A Simple Artificial Regression Based LM Test of Asymmetry in the Logit Model

By

Anthony Murphy
Department of Economics
University College Dublin

October 1974

Working Paper WP94/17

DEPARTMENT OF ECONOMICS
UNIVERSITY COLLEGE DUBLIN
BELFIELD DUBLIN 4
A SIMPLE ARTIFICIAL REGRESSION BASED LM TEST OF ASYMMETRY IN THE LOGIT MODEL

Anthony Murphy.

Department of Economics,
University College Dublin,
Dublin 4, Ireland.


Abstract: A simple and convenient artificial regression based LM tests of asymmetry in the logit model is derived. The test does not use the outer product gradient (OPG) form and is thus likely to have fairly good small sample properties.

Keywords: Logit Model, Asymmetry, LM Tests, Artificial Regression.

JEL No: C35.

Introduction

The logit model is probably the most popular binary choice model used in econometrics. As Kiefer and Skoog (1984) and others have shown, mis-specification of the logit model generally results in inconsistent parameter estimates\(^1\). Thus estimated logit models should be subjected to a range of mis-specification tests.

Davidson and MacKinnon (1984b) and Engle (1984) derive convenient artificial regression based LM tests for omitted variables and neglected heterogeneity in the logit model. Because the logit model, like the probit model, is symmetric it is useful to consider a simple non-symmetric alternative which nests the logistic distribution a special case.

In this paper a convenient artificial regression based Lagrange Multiplier (LM) test of asymmetry of the logit model is derived using the Burr distribution. The derivation follows the lines of Davidson and MacKinnon (1984b) and joint tests for omitted variables, neglected heterogeneity, incorrect functional form and asymmetry may be performed\(^2\).

\(^1\)This paper was written whilst attending the 1994 Warwick Summer Research Workshop. The Workshop was supported by the Economic and Social Research Council (UK) and the Human Capital and Mobility Programme of the European Community.

\(^2\)Omitted uncorrelated explanatory variables and neglected heterogeneity both generate inconsistent parameter estimates, which is different from the classical regression model.

\(^3\)Pagan and Vella (1989) suggest a RESET like test of incorrect functional form. They suggest testing for the omission of powers of \(x^3\).
Since the proposed test is a LM test, it has desirable properties\(^5\). It is simple and convenient since it is regression based. In addition, it is likely to have good small sample properties since the outer product gradient (OPG) form of the LM test is not used\(^4\). The information matrix is not approximated by the outer product of the matrix of contributions to the score. Instead it is obtained as the expectation of this outer product.

The Burr Type II Distribution

In the logit model the probability of success \(p\) is:

\[
p = \frac{1}{1 + \exp(-x'\beta)}
\]

where \(x\) and \(\beta\) are the vectors of explanatory variables and coefficients. The logistic distribution is symmetric. It is a special case of the Burr type II distribution used by Smith (1988) and Lechner (1991) inter alia\(^5\). With this more general distribution:

\[
p = \frac{1}{(1 + \exp(-x'\beta))^\alpha}, \quad \alpha > 0
\]

If \(\alpha = 1\) then the distribution reduces to the logistic distribution and it is symmetric. If \(\alpha < 1\) the left hand tail of the distribution is heavier and if \(\alpha > 1\) the right hand tail is heavier.

The Score and Information Matrix

With a random sample of \(N\) individuals, the log likelihood \(l\) is:

\[
l = \sum_{i=1}^{N} [y_i \ln(p_i) + (1 - y_i) \ln(1 - p_i)]
\]

where the subscript \(i\) denotes individuals, \(y_i\) is an indicator variable (ie. \(y_i\) is 1 if \(i\) is successful and 0 otherwise) and \(p_i = (1 + \exp(-x'\beta))^\alpha\). Under the null hypothesis \(\alpha = 1\) the score is:

---


\(^4\) The OPG approximation to the information matrix was originally proposed by Berndt et. al. (1974). Davidson and MacKinnon (1983,1984a,1993) and MacKinnon (1992), inter alia, discuss the poor small sample properties of the OPG form of LM test.

\(^5\) Stukel (1988) considers a more complicated and general non-symmetric alternative to the logistic distribution.
\[
\frac{\delta l}{\delta \beta} = \sum_{i} \frac{y_i - p_i}{p_i(1-p_i)} p_i(1-p_i) x_i
\]
\[
\frac{\delta l}{\delta \alpha} = \sum_{i} \frac{y_i - p_i}{p_i(1-p_i)} p_i \ln p_i
\]

and the information matrix is composed of the following elements:

\[
I_{\beta \beta} = \lim_{N \to \infty} \frac{1}{N} \sum_{i} \frac{1}{p_i(1-p_i)} p_i^2 (1-p_i)^2 x_i x_i'
\]
\[
I_{\beta \alpha} = \lim_{N \to \infty} \frac{1}{N} \sum_{i} \frac{1}{p_i(1-p_i)} p_i^2 \ln p_i (1-p_i) x_i
\]
\[
I_{\alpha \alpha} = \lim_{N \to \infty} \frac{1}{N} \sum_{i} \frac{1}{p_i(1-p_i)} p_i^2 (\ln p_i)^2
\]

The information matrix is assumed to be non-singular. Let \( \theta' = (\beta', \alpha) \) denote the vector of parameters and let \( \tilde{\theta}' = (\tilde{\beta}', 1) \) denote the restricted vector of parameter estimates.

**LM Test Statistic**

The LM test statistic of \( \alpha = 1 \) is:

\[
LM = \frac{1}{N} \frac{\delta l}{\delta \tilde{\theta}'} \tilde{I}_{\tilde{\theta} \tilde{\theta}}^{-1} \frac{\delta l}{\delta \tilde{\theta}}
\]

\[
= \frac{1}{N} \left( \frac{\delta l}{\delta \alpha} \right)^2 \left( \tilde{I}_{\alpha \alpha} - \tilde{I}_{\alpha \beta} \tilde{I}_{\beta \beta}^{-1} \tilde{I}_{\beta \alpha} \right)^{-1}
\]

using the score and observed information matrix evaluated at the restricted parameter estimates. The observed information matrix is:

\[
\tilde{I}_{\tilde{\theta} \tilde{\theta}} = E \frac{1}{N} \frac{\delta l}{\delta \tilde{\theta}} \frac{\delta l}{\delta \tilde{\theta}'}
\]

Under weak regularity conditions, the LM test statistic is distributed as \( \chi^2(1) \) under the null.
Artificial Regression

The observed score and information matrix may be rewritten as:

\[
\frac{\delta l}{\delta \beta} = \sum_i \tilde{r}_i \tilde{s}_i = 0
\]

\[
\frac{\delta l}{\delta \alpha} = \sum_i \tilde{r}_i \tilde{w}_i
\]

\[
I_{\beta\beta}^o = \frac{1}{N} \sum_i \tilde{s}_i \tilde{s}_i'
\]

\[
I_{\alpha\alpha}^o = \frac{1}{N} \sum_i \tilde{w}_i \tilde{w}_i
\]

\[
I_{\alpha\beta}^o = \frac{1}{N} \sum_i \tilde{w}_i^2
\]

where:

\[
\tilde{p}_i = \frac{1}{(1 + \exp(x_i/\tilde{\beta}))}
\]

\[
\tilde{r}_i = \frac{y_i - \tilde{p}_i}{\sqrt{\tilde{p}_i(1 - \tilde{p}_i)}}
\]

\[
\tilde{s}_i = \frac{y_i - \tilde{p}_i}{\sqrt{\tilde{p}_i(1 - \tilde{p}_i)}} x_i
\]

\[
\tilde{w}_i = \frac{\tilde{p}_i \ln \tilde{p}_i}{\sqrt{\tilde{p}_i(1 - \tilde{p}_i)}}
\]

\[\tilde{p}_i\] is the predicted probability evaluated using the restricted parameter estimates. The \(\tilde{r}_i\) and \(\tilde{s}_i\) are just scaled residuals and regressors as Davidson and MacKinnon (1984b) point out. Thus \(\tilde{r}_i\) has a mean of zero and a unit variance. The LM test statistic may be rewritten as:

\[
LM = (\sum_i \tilde{r}_i \tilde{w}_i)^2 (\sum_i \tilde{w}_i^2 - \sum_i \tilde{w}_i \tilde{s}_i' (\sum_i \tilde{s}_i \tilde{s}_i')^{-1} \sum_i \tilde{s}_i \tilde{w}_i)^{-1}
\]

Thus the LM test statistic is just the explained sum of squares from the
uncentered regression of $\tilde{r}_i$ on $\tilde{s}_i$ and $\tilde{w}_i$. This proposition holds since
\[ \delta l/\delta \tilde{b} = \sum_i \tilde{r}_i \tilde{s}_i = 0. \]
It is easy to show that $N$ times the $R^2$ from this regression is asymptotically equal to the LM test statistic since
\[ NR^2 = \frac{LM}{(\sum_i \tilde{r}_i^2/N)} \]
and
\[ \text{plim} \ \sum_i \tilde{r}_i^2/N = 1. \]

Preliminary Monte carlo results show that the proposed test has better size properties that the OPG based test. However neither test is very powerful for small departures from symmetry. With the Pitman drift
\[ \alpha = 1 + \frac{\delta}{\sqrt{N}} \]
the non-centrality coefficient for the LM test is:
\[ \delta^2 (I_{12} - I_{11} I_{11}^{-1} I_{12}) \]
which tends to be small in practice.

Conclusion

A simple and convenient artificial regression based LM test of asymmetry in the logit model is derived. The test is not of the OPG form and thus is likely to have fairly good small sample properties.

References


