STRATHCLYDE

DISCUSSION PAPERS IN ECONOMICS

THE CONVENIENT CALCULATION OF SOME TEST STATISTICS IN MODELS OF DISCRETE CHOICE

BY

DARRYL HOLDEN AND ROGER PERMAN

NO 14-10

DEPARTMENT OF ECONOMICS
UNIVERSITY OF STRATHCLYDE
GLASGOW

The Convenient Calculation Of Some Test Statistics In Models of Discrete Choice

Darryl Holden[a], Roger Perman[a,b]

[a]University of Strathclyde, United Kingdom, [b]Correspondence to: Department of Economics, Sir William Duncan Building, 130 Rottenrow, Glasgow G4 0GE, United Kingdom (r.perman@strath.ac.uk)

ABSTRACT The paper considers the use of artificial regression in calculating different types of score test when the log−likelihood is based on probabilities rather than densities. The calculation of the information matrix test is also considered. Results are specialised to deal with binary choice (logit and probit) models.

keywords score test information matrix artificial regression

1 Introduction

Wilde (2008) has recently pointed out that the Bera−Jarque−Lee (BJL) test of normality in the probit model can be straightforwardly calculated, as the explained sum of squares from an artificial regression. This is useful because the BJL test is a Lagrange multiplier (or score) test based on expected second derivatives of the log−likelihood ($\mathrm{LM_{ESD}}$) and in general, and certainly in testing for normality in the probit model, see for example Holden (2004), an $\mathrm{LM_{ESD}}$ has superior finite sample behaviour to first derivative based 'outer product of the gradient' ($\mathrm{LM_{OPG}}$) alternatives. Thus the frequent trade−off, that $\mathrm{LM_{ESD}}$ delivers superior performance to $\mathrm{LM_{OPG}}$ at the cost of computational inconvenience, does not apply in testing for normality in the probit model, as noted by Wilde, who also stresses the importance of testing the normality assumption. In fact the same conclusion might be taken to hold more generally since there will be an artificial regression giving $\mathrm{LM_{ESD}}$ as an explained sum of squares whenever the likelihood function is based on probabilities rather than densities. The required argument, presented in section 2, mirrors that of Murphy (1994, 1996), although our presentation does not use the information matrix equality connecting expected first and second derivatives of the log−likelihood. That equality is the basis of the information matrix test, the calculation of which is discussed in section 3. The argument in section 3 is more straightforward, and more general, than that in Orme (1998).

2 Different score tests

Let y_{ij} indicate which of $J+1$ alternatives, $j = 0, \ldots, J$, applies to individual i, $i = 1, \ldots, n$, so that for each i one and only one of the y_{ij} equals one, with the remainder equalling zero. Assuming independence across i the log likelihood is given by $\ell(\boldsymbol{\theta}) = \sum_i \sum_j y_{ij} \ln(p_{ij}(\boldsymbol{\theta}))$, where $p_{ij}(\boldsymbol{\theta})$ is $\Pr(y_{ij} = 1)$. Differentiation gives

$$\frac{\partial \ell}{\partial \boldsymbol{\theta}} = \sum_i \sum_j \frac{y_{ij}}{p_{ij}} \frac{\partial p_{ij}}{\partial \boldsymbol{\theta}}, \tag{1}$$

and

$$\frac{\partial^2 \ell}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} = \sum_i \sum_j \left[\frac{y_{ij}}{p_{ij}} \frac{\partial^2 p_{ij}}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} - \frac{y_{ij}}{p_{ij}^2} \frac{\partial p_{ij}}{\partial \boldsymbol{\theta}} \left[\frac{\partial p_{ij}}{\partial \boldsymbol{\theta}} \right]' \right]. \tag{2}$$

Then, as $\mathrm{E}(y_{ij}) = p_{ij}$, we have

$$\mathrm{E}\left[\frac{\partial^2 \ell}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \right] = \sum_i \sum_j \left[\frac{\partial^2 p_{ij}}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} - \frac{1}{p_{ij}} \frac{\partial p_{ij}}{\partial \boldsymbol{\theta}} \left[\frac{\partial p_{ij}}{\partial \boldsymbol{\theta}} \right]' \right].$$

But $\sum_j p_{ij} = 1$ holds for all i so that

$$\sum_j \frac{\partial^2 p_{ij}}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} = 0,$$

implying

$$\mathrm{E}\left[\frac{\partial^2 \ell}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \right] = -\sum_i \sum_j \frac{1}{p_{ij}} \frac{\partial p_{ij}}{\partial \boldsymbol{\theta}} \left[\frac{\partial p_{ij}}{\partial \boldsymbol{\theta}} \right]' = -\sum_i \sum_j \boldsymbol{V}_{ij} \boldsymbol{V}'_{ij} \tag{3}$$

given the definition

$$\boldsymbol{V}_{ij} = \frac{1}{\sqrt{p_{ij}}} \frac{\partial p_{ij}}{\partial \boldsymbol{\theta}}.$$

We also have

$$\frac{\partial \ell}{\partial \boldsymbol{\theta}} = \sum_i \sum_j \boldsymbol{V}_{ij} v_{ij} \tag{4}$$

with

$$v_{ij} = \frac{y_{ij}}{\sqrt{p_{ij}}} \tag{5}$$

or

$$v_{ij} = \frac{y_{ij} - p_{ij}}{\sqrt{p_{ij}}}. \tag{6}$$

The second definition of v_{ij} will be appropriate in view of $\sum_j \partial p_{ij}/\partial\boldsymbol{\theta} = 0$ holding for all i. It follows from (3) and (4) that the score test based on expected second derivatives,

$$\mathrm{LM}_{\mathrm{ESD}} = \left(\left.\frac{\partial \ell}{\partial \boldsymbol{\theta}}\right|_R\right)' \left[-\mathrm{E}\left[\frac{\partial^2 \ell}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right]\Big|_R\right]^{-1} \left(\left.\frac{\partial \ell}{\partial \boldsymbol{\theta}}\right|_R\right),$$

can be presented as

$$\mathrm{LM}_{\mathrm{ESD}} = \left(\textstyle\sum_i\sum_j \boldsymbol{V}_{ij} v_{ij}\Big|_R\right)' \left[\textstyle\sum_i\sum_j \boldsymbol{V}_{ij}\boldsymbol{V}'_{ij}\Big|_R\right]^{-1} \left(\textstyle\sum_i\sum_j \boldsymbol{V}_{ij} v_{ij}\Big|_R\right), \tag{7}$$

where $|_R$ indicates evaluation at the restricted estimated $\boldsymbol{\theta}$. $\mathrm{LM}_{\mathrm{ESD}}$ is the uncentred explained sum of squares from the regression of $v_{ij}|_R$ on $\boldsymbol{V}_{ij}|_R$. This regression involves 'observations' across i and across j and, with v_{ij} as in (6) rather than (5), is called the 'discrete choice artificial regression' by Davidson and MacKinnon (2004, page 472). The alternative statistic

$$\mathrm{LM}_{\mathrm{OPG}} = \left(\textstyle\sum_i \boldsymbol{P}_i|_R\right)' \left[\textstyle\sum_i \boldsymbol{P}_i\boldsymbol{P}'_i|_R\right]^{-1} \left(\textstyle\sum_i \boldsymbol{P}_i|_R\right), \tag{8}$$

where

$$\boldsymbol{P}_i = \frac{1}{p_{ij^*}}\frac{\partial p_{ij^*}}{\partial\boldsymbol{\theta}}$$

if j^* denotes the j which applies to i, requires 'observations' across i only. It can be obtained as the uncentred explained sum of squares from the regression of one on $\boldsymbol{P}_i$. The accumulated evidence suggests the computational advantage of (8) over (7), to the extent that it exists, is obtained at the cost of finite sample behaviour under the null. The choice of (5) or (6) for v_{ij} will matter if asymptotically equivalent alternatives to (7), based on the uncentered R^2 from the regression of $v_{ij}|_R$ on $\boldsymbol{V}_{ij}|_R$, are explored.

2.1 Binary choice (logit and probit)

When $J = 1$, as in the logit and probit models, it is natural to adopt a slightly different notation, letting $p_i = p_i(\boldsymbol{\theta})$ be $\Pr(y_i = 1)$, with $\Pr(y_i = 0) = 1 - p_i$ therefore. The calculation of (7) then requires

$$\boldsymbol{V}_{i0} = \frac{-1}{\sqrt{1-p_i}}\frac{\partial p_i}{\partial\boldsymbol{\theta}},\ \boldsymbol{V}_{i1} = \frac{1}{\sqrt{p_i}}\frac{\partial p_i}{\partial\boldsymbol{\theta}},\ v_{i0} = \frac{1-y_i}{\sqrt{1-p_i}},\ \text{and } v_{i1} = \frac{y_i}{\sqrt{p_i}}. \tag{9}$$

But it is natural to obtain

$$\sum_j \boldsymbol{V}_{ij}\boldsymbol{V}'_{ij} = \frac{1}{p_i(1-p_i)}\frac{\partial p_i}{\partial \boldsymbol{\theta}}\left[\frac{\partial p_i}{\partial \boldsymbol{\theta}}\right]' = \boldsymbol{V}_i\boldsymbol{V}'_i \tag{10}$$

and

$$\sum_j \boldsymbol{V}_{ij}v_{ij} = \frac{y_i - p_i}{p_i(1-p_i)}\frac{\partial p_i}{\partial \boldsymbol{\theta}} = \boldsymbol{V}_i v_i,$$

where

$$\boldsymbol{V}_i = \frac{1}{\sqrt{p_i(1-p_i)}}\frac{\partial p_i}{\partial \boldsymbol{\theta}}, \text{ and } v_i = \frac{y_i - p_i}{\sqrt{p_i(1-p_i)}}, \tag{11}$$

to obtain

$$\mathrm{LM}_{\mathrm{ESD}} = \left(\textstyle\sum_i \boldsymbol{V}_i v_i|_R\right)' \left[\textstyle\sum_i \boldsymbol{V}_i \boldsymbol{V}'_i|_R\right]^{-1} \left(\textstyle\sum_i \boldsymbol{V}_i v_i|_R\right).$$

Now the required quantity is the uncentred explained sum of squares from the regression of $v_i|_R$ on $\boldsymbol{V}_i|_R$. This is the regression termed the 'binary response model regression' by Davidson and MacKinnon (2004, page 461), who note the application to testing for the significance of variables and to testing for heteroscedasticity. The BJL normality test, as considered by Wilde (2008), is based on

$$p_i = 1 - F(-\boldsymbol{X}'_i\boldsymbol{\beta} : c_1, c_2)$$

where $F(a : c_1, c_2)$ is the cdf of the Pearson family of distributions and c_1 and c_2 are parameters of the family with $H_0 : c_1 = c_2 = 0$ leading to the standard normal distribution and the probit model via $F(a) =_{H_0} \Phi(a)$. The $\boldsymbol{V}_i$ and v_i of (11), where $\boldsymbol{\theta} = \begin{pmatrix}\boldsymbol{\beta}' & c_1 & c_2\end{pmatrix}'$, and therefore the artificial regression giving $\mathrm{LM}_{\mathrm{ESD}}$, as in Wilde, follow straightforwardly from

$$\frac{\partial F(a)}{\partial a} =_{H_0} \phi(a),$$

$$\frac{\partial F(a)}{\partial c_1} =_{H_0} \frac{\phi(a)(a^2-1)}{3},$$

and

$$\frac{\partial F(a)}{\partial c_2} =_{H_0} -\frac{a\phi(a)(a^2+3)}{4}.$$

Nothing further is required when the ordered probit is considered, as in Weiss (1997), Glewwe (1997), and Johnson (1996). None of these papers make the existence of an artificial regression leading to $\mathrm{LM}_{\mathrm{ESD}}$ explicit although only Glewwe might be described as suggesting a requirement for difficult computations.

2.2 Other applications

The normality test presented by Lahiri and Song (1999) is an $\mathrm{LM}_{\mathrm{OPG}}$ statistic. But they present sufficient detail to make obtaining $\mathrm{LM}_{\mathrm{ESD}}$ straightforward. The model has $J = 3$ with the y_{ij} corresponding to their $I_1I_2I_3$, $I_1I_2(1-I_3)$, $I_1(1-I_2)$, and $1-I_1$. Then, for example, their $\partial \ln L/\partial \boldsymbol{\beta}_j$ means we can immediately write, in their notation,

$$
\begin{aligned}
\frac{\partial p_{i0}}{\partial \boldsymbol{\beta}_s} &=_{H_0} \phi_1(z_s)\Phi_2(z^s_{s'}, z^s_{s''}; \rho_{s's''.s})\boldsymbol{X}_s \\
\frac{\partial p_{i1}}{\partial \boldsymbol{\beta}_s} &=_{H_0} (H^{12}_s\phi_1(z_s)\Phi_1(z^s_{s'}) - \phi_1(z_s)\Phi_2(z^s_{s'}, z^s_{s''}; \rho_{s's''.s}))\boldsymbol{X}_s \\
\frac{\partial p_{i2}}{\partial \boldsymbol{\beta}_s} &=_{H_0} (H^{1}_s\phi_1(z_s) - H^{12}_s\phi_1(z_s)\Phi_1(z^s_{s'}))\boldsymbol{X}_s \\
\frac{\partial p_{i3}}{\partial \boldsymbol{\beta}_s} &=_{H_0} -H^{1}_s\phi_1(z_s)\boldsymbol{X}_s,
\end{aligned}
$$

for $s = 1, 2, 3$, with the other derivatives being equally straightforward.
Murphy (2007) notes that his proposed $\mathrm{LM}_{\mathrm{ESD}}$ statistic can be obtained via an artificial regression, and the displayed equation following his Table 2 is enough to enable the various $\partial p_{ij}/\partial \boldsymbol{\theta}$ to be identified. However the T_1, T_2, and T_3 of his Table 1 uses ϕ for three different purposes and would be better presented as

$$
T_1 = s_1\phi(\boldsymbol{x}'_1\boldsymbol{\beta}_1)\Phi\left[\frac{s_2(\boldsymbol{x}'_2\boldsymbol{\beta}_2 - \rho\boldsymbol{x}'_1\boldsymbol{\beta}_1)}{\sqrt{1-\rho^2}}\right],\ T_2 = s_2\phi(\boldsymbol{x}'_2\boldsymbol{\beta}_2)\Phi\left[\frac{s_1(\boldsymbol{x}'_1\boldsymbol{\beta}_1 - \rho\boldsymbol{x}'_2\boldsymbol{\beta}_2)}{\sqrt{1-\rho^2}}\right],
$$
$$
\text{and } T_3 = s_1s_2\phi_2(\boldsymbol{x}'_1\boldsymbol{\beta}_1, \boldsymbol{x}'_2\boldsymbol{\beta}_2, \rho)
$$

where ϕ and Φ have their usual meaning and ϕ_2 is the density for a bivariate standard normal with correlation ρ.

3 The information matrix test (TIM)

From (1) and (2) we have

$$
\textstyle\sum_i \left[\frac{\partial \ell_i}{\partial \boldsymbol{\theta}}\left[\frac{\partial \ell_i}{\partial \boldsymbol{\theta}}\right]' + \frac{\partial^2 \ell}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}'}\right] = \sum_i\sum_j \frac{y_{ij}}{p_{ij}}\frac{\partial^2 p_{ij}}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}'}
$$

implying that the IM test is based on $\sum_i \boldsymbol{d}_i$ where

$$
\boldsymbol{d}_i = \sum_j \frac{y_{ij}}{p_{ij}}\boldsymbol{W}_{ij}
$$

with

$$\boldsymbol{W}_{ij} = \text{vech}\left[\frac{\partial^2 p_{ij}}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right],$$

as noted in Weiss (1997) for the ordered logit and probit models. We have

$$\text{TIM} = \frac{1}{n}\left(\sum_i \boldsymbol{d}_i|_U\right)'\left(\boldsymbol{\Lambda}|_U\right)^{-1}\left(\sum_i \boldsymbol{d}_i|_U\right),$$

where $|_U$ indicates evaluation at the unrestricted estimated $\boldsymbol{\theta}$, $\boldsymbol{\Lambda} = \boldsymbol{A} - \boldsymbol{B}\boldsymbol{C}^{-1}\boldsymbol{B}'$,

$$\boldsymbol{A} = \text{E}\left[\frac{1}{n}\sum_i \boldsymbol{d}_i\boldsymbol{d}_i'\right],\ \boldsymbol{B} = \text{E}\left[\frac{1}{n}\sum_i \boldsymbol{d}_i\left[\frac{\partial \ell_i}{\partial\boldsymbol{\theta}}\right]'\right],\ \text{and}\ \boldsymbol{C} = \text{E}\left[\frac{1}{n}\sum_i \frac{\partial \ell_i}{\partial\boldsymbol{\theta}}\left[\frac{\partial \ell_i}{\partial\boldsymbol{\theta}}\right]'\right]. \tag{12}$$

Different estimates of $\boldsymbol{\Lambda}$ lead to different versions of TIM. The OPG version, TIM_{OPG}, is the uncentred explained sum of squares from the regression of one on $\boldsymbol{P}_i|_U$ and $\boldsymbol{Q}_i|_U$, where

$$\boldsymbol{Q}_i = \frac{1}{p_{ij^*}}\boldsymbol{W}_{ij^*}$$

and $\boldsymbol{P}_i$ and j^* are as previously defined. TIM_{OPG} does not require evaluation of the expectations in (12). When the expectations, and therefore $\boldsymbol{\Lambda}$, are obtained we have what Orme (1990) calls the 'efficient form' of TIM, TIM_{EF}. Orme suggests TIM_{EF} is likely to have superior performance compared to other versions of TIM. To obtain it requires

$$\text{E}(\boldsymbol{d}_i\boldsymbol{d}_i') = \sum_j \frac{1}{p_{ij}}\boldsymbol{W}_{ij}\boldsymbol{W}_{ij}' = \sum_j \boldsymbol{Z}_{ij}\boldsymbol{Z}_{ij}', \tag{13}$$

where

$$\boldsymbol{Z}_{ij} = \frac{1}{\sqrt{p_{ij}}}\boldsymbol{W}_{ij},$$

$$\text{E}\left[\boldsymbol{d}_i\left[\frac{\partial \ell_i}{\partial\boldsymbol{\theta}}\right]'\right] = \sum_j \frac{1}{p_{ij}}\boldsymbol{W}_{ij}\left[\frac{\partial p_{ij}}{\partial\boldsymbol{\theta}}\right]' = \sum_j \boldsymbol{Z}_{ij}\boldsymbol{V}_{ij}', \tag{14}$$

and

$$\text{E}\left[\frac{\partial \ell_i}{\partial\boldsymbol{\theta}}\left[\frac{\partial \ell_i}{\partial\boldsymbol{\theta}}\right]'\right] = \sum_j \frac{1}{p_{ij}}\frac{\partial p_{ij}}{\partial\boldsymbol{\theta}}\left[\frac{\partial p_{ij}}{\partial\boldsymbol{\theta}}\right]' = \sum_j \boldsymbol{V}_{ij}\boldsymbol{V}_{ij}', \tag{15}$$

in accordance with (3). It follows from (13), (14), (15), $\boldsymbol{d}_i = \sum_j \boldsymbol{Z}_{ij}v_{ij}$, and $\sum_i\sum_j \boldsymbol{V}_{ij}v_{ij}|_U = \boldsymbol{0}$, given (4), that TIM_{EF} can be obtained as the uncentred explained sum of squares from the artificial regression of $v_{ij}|_U$ on $\boldsymbol{V}_{ij}|_U$ and $\boldsymbol{Z}_{ij}|_U$ with 'observations' across i and j.

3.1 Binary choice (logit and probit)

In the framework of section 2.1 the calculation of $\mathrm{TIM}_{\mathrm{EF}}$ outlined above requires adding

$$\boldsymbol{Z}_{i0} = \frac{-1}{\sqrt{1-p_i}}\mathrm{vech}\left[\frac{\partial^2 p_i}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right], \text{ and } \boldsymbol{Z}_{i1} = \frac{1}{\sqrt{p_i}}\mathrm{vech}\left[\frac{\partial^2 p_i}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right]$$

to (9) but it is natural to add

$$\sum_j \boldsymbol{Z}_{ij}\boldsymbol{Z}'_{ij} = \frac{1}{p_i(1-p_i)}\mathrm{vech}\left[\frac{\partial^2 p_i}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right]\mathrm{vech}\left[\left[\frac{\partial^2 p_i}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right]\right]' = \boldsymbol{Z}_i\boldsymbol{Z}'_i,$$

where

$$\boldsymbol{Z}_i = \frac{1}{\sqrt{p_i(1-p_i)}}\mathrm{vech}\left[\frac{\partial^2 p_i}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right],$$

and

$$\sum_j \boldsymbol{Z}_{ij}\boldsymbol{V}'_{ij} = \boldsymbol{Z}_i\boldsymbol{V}'_i, \text{ and } \sum_j \boldsymbol{Z}_{ij}v_{ij} = \frac{y_i - p_i}{p_i(1-p_i)}\mathrm{vech}\left[\frac{\partial^2 p_i}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right] = \boldsymbol{Z}_i v_i$$

to (10). We discover that $\mathrm{TIM}_{\mathrm{EF}}$ can be obtained as the uncentered explained sum of squares from the regression of $v_i|_U$ on $\boldsymbol{V}_i|_U$ and $\boldsymbol{Z}_i|_U$. This is the regression identified by Orme (1988) as required in the logit and probit cases.

4 Conclusions

The computational convenience recently discovered by Wilde (2008) follows from existing results regarding the calculation of Lagrange multiplier tests in models where the likelihood function is based on probabilities rather than densities, as in Murphy (1994, 1996), Holden (1993), and elsewhere. Given this, section 2 might be viewed as a reminder, and illustration, of important results in the literature which merit wider dissemination. Section 3 considers the calculation of 'the efficient form' of the information matrix test via an artificial regression, and is both more general and more straightforward than the existing discussion of this topic.

5 References

Bera, A.K., Jarque, C.M., Lee, L.F., 1984. Testing the Normality Assumption in Limited Dependent Variable Models. International Economic Review 25, 563-578.

Davidson, R., MacKinnon, J., 2004. Econometric Theory and Methods. Oxford University Press.

Glewwe, P., 1997. A Test of the Normality Assumption in the Ordered Probit Model. Econometric Reviews 16(1), 1-19.

Holden, D., 1993. Normality and Homoscedasticity Tests in the Ordered Probit Model. Strathclyde Papers in Economics 93/4.

Holden, D., 2004. Testing the Normality Assumption in the Tobit Model. Journal of Applied Statistics 31, 521-532,

Johnson, Paul A., 1996. A Test of the Normality Assumption in the Ordered Probit Model. Metron LIV, 213-21.

Lahiri, K., Song, Jae G.,1999. Testing for Normality in a probit model with double selection. Economics Letters 65, 33-39.

Murphy, A, 'Artificial Regression Based Mis−Specification Tests for Discrete Choice Models', The Economic and Social Review, 26, 69−74, 1994.

Murphy, A., 1996. Simple LM tests of mis−specification for ordered logit models. Economics Letters 52, 137-141.

Murphy, A., 2007. Score Tests of Normality in Bivariate Probit Models. Economics Letters 95, 374-379,

Orme, C., 1988. The calculation of the information matrix test for binary data models. The Manchester School54, 370-376.

Orme, C., 1990. The Small−Sample Performance of the Information Matrix Test. Journal of Econometrics 46, 309-331.

Weiss, Andrew A., 1997. Specification Tests in Ordered Logit and Probit Models. Econometric Reviews 16(4), 361-391.

Wilde, J., 2008. A Simple Representation of the Bera−Jarque−Lee Test for Probit Models. Economics Letters 101, 119-121,