-order bias in $\sqrt{n}LR_n$.
- How influential is the higher-order bias?
  - It depends on the relative magnitude of $LR1_n$ and $-n^{-1/2}LR2_n$
  - $-n^{-1/2}LR2_n$ is important if
    - $n\omega^2$ small

Bias in LRn

- $-n^{-1/2}E\left[LR2_n\right]$ is the higher-order bias in $\sqrt{n}LR_n$.
- How influential is the higher-order bias?
  - It depends on the relative magnitude of $LR1_n$ and $-n^{-1/2}LR2_n$
  - $-n^{-1/2}LR2_n$ is important if
    - $n\omega^2$ small
    - $\left|E\left[LR2_n\right]\right|$ large.

Bias in LRn - Asymptotic Form of E[LR2n]

Let $\Lambda_i(\phi) = \log f(X_i, \theta) - \log g(X_i, \beta)$; $\phi = (\theta', \beta')'$.

$$\sqrt{n}\left(\hat{\phi}_n - \phi_*\right) \rightarrow_d A^{-1} Z_\phi \equiv A^{-1} \cdot N(0, B),$$

where

$$A = E\left[\frac{\partial^2 \Lambda_i(\phi_*)}{\partial\phi\partial\phi'}\right], \quad B = E\left[\frac{\partial \Lambda_i(\phi_*)}{\partial\phi} \cdot \frac{\partial \Lambda_i(\phi_*)}{\partial\phi'}\right].$$

Lemma

Under standard conditions,

$$LR2_n \rightarrow_d \frac{Z_\phi' A^{-1} Z_\phi}{2}$$

Bias in LRn - Asymptotic Form of Bias

$$E\left[LR2_n\right] = \frac{trace\left(A^{-1}B\right)}{2}$$

$$= \frac{trace\left(A_1^{-1}B_1\right) - trace\left(A_2^{-1}B_2\right)}{2},$$

where A_j and B_j are respectively the Hessian and the outer-product versions of the information matrix of model j.

- **Special case**: under mild or no misspecification: bias$=\left(d_\theta - d_\beta\right)/2$.
- It can be quite large (relative to $n\omega^2$), and it favors the model with more parameters.
- AIC and BIC correct too much and result in an opposite bias.

Outline

- Bias in LR_n
- Over-rejection of the Vuong tests
- Modified Test
- Examples
- Extensions to GMM Models

Over-rejection of the Vuong Tests

- (Mainly) due to the bias in LR_n, the Vuong tests can over-reject the null.
- The over-rejection can be arbitrarily large (close to $1-\alpha$) – far worse than illustrated in previous graph.
- The over-rejection can be captured asymptotically by considering a drifting sequence of null DGPs $\{P_n\}$
- $n\omega^2_{P_n} \to \sigma^2 \in [0,\infty]$, $A_{P_n} \to A$, $B_{P_n} \to B$, and

$$\rho^*_{P_n} = E_{P_n}\left[\Lambda_i(\phi_*) \cdot \frac{\partial \Lambda_i(\phi_*)}{\partial \phi}\right] \to \rho^*$$

Over-rejection of the Vuong Tests

Lemma

Under $\{P_n\}$ and standard MLE conditions

$$\begin{pmatrix} nLR_n \\ n\hat{\omega}_n^2 \end{pmatrix} \to_d \begin{pmatrix} \sigma Z_0 - 2^{-1} Z_1' V Z_1 \\ \sigma^2 - 2\sigma\rho V Z_1 + Z_1' V^2 Z_1 \end{pmatrix}.$$

where $[Q, V] = eig\left(A^{-1}B\right)$, $(Z_0, Z_1) \sim N\left(0, [1, \rho'; \rho, I]\right)$ *and* $\rho = Q' \left[\Omega^{1/2}\right]^{+} \rho^{*}$.

- $\sqrt{n}LR_n/\hat{\omega}_n$ is close to $N(0,1)$ if σ is large relative to $trace\,(V)$
- the bias dominates if $trace\,(V)$ is large relative to σ

Over-rejection of the Vuong Tests

Theorem

Under $\{P_{n,k}\}_{n,k=1}^{\infty}$ *such that* H_0 *holds and*
(i) for all k, $\left(n\omega^2_{P_{n,k}}, A_{P_{n,k}}, B_{P_{n,k}}, \rho_{P_{n,k}}\right) \rightarrow \left(\sigma^2_k, A_k, B, \rho\right)$
(ii) $\frac{-tr(V_k)}{\sigma_k} \rightarrow \infty$, $\frac{-tr(V_k)}{\sqrt{tr\left(V_k^2\right)}} \rightarrow \infty$, *and* $\frac{tr\left(V_k^4\right)}{\left[tr\left(V_k^2\right)\right]^2} \rightarrow 0$
then

$$\lim_{k\rightarrow\infty}\lim_{n\rightarrow\infty} \Pr\left(\frac{\sqrt{n}LR_n}{\hat{\omega}_n} > z_{\alpha/2}\right) = 1.$$

If in addition, $\frac{\sigma_k^2}{tr\left(V_k^2\right)} \rightarrow \infty$, *then we also have*

$$\lim_{k\rightarrow\infty}\lim_{n\rightarrow\infty} \Pr\left(n\hat{\omega}_n > c_n\left(1-\alpha\right) \ \& \ \frac{\sqrt{n}LR_n}{\hat{\omega}_n} > z_{\alpha/2}\right) = 1$$

Over-rejection of the Vuong Tests

- Implications of the Theorem:
  - by increasing the number of parameters of one model, one can always make the Vuong tests pick this model, even if this model is no better than the other.
  - "no better than" can be replaced with "worse".
- What about AIC and BIC corrections (suggested by various authors)?
  - correct too much
  - By increasing the number of parameters of one model, one can always make the Vuong tests **reject** this model, even if this model is no worse than the other
  - OK if objective is forecasting; not OK if want to take Vuong tests as hypothesis tests seriously.

Outline

- Bias in LR_n
- Over-rejection of the Vuong tests
- Modified Test
- Examples
- Extensions to GMM Models

Modified Test

- Modification contains three parts:
  - modified LR_n: $LR_n^{\mathrm{mod}} = LR_n + tr\left(\hat{V}_n\right)/(2n)$,
  - modified $\hat{\omega}_n^2$: $\left(\hat{\omega}_n^{\mathrm{mod}}\right)^2 = \hat{\omega}_n^2 + n^{-1} tr\left(\hat{V}_n^4\right)/tr\left(\hat{V}_n^2\right)$,
  - modified critical value (discussed later): $z_{\alpha/2}^{\mathrm{mod}}$.
- Modification to LR_n removes most of the over-rejection,
- But $tr\left(\hat{V}_n\right)/(2n)$ introduces slight new over-rejection when $\hat{V}_n$ has one dominating element – solved by the modification of $\hat{\omega}_n^2$,
- $\sqrt{n} LR_n^{\mathrm{mod}}/\hat{\omega}_n^{\mathrm{mod}}$ has little bias and is close to $N(0,1)$, but still not exactly $N(0,1)$ – fortunately we know what it is (asymptotically).

Asymptotic Distribution of Modified Statistic

Lemma

Under $\{P_n\}$ and standard MLE conditions

$$\frac{n^{1/2} LR_n^{\mathrm{mod}}}{\hat{\omega}_n^{\mathrm{mod}}} \rightarrow_d J_{\sigma,\rho,V}$$
$$= \frac{\sigma Z_0 - 2^{-1}\left(Z_1' V Z_1 - tr\left(V\right)\right)}{\sqrt{\sigma^2 - 2\sigma\rho V Z_1 + Z_1' V^2 Z_1 + tr\left(V^4\right)/tr\left(V^2\right)}}.$$

- Modified critical value:

$$z_{\alpha/2}^{\mathrm{mod}} = \sup_{\sigma \in [0,\infty)} \mathrm{Quantile}(|J_{\sigma,\hat{\rho}_n,\hat{V}_n}|, 1-\alpha).$$

  - where $\hat{\rho}_n$, $\hat{V}_n$ are consistent estimators of ρ, V,
  - σ cannot be consistently estimated.

Modified Test

- Modified Test: reject H_0 if $T_n^{\text{mod}} \equiv \left| \frac{n^{1/2} LR_n^{\text{mod}}}{\hat{\omega}_n^{\text{mod}}} \right| > z_{\alpha/2}^{\text{mod}}$.

Theorem

For a set of null DGPs $\mathcal{H}_0$, suppose the standard MLE conditions hold uniformly over the set, then

$$\limsup_{n\to\infty} \sup_{P\in\mathcal{H}_0} \Pr_P \left(\left| \frac{n^{1/2} LR_n^{\text{mod}}}{\hat{\omega}_n^{\text{mod}}} \right| > z_{\alpha/2}^{\text{mod}} \right) \leq \alpha.$$

- In words: the asymptotic size of the modified test is less than or equal to α.
- In other words: the null rejection probability is uniformly well-controlled.

Discussion of the Critical Value

- $z^{\mathrm{mod}}_{\alpha/2}$ is in a sense a worst-case critical value.
- How conservative is it?
  - in the scenario when the classical Vuong tests over-rejection is the worst, $z^{\mathrm{mod}}_{\alpha/2} = z_{\alpha/2}$.
  - in other cases, $z^{\mathrm{mod}}_{\alpha/2}$ could be bigger, but not much bigger. For example $z^{\mathrm{mod}}_{0.05/2}$ is up to around $z_{0.01/2}$.
  - in the later cases, the modified test is much more powerful than the two-step Vuong test, and does not over-reject as the one-step Vuong test.
- How difficult is the computation?
  - fast (because only maximizing over a scalar)
  - convenient (because $\hat{\rho}_n$ and $\hat{V}_n$ can be easily obtained from the maximum likelihood routines).

Outline

- Bias in LR_n
- Over-rejection of the Vuong tests
- Modified Test
- **Examples**
- Extensions to GMM Models

Example 1 - Normal Regression

M1. $Y = \beta_0 + \sum_{j=1}^{d_1-1} \beta_j X_{1,j} + v$, $v \sim N(0, \sigma_2^2)$.

M2. $Y = \theta_0 + \sum_{j=1}^{d_2-1} \theta_j X_{2,j} + u$, $u \sim N(0, \sigma_1^2)$;

- DGP:

$$Y = 1 + \frac{a_1 \sum_{j=1}^{d_1-1} X_{1,j}}{\sqrt{d_1 - 1}} + \frac{a_2 \sum_{j=1}^{d_2-1} X_{2,j}}{\sqrt{d_2 - 1}} + \varepsilon$$

$$\left(X_{1,1}, ..., X_{1,d_1-1}, X_{2,1}, ..., X_{2,d_2-1}, \varepsilon\right) \sim N\left(0, I\right)$$

- Null: $a_1 = a_2 = 0.25$; Alterative: $a_1 = 0$, $a_2 = 0.25$

- Base case: $d_1 = 10$, $d_2 = 2$, $n = 250$.

Table 1. Rej. Prob. of Original and Modified Tests ($\alpha = 0.05$)

	Original Tests			Modified Test	
	2-Step	1-Step	Var. Test	Sel. Prob	Max c_n^{mod}
Null DGP					
Base	(.087,.004)	(.088,.004)	.949	(.015,.022)	2.00
$d_1 = 20$	(.205,.000)	(.283,.000)	.680	(.015,.014)	2.00
$d_1 = 5$	(.037,.010)	(.037,.010)	.990	(.018,.018)	2.04
$n = 500$	(.067,.005)	(.067,.005)	1	(.020,.019)	1.98
$n = 100$	(.051,.000)	(.136,.001)	.276	(.012,.013)	2.17
Alternative DGP (M2 true)					
Base	(.000,.032)	(.000,.032)	.625	(.000,.281)	2.00
$d_1 = 20$	(.001,.000)	(.001,.000)	.249	(.000,.187)	2.00
$d_1 = 5$	(.000,.204)	(.000,.204)	.830	(.000,.336)	2.10
$n = 500$	(.000,.315)	(.000,.315)	.971	(.000,.724)	2.00
$n = 100$	(.003,.001)	(.004,.001)	.109	(.001,.051)	2.10

Example 2 - Joint Normal Location Model

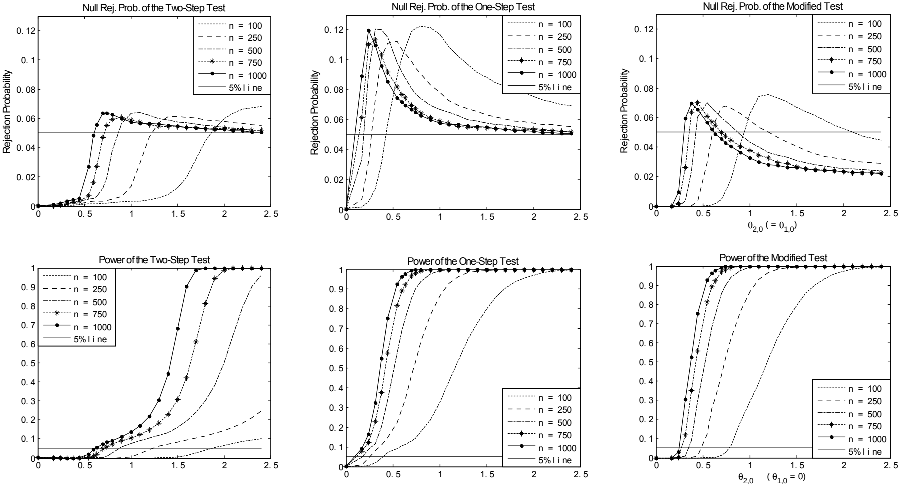
M1. $(Y_1, Y_2) \sim N((\theta_1, 0), I_2)$, $\theta_1 \in R$;

M2. $(Y_1, Y_2) \sim N((0, \theta_2), I_2)$, $\theta_2 \in R$.

- DGP:

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} \sim N\left(\begin{pmatrix} \theta_{1,0} \\ \theta_{2,0} \end{pmatrix}, \begin{pmatrix} 25 & 0 \\ 0 & 1 \end{pmatrix}\right)$$

- $LR = \theta_{1,0}^2 - \theta_{2,0}^2$.
- nominal size $\alpha = 0.05$.


Null Rej. Prob. of the Two-Step Test
Null Rej. Prob. of the One-Step Test
Null Rej. Prob. of the Modified Test
Rejection Probability
n = 100
n = 250
n = 500
n = 750
n = 1000
5% line
$\theta_{2,0}$ (= $\theta_{1,0}$)
Power of the Two-Step Test
Power of the One-Step Test
Power of the Modified Test
$\theta_{2,0}$ ($\theta_{1,0}$ = 0)

Outline

- Bias in LR_n
- Over-rejection of the Vuong tests
- Modified Test
- Examples
- **Extensions to GMM Models**

GMM Models and GEL Criteria

- GMM models (or moment condition models):

$$\begin{aligned} \text{M1} &: \quad Em_f\left(x, \psi_f\right) = 0 \text{ for some } \psi_f \in \Psi_f \subset R^{d_{\psi_f}}, \\ \text{M2} &: \quad Em_g\left(x, \psi_g\right) = 0 \text{ for some } \psi_g \in \Psi_g \subset R^{d_{\psi_g}}, \end{aligned} \quad (1)$$

where m_f and m_g are known moment functions and ψ_f and ψ_g are unknown parameters.

- Generalized Empirical Likelihood criteria: H_0:

$$\begin{aligned} GELR &\equiv \max_{\psi_f \in \Psi_f} \min_{\gamma_f} E\left[\kappa\left(\gamma_f' m_f\left(X_i, \psi_f\right)\right)\right] - \\ &\quad \max_{\psi_g \in \Psi_g} \min_{\gamma_g} E\left[\kappa\left(\gamma_g' m_g\left(X_i, \psi_g\right)\right)\right] \\ &= 0. \end{aligned}$$

EL: $\kappa(v) = -\log(1-v)$, ET (exponential tilting): $\kappa(v) = e^v$.

General Framework

- In previous analysis,
  - replace $\log f\left(x, \theta\right)$ and $\log g\left(x, \beta\right)$ with $\kappa\left(\gamma_f' m_f\left(X_i, \psi_f\right)\right)$ and $\kappa\left(\gamma_g' m_g\left(X_i, \psi_g\right)\right)$
  - replace θ_* and β_* with $\left(\gamma_{f,*}', \psi_{f,*}'\right)'$ and $\left(\gamma_{g,*}', \psi_{g,*}'\right)'$
  - then everything go through.

Summary

- Discover the higher-order bias in the Vuong test statistic
- Show that the bias cause (sometimes severe) over-rejection
- Propose a uniformly valid modified Vuong test
- Modified Vuong test is easy to compute and has good power.