THE EFFECT OF HETEROSCEDASTICITY ON THE ACTUAL SIZE OF THE CHOW TEST

Toshio Honda* and Akimichi Takemura**

The Chow test is a test of equality of two sets of regression coefficients in two regression models under the assumption of homoscedasticity. Toyoda (1974) studied the actual size of the Chow test under heteroscedasticity. In this paper, we reexamine the same problem using the method of Takemura and Honda (1994), and pay particular attention to cases of small heteroscedasticity.

1. Introduction

Consider two linear regression models with \( n_1 \) and \( n_2 \) observations

\[
\begin{align*}
(1.1) \quad y_1 &= X_1 \beta_1 + \varepsilon_1, \quad \varepsilon_1 \sim N_{n_1}(0, \sigma^2_1 I_{n_1}), \\
(1.2) \quad y_2 &= X_2 \beta_2 + \varepsilon_2, \quad \varepsilon_2 \sim N_{n_2}(0, \sigma^2_2 I_{n_2}),
\end{align*}
\]

where \( y_i \) are \( n_i \times 1 \) vectors of dependent variables, \( X_i \) are \( n_i \times k \) matrices of fixed regressors, \( \beta_i \) are \( k \times 1 \) vectors of unknown regression coefficients, \( \varepsilon_i \) are \( n_i \times 1 \) vectors of errors, \( i = 1, 2 \), and \( N_m(\mu, \Sigma) \) stands for the \( m \)-variate normal distribution with mean \( \mu \) and covariance matrix \( \Sigma \). The Chow test is a test of equality of regression coefficients \( \beta_1 \) and \( \beta_2 \) in the above linear regression models under the homoscedasticity assumption \( \sigma^2_1 = \sigma^2_2 \).

Even if the homoscedasticity assumption is slightly violated, researchers will still use the Chow test for testing the equality of regression coefficients. This is the case we will study here. Toyoda (1974) studied the actual size of the Chow test under heteroscedasticity and Schmidt and Sickles (1977) reexamined the numerical values of Toyoda (1974).

The present paper applies the results of Takemura and Honda (1994) to this problem. Based on the Taylor expansion, the method is extensively useful for theoretical examination of the effect of error covariance misspecification. We clarify the relationship between the values of regressors and the effect of heteroscedasticity on the actual size and show that the actual size will be smaller than the nominal one in some important cases. This means that the numerical examples of balanced cases given in Schmidt and Sickles (1977) are largely misleading. Similar ambiguous accounts on this issue are also found in some standard textbooks on econometrics, e.g. Johnston (1984) or Greene (1993).

In Section 3 we confirm our theoretical results by some simulation study. We add that other methods might be more suitable for numerical purposes, e.g. Imhof (1961).

If researchers are well aware of the presence of heteroscedasticity \( \sigma^2_1 \neq \sigma^2_2 \),

\begin{flushright}
Received May, 1995. Revised February, 1996, April, 1996. Accepted April, 1996.
* Institute of Social Sciences, University of Tsukuba, Ibaraki 305, Japan
** Faculty of Economics, University of Tokyo, Bunkyo-ku, Tokyo 113, Japan
\end{flushright}
then they can use other test statistics, e.g. the asymptotic Chow test by Goldfeld and Quandt (Goldfeld and Quandt (1978)) or Jayatissa’s exact test (Jayatissa (1977)). Tsurumi and Sheflin (1985) proposed a test based on the posterior mean of the ratio of standard deviations of error terms. An exact $F$-test for paired data cases is also given there, although their numerical results show that this $F$-test is not powerful. See Thursby (1992) for other test statistics and numerical comparisons among those tests.

Some procedures for testing structural change have been proposed for the case of unspecified change points. Brown et al. (1975) proposed the CUSUM test for stability in the coefficients of a linear regression, and this was based on the cumulative sum of recursively computed regression residuals. Ploberger and Krämer (1992) extended the CUSUM test to the OLS regression residuals. Andrews (1993) proposed several tests of structural stability based on GMM and examined the asymptotic properties under general conditions. Likelihood ratio test statistics for changes in the parameters of the normal population are investigated in Horváth (1993). Other relevant works on the subjects are given in the above papers.

Other related works include, Perron (1989) and Zivot and Andrews (1992) who considered the tests of the unit root hypothesis under breaking deterministic trends. Bai (1995) dealt with the estimation of the change point in the regression coefficients by LAD (least absolute deviation) and obtained the asymptotic distribution of the estimator of the change point. See Bai (1995) for literature on the estimation of the change point. In addition see Lai (1995) for a brief survey of sequential change point detection procedures.

2. Actual significance level of the Chow test

The regression models (1.1) and (1.2) can be written as

$$y = (X_1 \begin{pmatrix} \beta_1 \\ \beta_2 - \beta_1 \end{pmatrix}) + \varepsilon, \quad \varepsilon \sim N(0, \Sigma),$$

where $y = (y_1, y_2)$, $\varepsilon = (\varepsilon_1, \varepsilon_2)$, and $n = n_1 + n_2$. $\Sigma$ in (2.1) is represented as

$$\Sigma = \begin{pmatrix} \sigma_1^2 I_{n_1} & 0 \\ 0 & \sigma_2^2 I_{n_2} \end{pmatrix} = \begin{pmatrix} I_{n_1} & 0 \\ 0 & \frac{\sigma_2^2}{\sigma_1^2} I_{n_2} \end{pmatrix}.$$

Our null hypothesis is

$$H_0: \beta_2 = \beta_1 = 0$$

in (2.1) and the Chow test coincides with the ordinary $F$-test statistic, which is denoted by $F_c$. In Takemura and Honda (1994) we considered the effect of misspecified error covariance matrices on the actual size of the $F$-test in a normal linear regression model. The problem of heteroscedasticity was treated in Section 3.2. Here we will apply the results to this problem. Honda and Takemura (1994), reproduces the relevant results of Takemura and Honda
THE ACTUAL SIZE THE CHOW TEST

(1994) in English and is available from the first author.

We now introduce some notations and assumptions. For simplicity let $X_1$ and $X_2$ be of full rank hereafter, although our approach can be easily extended to the case where either $X_1$ or $X_2$ is not of full rank with appropriate changes of notation and degrees of freedom. Without loss of generality we can take $\sigma^2 = 1$. Then we have from (2.2)

$$\Sigma^{-1} - I_n = \begin{pmatrix} 0 & 0 \\ 0 & \left( 1 - \frac{1}{\sigma^2} \right) I_n \end{pmatrix} = \rho S,$$

where $\rho = \frac{1}{\sigma^2} - 1$ and $S$ is clearly defined. The argument which is analogous to Section 3.1 of Takemura and Honda (1994) shows that the actual significance level is represented uniquely as an absolutely convergent power series in $\rho \in (-1, 1)$

$$P(F_c \geq F_a(k, n-2k)) = \alpha + \sum_{i=1}^{\infty} c_i \rho^i,$$

where $F_a(k, n-2k)$ denotes the upper 100\% point of $F$-distribution with $k$ and $n-2k$ degrees of freedom. We showed in Takemura and Honda (1994) that in some cases, e.g. AR(1) errors, the explicit forms of $c_1$ and $c_2$ give interesting implications on the effect of misspecified error covariance matrices.

In our problem, the explicit forms of $c_1$ and $c_2$ reveal the relation between the values of regressors and the effect of heteroscedasticity. Let the following four orthogonal projection matrices be defined as

$$P_0 = \begin{pmatrix} X_1(X_1'X_1)^{-1}X_1 & 0 \\ 0 & X_2(X_2'X_2)^{-1}X_2 \end{pmatrix},$$

$$P_1 = \begin{pmatrix} X_1(X_1X_1 + X_2X_2)^{-1}X_1 & X_1(X_1X_1 + X_2X_2)^{-1}X_2 \\ X_2(X_1X_1 + X_2X_2)^{-1}X_1 & X_2(X_1X_1 + X_2X_2)^{-1}X_2 \end{pmatrix},$$

$$P_2 = P_0 - P_1, \quad \text{and} \quad P_3 = I_n - P_1 - P_2.$$

Note that $P_0$ and $P_1$ are the orthogonal projection matrices to the subspaces spanned by the columns of the regressor matrix in (2.1) and $(X_1', X_2')$ respectively. Besides, for notational convenience we define $P_{i,j}$ as

$$P_{i,j} = P(X_i^2 \geq d\chi^2_i),$$

where $\chi^2_i$ and $\chi^2_j$ are two independent $\chi^2$ variables with d.f. $i$ and $j$ respectively and

$$d = \frac{k}{n-2k} F_a(k, n-2k).$$

Now we make a close examination of $c_1$. General formulas of $c_i$ are given in (45) of Takemura and Honda (1994). The explicit form of $c_1$ in this case reduces to
Traces in (2.8) can be represented as
\[
\begin{align*}
\text{tr } S &= k, \\
\text{tr } S_\delta &= n_2 = n \left( \frac{1}{2} + \delta_0 \right), \\
\text{tr } S_\delta S_\delta &= \sum_{i=1}^{k} \left( \delta_i + \frac{1}{2} \right),
\end{align*}
\]
where $\delta_i$ is clearly defined. Denoting the eigenvalues of $X_\delta(X_1X_1 + X_2X_2)^{-1}X_2$ by $\lambda_i, i=1, \cdots, k$, we can rewrite $\text{tr } S_\delta$ as
\[
\text{tr } S_\delta = \sum_{i=1}^{k} \lambda_i = \sum_{i=1}^{k} \left( \delta_i + \frac{1}{2} \right) = \sum_{i=1}^{k} \delta_i + \frac{k}{2},
\]
where $\delta_i, i=1, \cdots, k$, are clearly defined. Finally we have
\[
(2.9) \quad c_1 = \frac{1}{2} \left( \sum_{i=1}^{k} \delta_i + \frac{n_k}{n-k} \delta_0 \right) (P_{k+2,n-2k} - P_{k,n-2k+2}).
\]
Note that the difference of probabilities in (2.9) is positive. The numerical values of the difference of probabilities in (2.9) are given in Table 1 for reference.

<table>
<thead>
<tr>
<th>$k$</th>
<th>$n$</th>
<th>$P_{k+2,n-2k} - P_{k,n-2k+2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>20</td>
<td>0.14055</td>
</tr>
<tr>
<td>2</td>
<td>30</td>
<td>0.14407</td>
</tr>
<tr>
<td>2</td>
<td>40</td>
<td>0.14565</td>
</tr>
</tbody>
</table>

$c_1$ in (2.9) gives us the following useful implication. When $\sigma_2 > \sigma_1$, i.e. $\rho$ is negative, $n_1 < n_2$, i.e. $\delta_0 \geq 0$, and $\sum_{i=1}^{k} \delta_i > 0$, then the actual significance level is less than the nominal significance level $\alpha$.

As an illustration we will consider a simple regression model where $k=2$ and
\[
(2.10) \quad X_1X_1 = \begin{pmatrix} 1 & \mu \\ \mu & 1 \end{pmatrix} \quad \text{and} \quad X_2X_2 = \frac{n_2}{n_1} \begin{pmatrix} 1 & \gamma \mu \\ \gamma \mu & \gamma^2 \end{pmatrix} = r \begin{pmatrix} 1 & \gamma \mu \\ \gamma \mu & \gamma^2 \end{pmatrix},
\]
where $r = n_2/n_1 = (1 + 2\delta_0)/(1 - 2\delta_0)$. Note that we standardized the second column of $X_1$ and $X_2$ using an Euclidean norm of the second column of $X_1$. It is then easy to derive
\[
(2.11) \quad \sum_{i=1}^{k} \delta_i = \frac{(r^2 \gamma^2 - 1)(1 - \mu^2)}{(1 + r)(1 + r \gamma^2) - (1 + r \gamma)^2 \mu^2}.
\]
We can suppose that $|\mu|<1$ assuming $X_1$ to be full rank. Therefore \((2.11)\) shows that $\sum_{i=1}^{k} \delta_i > 0$ for $r > 1$ and $c_1$ is positive when $r \geq 1$ and $\gamma > 1$. We usually have a negative $\rho$ with $\gamma > 1$ in this simple model. The actual significance level is then smaller than the nominal significance level with $r \geq 1$ because $\sum_{i=1}^{k} \delta_i > 0$, $\delta_0 \geq 0$, and $\rho$ is negative and not large. This is quite often the case for practical economic data and in statistical experiments with $n_1 \equiv n_2$.

It was previously found that the actual size could be less than the nominal size by numerical studies in Ohtani and Toyoda (1985). They ignored the effect of the regressors, however, by assuming $X_iX_i = X_2X_2$ and gave no analytical explanation. In addition, misleading accounts are found in some standard textbooks on econometrics. For example, see p. 217 of Johnston (1984).

Now we will present an explicit form of $c_2$. By tedious calculation we have from (45) of Takemura and Honda (1994)

\[
(2.12) \quad c_2 = \frac{1}{8} \left\{ a_1 \left( \sum_{i=1}^{k} \left( \delta_i - \frac{1}{4} \right) \right) + a_2 \left( \sum_{i=1}^{k} \delta_i \right)^2 + a_3 \sum_{i=1}^{k} \delta_i + a_4 \delta_0 \delta_2 + a_5 \delta_0 + a_6 \right\},
\]

where

\[
\begin{align*}
    a_1 &= 2a_2, \quad a_2 = P_{k,n-2k} + P_{k+4,n-2k} - 2P_{k+2,n-2k}, \\
    a_3 &= 2(P_{k+2,n-2k} - P_{k+4,n-2k}), \\
    a_4 &= 2k(-P_{k,n-2k} + 2P_{k+2,n-2k} - 2P_{k+2,n-2k} + P_{k+4,n-2k}) \\
    &\quad + 4(P_{k+4,n-2k} - P_{k+2,n-2k}), \\
    a_5 &= k^2(P_{k,n-2k} - 4P_{k,n-2k} + 2P_{k+2,n-2k} - 4P_{k+2,n-2k} + P_{k+4,n-2k} \\
    &\quad + 4P_{k+4,n-2k}) + 2k(3P_{k,n-2k} - P_{k+2,n-2k} + P_{k+2,n-2k} + 2P_{k+4,n-2k} - 4P_{k,n-2k}) \\
    &\quad + (P_{k+4,n-2k} - 4P_{k,n-2k}) - 4(P_{k,n-2k} - P_{k,n-2k} + 4P_{k,n-2k}), \\
    a_6 &= -2k(P_{k,n-2k} - 2P_{k,n-2k} + P_{k+2,n-2k}), \\
    a_7 &= \frac{k}{2}(-P_{k,n-2k} + P_{k+2,n-2k} - P_{k,n-2k} - 2P_{k,n-2k} + 2P_{k,n-2k}) \\
    &\quad + (P_{k,n-2k} - P_{k,n-2k}).
\end{align*}
\]

The numerical values of $a_i$ are given in Table 2 for reference.

<table>
<thead>
<tr>
<th>$k$</th>
<th>$n$</th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$a_3$</th>
<th>$a_4$</th>
<th>$a_5$</th>
<th>$a_6$</th>
<th>$a_7$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>20</td>
<td>0.01267</td>
<td>0.00633</td>
<td>-0.04390</td>
<td>0.14142</td>
<td>0.24389</td>
<td>-0.07808</td>
<td>0.00487</td>
</tr>
<tr>
<td>2</td>
<td>30</td>
<td>0.01474</td>
<td>0.00737</td>
<td>-0.04818</td>
<td>0.14521</td>
<td>0.23828</td>
<td>-0.07718</td>
<td>0.00344</td>
</tr>
<tr>
<td>2</td>
<td>40</td>
<td>0.01574</td>
<td>0.00787</td>
<td>-0.05024</td>
<td>0.14665</td>
<td>0.23506</td>
<td>-0.07666</td>
<td>0.00264</td>
</tr>
</tbody>
</table>

3. Simulation study

In this section we investigate how $\delta_i$ amplifies the discrepancy between the nominal and actual size in the presence of heteroscedasticity. In addition, we compare the 1st and 2nd order approximations.
with the actual size and examine the accuracy. The actual sizes are obtained by simulation of 200000 iterations. We deal here with the simple model in (2.10) and consider the cases where \( n_1 \leq n_2 \), \( \sigma_1^2 \leq \sigma_2^2 \), and \( \gamma \geq 1 \), which are of more practical interest. We take \( \mu = \sqrt{1/2} \) and \( \gamma = 1, \sqrt{2}, \) or 2 in (2.10). We carried out simulations for each case with \( n_1 = n_2 = 20 \) and \( n_1 = 20 \) and \( n_2 = 30 \). The numerical values are given in Tables 3-8.

We can see from Tables 3-8 that the first order approximation is accurate enough and the remarks following (2.10) indeed hold. Tables 3-4 show that the effect of heteroscedasticity is not so serious when \( n_1 = n_2 \) and \( \gamma \) is not large. We know from Tables 6-8, however, that we should be careful in using the Chow test unless \( n_1 \approx n_2 \). This is also the case when \( \gamma \) is large. We could say

<table>
<thead>
<tr>
<th>Table 3. Values of ( n_1 = n_2 = 20, \delta's = (0.0, 0.0, 0.0), \gamma = 1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma_1^2 / \sigma_1^2 )</td>
</tr>
<tr>
<td>Actual size</td>
</tr>
<tr>
<td>1st order approx.</td>
</tr>
<tr>
<td>2nd order approx.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 4. Values of ( n_1 = n_2 = 20, \delta's = (0.0, 0.20, -0.03), \gamma = \sqrt{2} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma_1^2 / \sigma_1^2 )</td>
</tr>
<tr>
<td>Actual size</td>
</tr>
<tr>
<td>1st order approx.</td>
</tr>
<tr>
<td>2nd order approx.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 5. Values of ( n_1 = n_2 = 20, \delta's = (0.0, 0.34, -0.07), \gamma = 2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma_1^2 / \sigma_1^2 )</td>
</tr>
<tr>
<td>Actual size</td>
</tr>
<tr>
<td>1st order approx.</td>
</tr>
<tr>
<td>2nd order approx.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 6. Values of ( n_1 = 20, n_2 = 30, \delta's = (0.1, 0.1, 0.1), \gamma = 1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma_1^2 / \sigma_1^2 )</td>
</tr>
<tr>
<td>Actual size</td>
</tr>
<tr>
<td>1st order approx.</td>
</tr>
<tr>
<td>2nd order approx.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 7. Values of ( n_1 = 20, n_2 = 30, \delta's = (0.1, 0.28, 0.06), \gamma = \sqrt{2} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma_1^2 / \sigma_1^2 )</td>
</tr>
<tr>
<td>Actual size</td>
</tr>
<tr>
<td>1st order approx.</td>
</tr>
<tr>
<td>2nd order approx.</td>
</tr>
</tbody>
</table>
Table 8. Values of $n_1=20$, $n_2=30$, $\delta_s=(0.1, 0.39, 0.03)$, $\gamma=2$

<table>
<thead>
<tr>
<th>$\sigma^2_{\delta}$</th>
<th>1.0</th>
<th>1.5</th>
<th>2.0</th>
<th>2.5</th>
<th>3.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actual size</td>
<td>0.061</td>
<td>0.034</td>
<td>0.026</td>
<td>0.022</td>
<td>0.019</td>
</tr>
<tr>
<td>1st order approx.</td>
<td>0.050</td>
<td>0.035</td>
<td>0.028</td>
<td>0.023</td>
<td>0.020</td>
</tr>
<tr>
<td>2nd order approx.</td>
<td>0.050</td>
<td>0.034</td>
<td>0.026</td>
<td>0.020</td>
<td>0.017</td>
</tr>
</tbody>
</table>

that we should avoid using the Chow test when $|n_1 - n_2|$ is large or $\sum_{i=1}^{k} \delta_i$ is large.

Here we considered only the cases where the actual size is less than the nominal size. When $c_1 < 0$ and $\rho < 0$, however, the actual size is larger than the nominal size. The above conclusion will be true in those cases as well.

Acknowledgements

The authors are grateful to an anonymous referee whose comments improved the earlier version of this paper.

REFERENCES

heteroscedasticity, *J. of Econometrics*, 27, 221-234.

