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

- Likelihood ratio statistic:
  \[
  LR_n = n^{-1} \sum_{i=1}^{n} \left[ \log f(X_i, \hat{\theta}_n) - \log g(X_i, \hat{\beta}_n) \right].
  \]
Vuong Test (Vuong, 1989)

- If the two models are nonnested, under $H_0$:
  \[
  \sqrt{nLR_n} \overset{p}{\rightarrow} N(0, \omega^2)
  \]
  where $\omega^2 = E \left[ \log f(X_i, \theta^*_i) - \log g(X_i, \beta^*_i) \right]^2$.

- One-Step Test: ($\hat{\omega}_n^2$: sample version of $\omega^2$)
  Reject $H_0$ if
  \[
  \left| \frac{\sqrt{nLR_n}}{\hat{\omega}_n} \right| > z_{\alpha/2}.
  \]

- Two-step Test: reject $H_0$ if
  \[
  n\hat{\omega}_n^2 > c_n (1 - \alpha) \quad \text{and} \quad \left| \frac{\sqrt{nLR_n}}{\hat{\omega}_n} \right| > z_{\alpha/2}.
  \]
Approximation Quality of Normal (n=1000)

\[ n^{1/2} \frac{LR_n}{\omega_n} \]

\[ N(0,1) \]
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)

\[
\begin{align*}
&n^{1/2}LR_n/\omega_n \\
&N(0,1) \\
&T_{n}^{\text{mod}}
\end{align*}
\]
Bias in $LR_n$

Over-rejection of the Vuong tests

Modified Test

Examples

Extensions to GMM Models
Bias in LRn

\[ \sqrt{n}LR_n = n^{-1/2} \sum_{i=1}^{n} \left[ \log f (X_i, \hat{\theta}_n) - \log g (X_i, \hat{\beta}_n) \right] \]

\[ = n^{-1/2} \sum_{i=1}^{n} \left[ \log f (X_i, \theta_*) - \log g (X_i, \beta_*) \right] - \frac{1}{2\sqrt{n}} \cdot \sqrt{n} (\hat{\phi}_n - \phi_*)' A\sqrt{n} (\hat{\phi}_n - \phi_*) + o_p \left( n^{-1} \right) \]

\[ \equiv LR1_n - n^{-1/2} LR2_n + o_p \left( n^{-1} \right). \]

Under \( H_0, E[LR1_n] = 0 \), but \( E[LR2_n] \neq 0 \)
- $n^{-1/2} E[LR2_n]$ is the higher-order bias in $\sqrt{n}LR_n$.

- How influential is the higher-order bias?
- $-n^{-1/2} E [LR2_n]$ 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 $-n^{-1/2} LR2_n$. 

X. Shi (UW-Mdsn)
Bias in LRn

- $-n^{-1/2} E[LR2_n]$ 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 $\omega^2$ small and $E[LR2_n]$ large.
Bias in LRn

- \(-n^{-1/2} E[LR2_n]\) 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 $LR_n$

- $-n^{-1/2} E[LR_{2n}]$ is the higher-order bias in $\sqrt{n}LR_n$.

How influential is the higher-order bias?

- It depends on the relative magnitude of $LR_{1n}$ and $-n^{-1/2} LR_{2n}$
- $-n^{-1/2} LR_{2n}$ is important if
  - $n\omega^2$ small
  - $|E[LR_{2n}]|$ large.
Let \( \Lambda_i (\phi) = \log f (X_i, \theta) - \log g (X_i, \beta); \ \phi = (\theta', \beta')' \).

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

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

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

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

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 $\{P_n\}$

$$n\omega^2_{P_n} \rightarrow \sigma^2 \in [0, \infty], \ A_{P_n} \rightarrow A, \ B_{P_n} \rightarrow B, \text{ and}$$

$$\rho^*_P = E_{P_n} \left[ \Lambda_i (\phi_*) \cdot \frac{\partial \Lambda_i (\phi_*)}{\partial \phi} \right] \rightarrow \rho^*$$
Over-rejection of the Vuong Tests

**Lemma**

Under \( \{P_n\} \) and standard MLE conditions

\[
\left( \begin{array}{c} nLR_n \\ n\hat{\omega}_n^2 \end{array} \right) \xrightarrow{d} \left( \begin{array}{c} \sigma Z_0 - 2^{-1} Z'_1 V Z_1 \\ \sigma^2 - 2\sigma \rho V Z_1 + Z'_1 V^2 Z_1 \end{array} \right).
\]

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

- \(\sqrt{nLR_n/\hat{\omega}_n}\) is close to \(N (0, 1)\) if \(\sigma\) is large relative to \(\text{trace} (V)\)
- the bias dominates if \(\text{trace} (V)\) is large relative to \(\sigma\)
Over-rejection of the Vuong Tests

**Theorem**

Under \( \{ P_{n,k} \}_{n,k=1}^{\infty} \) such that \( H_0 \) holds and
(i) for all \( k \), \((n\omega_{P_{n,k}}^2, A_{P_{n,k}}, B_{P_{n,k}}, \rho_{P_{n,k}}) \rightarrow (\sigma_k^2, A_k, B, \rho) \)

(ii) \( \frac{-\text{tr}(V_k)}{\sigma_k} \rightarrow \infty \), \( \frac{-\text{tr}(V_k)}{\sqrt{\text{tr}(V_k^2)}} \rightarrow \infty \), and \( \frac{\text{tr}(V_k^4)}{[\text{tr}(V_k^2)]^2} \rightarrow 0 \)

then

\[
\lim_{k \to \infty} \lim_{n \to \infty} \text{Pr} \left( \frac{\sqrt{nLR_n}}{\hat{\omega}_n} > z_{\alpha/2} \right) = 1.
\]

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

\[
\lim_{k \to \infty} \lim_{n \to \infty} \text{Pr} \left( n\hat{\omega}_n > c_n (1 - \alpha) \ & \frac{\sqrt{nLR_n}}{\hat{\omega}_n} > z_{\alpha/2} \right) = 1
\]
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.
Bias in $LR_n$

Over-rejection of the Vuong tests

Modified Test

Examples

Extensions to GMM Models
Modified Test

- Modification contains three parts:
  - modified $LR_n$: $LR_n^{\text{mod}} = LR_n + tr \left( \hat{V}_n \right) / (2n)$,
  - modified $\hat{\omega}^2_n$: $(\hat{\omega}^{\text{mod}}_n)^2 = \hat{\omega}^2_n + n^{-1} tr (\hat{V}^4_n) / tr (\hat{V}^2_n)$,
  - modified critical value (discussed later): $z^{\text{mod}}_{\alpha/2}$.

- 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 $\hat{\omega}^2_n$,

- $\sqrt{n}LR_n^{\text{mod}} / \hat{\omega}^{\text{mod}}_n$ 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

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

- Modified critical value:

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

- where \( \hat{\rho}_n, \hat{V}_n \) are consistent estimators of \( \rho, V \),

- \( \sigma \) 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

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

- In words: the asymptotic size of the modified test is less than or equal to $\alpha$.
- In other words: the null rejection probability is uniformly well-controlled.
Discussion of the Critical Value

- $z_{\alpha/2}^{\text{mod}}$ 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, $z_{\alpha/2}^{\text{mod}} = z_{\alpha/2}$.
  - in other cases, $z_{\alpha/2}^{\text{mod}}$ could be bigger, but not much bigger. For example $z_{0.05/2}^{\text{mod}}$ 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).
Bias in $LR_n$

Over-rejection of the Vuong tests

Modified Test

Examples

Extensions to GMM Models
Example 1 - Normal Regression

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

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

- **DGP:**
  \[
  Y = 1 + \frac{a_1 \sum_{j=1}^{d_1-1} X_{1,j}}{\sqrt{d_1 - 1}} + \frac{a_2 \sum_{j=1}^{d_2-1} X_{2,j}}{\sqrt{d_2 - 1}} + \varepsilon
  \]
  
  \((X_{1,1}, \ldots, X_{1,d_1-1}, X_{2,1}, \ldots, X_{2,d_2-1}, \varepsilon) \sim N(0, I)\)

- **Null:** \( a_1 = a_2 = 0.25; \ **Alternative:** \( a_1 = 0, \ a_2 = 0.25 \)

- **Base case:** \( d_1 = 10, \ d_2 = 2, \ n = 250. \)
Table 1. Rej. Prob. of Original and Modified Tests ($\alpha = 0.05$)

<table>
<thead>
<tr>
<th></th>
<th>Original Tests</th>
<th>Modified Test</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2-Step</td>
<td>1-Step</td>
<td>Var. Test</td>
</tr>
<tr>
<td><strong>Null DGP</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base</td>
<td>(.087,.004)</td>
<td>(.088,.004)</td>
<td>.949</td>
</tr>
<tr>
<td>$d_1 = 20$</td>
<td>(.205,.000)</td>
<td>(.283,.000)</td>
<td>.680</td>
</tr>
<tr>
<td>$d_1 = 5$</td>
<td>(.037,.010)</td>
<td>(.037,.010)</td>
<td>.990</td>
</tr>
<tr>
<td>$n = 500$</td>
<td>(.067,.005)</td>
<td>(.067,.005)</td>
<td>1</td>
</tr>
<tr>
<td>$n = 100$</td>
<td>(.051,.000)</td>
<td>(.136,.001)</td>
<td>.276</td>
</tr>
<tr>
<td><strong>Alternative DGP (M2 true)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base</td>
<td>(.000,.032)</td>
<td>(.000,.032)</td>
<td>.625</td>
</tr>
<tr>
<td>$d_1 = 20$</td>
<td>(.001,.000)</td>
<td>(.001,.000)</td>
<td>.249</td>
</tr>
<tr>
<td>$d_1 = 5$</td>
<td>(.000,.204)</td>
<td>(.000,.204)</td>
<td>.830</td>
</tr>
<tr>
<td>$n = 500$</td>
<td>(.000,.315)</td>
<td>(.000,.315)</td>
<td>.971</td>
</tr>
<tr>
<td>$n = 100$</td>
<td>(.003,.001)</td>
<td>(.004,.001)</td>
<td>.109</td>
</tr>
</tbody>
</table>
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 \( \alpha = 0.05. \)
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):
  
  \( M1 : \quad Em_f (x, \psi_f) = 0 \) for some \( \psi_f \in \Psi_f \subset R^{d_{\psi_f}} \),
  
  \( M2 : \quad Em_g (x, \psi_g) = 0 \) for some \( \psi_g \in \Psi_g \subset R^{d_{\psi_g}} \),

  \[ \text{(1)} \]

  where \( m_f \) and \( m_g \) are known moment functions and \( \psi_f \) and \( \psi_g \) are unknown parameters.

- Generalized Empirical Likelihood criteria: \( H_0 : \)

  \[
  GELR \equiv \max_{\psi_f \in \Psi_f} \min_{\gamma_f} \min E \left[ \kappa \left( \gamma'_f m_f (X_i, \psi_f) \right) \right] - \\
  \max_{\psi_g \in \Psi_g} \min_{\gamma_g} \min E \left[ \kappa \left( \gamma'_g m_g (X_i, \psi_g) \right) \right] \\
  = 0.
  \]

  EL: \( \kappa (\nu) = - \log (1 - \nu) \), ET (exponential tilting): \( \kappa (\nu) = e^\nu \).
In previous analysis,

- replace \( \log f(x, \theta) \) and \( \log g(x, \beta) \) with \( \kappa \left( \gamma'_f m_f(X_i, \psi_f) \right) \) and \( \kappa \left( \gamma'_g m_g(X_i, \psi_g) \right) \)

- replace \( \theta_* \) and \( \beta_* \) 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.