A Test for Ordered Probit Models Against Multinomial Probit Models

Masahito Kobayashi*

Faculty of Economics, Yokohama National University

Key words: Bond Ratings; Hypothesis Testing; Multinomial Probit Model; Ordered Probit Model,

JEL C35

Abstract

This paper shows that the ordered probit model with three categories is a special case of the multinomial probit model where the disturbance distribution is degenerate. It is shown that the conventional test statistics are unavailable for this problem because the derivative of the log likelihood with respect to disturbance correlation coefficient converges to zero identically under the null hypothesis, and a new feasible test statistic is proposed. The test is applied to the ratings of the corporate bonds and the ordered probit model is found to be inconsistent with the data.

The author would like to thank T. Amemiya, L.-F. Lee, M.-J. Lee, K. Nawata, and seminar participants at the Bank of Japan for their helpful comments. Part of this research was done while the author was a visiting researcher at the Institute of Monetary Economics, the Bank of Japan.

Corresponding Author: Masahito Kobayashi,

Faculty of Economics, Yokohama National University, Hodogaya-ku, Yokohama 240-8501, Japan,

Phone: +81-45-339-3544, Fax:+81-45-339-3518,E-mail: mkoba@ynu.ac.jp

1. Introduction

The ordered probit and logit models have been used for modeling bond ratings by several authors, for example, by Kaplan and Urwitz (1979), Ederington (1985), Cheung (1996), and Blume et al. (1998). These models assume that corporate bonds are rated by comparing a one-dimensional latent variable, which is expressed as a linear function of several financial variables, with threshold values. It is possible, however, that the ratings depends upon more latent variables; for example, debt might be essential in telling bad bonds from average bonds, although cash flow is important in telling good bonds from average bonds. In this case the multinomial probit and logit analysis are more appropriate.

The relation between the ordered and unordered models, which is this paper's main theme, has been almost neglected; a short reference to this problem can be found only in Amemiya (1985, p.293). No formal test for the ordered models against less restricted models has been proposed, although it is be very important to check the validity of the ordered models in practice.

We show that the ordered probit model is the limiting case of the multinomial probit model with three categories. Kobayashi (1999), which is the previous version of this paper, gave the proof, and Nawata (2000) generalized this result to the case of more than three categories, although he did not consider hypothesis testing.

The conventional tests such as the Lagrange multiplier test are unavailable for this problem, because the derivative of the log likelihood with respect to the correlation coefficient converges to zero under the null. Our case can be handled neither by the parameter transormation suggested by Cox and Hinkley (1974, pp.117-118) nor by the use of higher-order derivatives suggested by Lee and Chesher (1986), because the log likelihood function is a complicated function of parameters and explanatory variables. Then, in the case of three categories, we propose a feasible test for the ordered probit model against the multinomial probit model, which can be estimated without difficulty when the number of categories is three. The extension of the test to the general case is not pursued here because it is algebraically intractable.

In the next section the model and assumptions are illustrated, and the test is proposed in Section 3. In Section 4 the modeling of corporate bond ratings is analyzed and the ordered probit model is found to be inconsistent with the data.

2. Models and Assumptions

Assume that the latent variable for Categories A, B, and C are expressed as linear functions of explanatory variables:

Latent variable for Category A : $\alpha + \beta' x_i + e_{Ai}$

Latent variable for Category B: $0,$ (1)

Latent variable for Category C: $\gamma + \eta' x_i + e_{Ci},$

where x_i, $(i=1,...,n)$ is a $k\times1$ vector of explanatory variables, β and η are $k\times1$ coefficient vectors, and (e_{Ai}, e_{Ci}), $(i=1,...,n)$, follow independent bivariate normal distribution with zero means and covariance matrix

$$\begin{bmatrix} 1 & \rho\sigma \\ \rho\sigma & \sigma^2 \end{bmatrix}.$$

For the sake of standardization the latent variable for Category B is set at 0. The suffix i, which corresponds to the i-th individual, is dropped where there is no fear of ambiguity.

The multinomial probit model is defined by assuming that the category with the highest value of the latent variable is chosen; the dependent variable $y_{Ai}=1$ when Category A is chosen, and $y_{Ai}=0$ otherwise.

Definition (Multinomial Probit Model):

$$
\begin{aligned}
y_{Ai} &= 1 \text{ if } \alpha + \beta' x_i + e_{Ai} > 0 \text{ and } \gamma + \eta' x_i + e_{Ci} < \alpha + \beta' x_i + e_{Ai}, \\
&= 0 \text{ otherwise}, \\
y_{Bi} &= 1 \text{ if } \gamma + \eta' x_i + e_{Ci} < 0 \text{ and } \alpha + \beta' x_i + e_{Ai} < 0, \qquad (2) \\
&= 0 \text{ otherwise}, \\
y_{Ci} &= 1 \text{ if } \gamma + \eta' x_i + e_{Ci} > 0 \text{ and } \gamma + \eta' x_i + e_{Ci} > \alpha + \beta' x_i + e_{Ai}, \\
&= 0 \text{ otherwise}.
\end{aligned}
$$

In Figure 1 Category A is chosen when (e_A, e_C) is in the upper left-hand corner, namely when e_A is sufficiently large. Category B is chosen when (e_A, e_C) is in the lower left-hand corner, namely when neither of the disturbances are sufficiently large. Category C is chosen when (e_A, e_C) is in the lower right-hand corner, namely when e_C is sufficiently large.

When $\rho = -1$ and $\sigma = 1$, the distribution of (e_A, e_C) is degenerate and the probability mass concentrates on the line angled at 45 degrees from the upper left-hand corner to the lower right-hand corner. It follows from the conditions $\beta = -\eta$, $\rho = -1$ and $\sigma = 1$ that $\beta' x_i + e_{Ai} = -\eta' x_i - e_{Ci}$, and hence the unique latent variable $\beta' x_i + e_{Ai} = -\eta' x_i - e_{Ci}$ determines the choice. Then we have the ordered probit model. This intuitive explanation is stated formally in the next proposition.

Proposition 1: Assume that $\sigma = 1$, $\eta = -\beta$, $\alpha < -\gamma$, and $\rho = -1$. Then the multinomial probit model is reduced to the next ordered probit model:

$$
\begin{aligned}
y_A &= 1 \text{ if } \beta' x + e_A > -\alpha, \\
&= 0 \text{ otherwise}, \\
y_B &= 1 \text{ if } \beta' x + e_A > \gamma \text{ and } \alpha + \beta' x + e_A < -\alpha, \qquad (3)
\end{aligned}
$$

$= 0$ otherwise,

$y_C = 1$ if $\beta' x + e_A < \gamma$,

$= 0$ otherwise.

Proof:

When $\alpha + \beta' x + e_A > 0$, we have $\gamma + \eta' x + e_C < \alpha + \beta' x + e_A$, because we have $(\alpha + \beta' x + e_A) - (\gamma + \eta' x + e_C) > 2\alpha + 2\beta' x + 2e_A > 0$; the first inequality follows from $\beta' x + e_A = -(\gamma + \eta' x)$ and $-\gamma > \alpha$. Then we have $y_A = 1$ when $\alpha + \beta' x + e_A > 0$. It also follows, under Assumption 1, that $\gamma + \eta' x + e_C > \alpha + \beta' x + e_A$ from $\gamma + \eta' x + e_C > 0$, because we have $(\gamma + \eta' x + e_C) - (\alpha + \beta' x + e_A) > 2\gamma + 2\eta' x + 2e_C > 0$. Then we have $y_C = 1$ when $-\gamma + \beta' x + e_A < 0$. We can easily see that $y_B = 1$ when $-\gamma + \beta' x + e_A > 0$ and $\alpha + \beta' x + e_A < 0$, using $\eta = -\beta$ and $e_A = -e_C$. ■

3.Test Statistic

We here propose a test for the ordered probit model against the multinomial probit model in the case of three alternatives, for the null hypothesis $\rho = -1$, $\sigma = 1$, and $\eta = -\beta$. The test is is a modified version of the Lagrange multiplier test, which is based on the derivatives of the log likelihood given below. The algebraic detail is given in Appendix.

Let us denote P_{Ai}, for example, by the probability that Category A is chosen in the i-th observation. Then the log likelihood for the i-th observation is denoted as

$$\ell_i = y_{Ai} \log P_{Ai} + y_{Bi} \log P_{Bi} + y_{Ci} \log P_{Ci}. \tag{4}$$

Let us denote $\mu = -\alpha - \beta' x$ and $\tau = -\gamma + (\beta - \delta)' x$, where $\delta = \eta + \beta$, for the sake of notational simplicity. Also denote the derivative of the log likelihood with respect to ρ, for example, by $\partial \ell / \partial \rho = \ell_\rho$, neglecting the suffix i. Then we have the following derivatives:

Proposition 2:

The derivatives of the log likelihood evaluated under the null hypothesis are

$\ell_\sigma = \partial\ell/\partial\sigma = \phi(\tau)\,(-y_B/P_B + y_C/P_C)\tau,$

$\ell_\rho = \partial\ell/\partial\rho$

$= \lim_{\rho\to -1} (1/2)(2\pi)^{-1}(1-\rho^2)^{-1/2}\exp[-(1-\rho^2)^{-1}(\mu^2+\tau^2-2\rho\tau\mu)/2](-y_A/P_A + 2y_B/P_B - y_C/P_C)$

$= 0,$

$$\ell_\delta = \partial\ell/\partial\delta = \phi(\tau)(-y_B/P_B \;\; + y_C/P_C)x, \qquad (5)$$

$\ell_\alpha = \partial\ell/\partial\alpha = \phi(\mu)(y_A/P_A - y_B/P_B),$

$\ell_\gamma = \partial\ell/\partial\gamma = \phi(\tau)(y_C/P_C - y_B/P_B),$

$\ell_\beta = \partial\ell/\partial\beta = [\phi(\mu)y_A/P_A - (\phi(\mu) - \phi(\tau))\, y_B/P_B - \phi(\tau)y_C/P_C]x.$

$\phi(\mu)\,[y_A/P_A - y_B/P_B]x + \phi(\tau)\,[\,y_B/P_B - y_C/P_C]x,$

where $\delta=\beta+\eta$, $\mu = -\alpha - \beta'x$, $\tau = -\gamma - \eta'x$, $\phi()$ is the density of N(0,1), $\Phi()$ is the distribution function of N(0,1), $P_A = \Phi(\mu)$, $P_C = \Phi(-\tau)$, and $P_B = 1 - \Phi(-\tau) - \Phi(\mu)$ under the null hypothesis.

The derivative of the log likelihood with respect to ρ converges to zero identically under the null hypothesis, and hence the Lagrange multiplier test cannot be defined for this problem. Lee and Chesher (1986) proposed the use of high-order derivatives when the first-order derivative is identically zero. In our case, however, the higher-order derivatives also converge to zero exponentially, so that their method is inapplicable. The parameter transformation proposed by Cox and Hinkley (1976) is also inapplicable to our problem, because $\partial\ell/\partial\rho$ is a complicate function of parameters and explanatory variables. Then we now propose a test statistic by replacing the coefficients on y_A/P_A, y_B/P_B, and y_C/P_C in ℓ_ρ by their relative ratios, 1, -2, and 1, namely by using

$$\ell_0 = y_A/P_A - 2y_B/P_B + y_C/P_C, \tag{6}$$

instead of ℓ_ρ.

We now test the hypothesis by checking whether or not $\Sigma_{i=1}{}^n L_{0i}$ and $\Sigma_{i=1}{}^n L_{\delta i}$ are close to zero under the null hypothesis, where, for example, $L_{\delta i}$ is the estimator of $\ell_{\delta i}$ obtained by substituting the maximum likelihood estimator of α, β, and γ, and σ=1, δ=0, and ρ=−1. In the test statistics $\Sigma_{i=1}{}^n L_{\sigma i}$ is dropped, because $L_{\sigma i}$, $L_{\delta i}$, and $L_{\gamma i}$ are linearly dependent and $\Sigma_{i=1}{}^n L_{\gamma i}$ is zero because γ is estimated by the ML method. Then the proposed test statistic is

$$U'[\mathrm{var}(U)]^{-1}U, \tag{7}$$

where $U = n^{-1/2}(\Sigma_{i=1}{}^n L_{0i}, \Sigma_{i=1}{}^n L_{\delta i})'$. Because ℓ_0 obeys the condition to be satisfied by the derivatives of the conventional log likelihood function, the asymptotic variance of U is expressed as $\Sigma_{11} - \Sigma_{12}\Sigma_{22}{}^{-1}\Sigma'_{12}$, where

$$v_1 = n^{-1/2}(\Sigma_{i=1}{}^n \ell_{0i}, \Sigma_{i=1}{}^n \ell_{\delta i})',\ v_2 = n^{-1/2}(\Sigma_{i=1}{}^n \ell_{\alpha i}, \Sigma_{i=1}{}^n \ell_{\beta i}, \Sigma_{i=1}{}^n \ell_{\gamma i})',$$

$$\Sigma_{11} = E(v_1 v_1') - E(v_1)E(v_1'),$$

$$\Sigma_{22} = E(v_2 v_2') - E(v_2)E(v_2'),\ \Sigma_{12} = E(v_1 v_2') - E(v_1)E(v_2'), \tag{8}$$

and hence $U'\Sigma^{11}U$ follows the chi-square distribution with degrees of freedom k+1 asymptotically under the null hypothesis, where $\Sigma^{11} = (\Sigma_{11} - \Sigma_{12}\Sigma_{22}{}^{-1}\Sigma'_{12})^{-1}$ and k is the number of explanatory variables. The derivation is similar to that of the Lagrange multiplier test, so that it is not given here. Note that Σ^{11} is the (k+1)×(k+1) matrix in the upper left-hand corner of the inverse of the variance-covariance matrix of v_1 and v_2. Then Σ^{11} can be estimated by the corresponding sub-matrix of the inverse matrix of

$$H = n^{-1}\Sigma_{i=1}{}^n L_i L_i', \tag{9}$$

where $L_i = (L_{0i}, L_{\delta i}, L_{\alpha i}, L_{\beta i}, L_{\gamma i})'$. We then define the test statistic by

$$T = v_1{}'H^{11}v_1, \tag{10}$$

where $\begin{bmatrix} H^{11} & H^{12} \\ H^{21} & H^{22} \end{bmatrix} = H^{-1}$.

4. An Application to Corporate Bonds

We here estimate the ordered probit model to the ratings of the corporate bonds of 302 Japanese manufacturing companies given by Rating and Investment Information, Inc. (R & I) in 1997. The ratings are merged into three categories (AA and higher, A, BBB and lower), and explained by the following six explanatory variables, which are stated by R & I publicly to be referenced in their ratings: cash flow, debt / cash flow (in natural log), stockholders' equity ratio, ordinary income / total assets, interest coverage, ordinary transaction flow ratio, taken from the Nikkei Needs financial database. Four companies with negative cash flow are excluded from the sample, so that the sample size is 298.

In Table 1 the results of the estimation are given. It is strange the last three variables of the ordered probit model are not significant, since they have been referenced in ratings. The value of the proposed test statistic is 23.2; the ordered probit model is rejected, since the test statistic follows a chi-distribution with degree of freedom 7, and the multinomial probit model is suggested. The maximum likelihood estimation failed, however, because the likelihood function has no global maximum with respect to ρ and σ, as is shown in Figure 2. Then the multinomial probit model was fitted under the assumption of ρ=0 and σ=1, which had been accepted by the Lagrange Multiplier test, and the result is given in Table 1, and the predicted and actual ratings are illustrated in Figure 3.

The last three variables, which are not significant in the ordered probit model, are insignificant in the second latent variable, which corresponds to the ratings BBB and lower, but are significant in the first latent variable, which corresponds to the ratings AA and higher. It also shows that the differences between the coefficients are all significant or marginally significant for the three variables. Thus we can conclude that the three variables are unimportant in evaluating negatively appraised bonds but important in evaluating positively appraised bonds.

5. Extension to the Case of More Than Three Categories

It seems to me hopeless to generalize the test to the case of more than three categories, although it can be shown that the ordered probit model is nested in the multinomial probit model generally; the variance-covariance structure of the latter model is complicated when the number of categories is more than three, and I could not find a natural restriction on it. I have also failed to make a more general test by dividing the lower and upper categories into ordered subcategories under the alternative hypothesis. In practice, it is recommended to merge the categories into three in checking the validity of the ordered probit model with more than three categories.

References

Amemiya, T., 1996, *Advanced Econometrics*, Massachusetts, Harvard University Press,1985.

Blume, M.E., F. Lin, and A.C. MacKinlay, 1998, "The Declining Credit Quality of U.S. Corporate Debt: Myth and Reality," *Journal of Finance*, 53, 1389-1413.

Cheung, Stella, 1996, "Provincial Credit Ratings in Canada: An Ordered Probit Analysis," Working paper 96-6, Bank of Canada.

Cox, D. R. and D. V. Hinkley, 1976, *Theoretical Statistics*, London, Chapman and Hall̃.

Ederington, L.H. 1985, "Classification Models and Bonds Ratings," *The Financial Review*, 20, 237-262,.

Greene, W., *Econometric Analysis*, 3rd ed., Englewood Cliffs, Prentice Hall, 1997.

Kaplan, R. S. and G. Urwitz, 1979, "Statistical Models of Bond Ratings: A Methodological Inquiry," *Journal of Business*, 52, 231-261.

Kobayashi, M., 1999, "A Specification Test for Ordered Probit Models Against Multinomial Probit Models," Discussion Paper 1999-F-5, Faculty of Economics, Yokohama National University.

Lee, L. F. , and A. Chesher, 1986, “Specification Testing When Score Test Statistics are Identically Zero,” *Journal of Econometrics*, 31, 121-149.

Nawata, K., 2000, "Tako purobitto moderu to junnjo prubitto moderu no kanren ni tsuite,"(in Japanese), mimeo.

Appendix: Algebraic Details

We here obtain the derivatives in Proposition 3 under the null hypothesis, namely when $\rho = -1$, $\sigma = 1$ and $\eta = -\beta$. Let us denote

$$\mu = -\alpha - \beta'x, \ \tau = -\gamma - \eta'x, \tag{A.1}$$

for the sake of notational convenience, and we have $\mu + \tau > 0$ from $\alpha + \gamma < 0$. The correlated random variables e_A and e_C can be expressed as

$$e_A = u, \ e_C = \sigma[\rho u + (1-\rho^2)^{1/2} v], \tag{A.2}$$

where the random variables u and v follow N(0,1) independently. Then we have that

$$\begin{aligned}
P_A &= \Pr(-\mu + e_A > 0,\ -\tau + e_C < -\mu + e_A) = \int_{\mu}^{\infty}\int_{-\infty}^{\vartheta(u)} \phi(v)\phi(u)dvdu, \\
P_B &= \Pr(-\tau + e_C < 0,\ -\mu + e_A < 0) = \int_{-\infty}^{\mu}\int_{-\infty}^{\Psi(u)} \phi(v)\phi(u)dvdu, \\
P_C &= \Pr(-\tau + e_C > 0,\ -\tau + e_C > -\mu + e_A) = \int_{\tau/\sigma}^{\infty}\int_{-\infty}^{\Omega(u)} \phi(v)\phi(u)dvdu,
\end{aligned} \tag{A.3}$$

where

$$\begin{aligned}
\vartheta(u) &= (1-\rho^2)^{-1/2}\,[(u+\tau-\mu)/\sigma - \rho u], \\
\Psi(u) &= (1-\rho^2)^{-1/2}(\tau/\sigma - \rho u), \\
\Omega(u) &= (1-\rho^2)^{-1/2}(u\sigma + \mu - \tau - \rho u).
\end{aligned} \tag{A.4}$$

(1) Derivatives with respect to σ

The derivatives of the log likelihood can be obtained from the formula

$$\partial\ell/\partial\sigma = y_A(\partial P_A/\partial\sigma)/P_A + y_B(\partial P_B/\partial\sigma)/P_B + y_C(\partial P_C/\partial\sigma)/P_C.$$

When $\sigma = 1$,

$$\partial P_A/\partial\sigma = \int_{\mu}^{\infty}(\partial\vartheta(u)/\partial\sigma)\phi(\vartheta(u))\phi(u)du,$$

where

$\partial\vartheta(u)/\partial\sigma = (-1)(1-\rho^2)^{-1/2}(u+\tau-\mu)$,

$\phi(\vartheta(u)) = [2(1-\rho^2)^{-1/2}(2\pi)^{-1/2}\exp\{-(1-\rho^2)^{-1}2^2[u-(\mu-\tau)/2]^2/2\}]\,(1/2)(1-\rho^2)^{1/2}$.

Noting that the function in the bracket is a normal density with mean $(\mu-\tau)/2$ and infinitely small variance when $\rho\to-1$, which is denoted by $\delta(u-(\mu-\tau)/2)$ hereafter. Noting that, for a wide class of functions, say g(x), we have $\int_{-\infty}^{\infty}\delta(u-(\mu-\tau)/2)g(u)du = g((\mu-\tau)/2)$. We than have that

$$\partial P_A/\partial\sigma = \int_{\mu}^{\infty}\delta(u-(\mu-\tau)/2)(1/2)(u+\tau-\mu)\phi(u)du = 0$$

because $\mu+\tau>0$ ensures that $(\mu-\tau)/2$ is outside the integration region (μ, ∞).

We also have that

$$\partial P_B/\partial\sigma = \int_{-\infty}^{\mu}(\partial\Psi(u)/\partial\sigma)\,\phi(u)\phi(\Psi(u))du,$$

where

$\partial\Psi(u)/\partial\sigma = (-1)(1-\rho^2)^{-1/2}\tau$,

$\phi(\Psi(u)) = [(1-\rho^2)^{-1/2}(2\pi)^{-1/2}\exp(-(1-\rho^2)^{-2}(u+\tau)^2/2)](1-\rho^2)^{1/2}$.

Noting that he function in the bracket is a normal density with mean $-\tau$ ($<\mu$) and infinitely small variance, we have $\partial P_B/\partial\sigma = -\tau\phi(\tau)$. We can easily see that

$$\partial P_C/\partial\sigma = \tau\phi(\tau) \tag{A.5}$$

from the identity $\partial P_A/\partial\sigma + \partial P_B/\partial\sigma + \partial P_C/\partial\sigma = 0$.

(2) Derivatives with respect to α, γ, β, and $\delta = \beta+\eta$.

The derivatives of P_A with respect to μ is expressed as

$$P_A = \Pr(-\mu + e_A > 0,\ -\tau + e_C < -\mu + e_A) = \int_{\mu}^{\infty}\int_{-\infty}^{\vartheta(u)}\phi(v)\phi(u)dvdu,$$

$$\partial P_A/\partial\mu = -\phi(\mu)\int_{-\infty}^{\vartheta(\mu)}\phi(v)dv + \int_{\mu}^{\infty}\phi(\vartheta(u))\phi(u)[\partial\vartheta(u)/\partial\mu]du$$

$$= -\phi(\mu)\Phi(\vartheta(\mu)) - (1-\rho^2)^{-1/2}\int_{\mu}^{\infty}\phi(\vartheta(u))\phi(u)du,$$

The first term converges to $-\phi(\mu)$, since $\vartheta(\mu)=(1-\rho^2)^{-1/2}(\tau+\mu)$ increases infinitely and hence $\Phi(\vartheta(\mu))$ converges to unity under the null, and the second term disappears, because, as shown before, the density

$$(1-\rho^2)^{-1/2}\phi(\vartheta(u))=(1/2)(1-\rho^2)^{1/2}\delta(u-(\mu-\tau)/2)$$

has the probability mass outside the integration region (μ, ∞). Then $\partial P_A/\partial\mu = -\phi(\mu)$.

The derivative of P_B with respect to μ is expressed as

$$\partial P_B/\partial\mu = (\partial/\partial\mu)\int_{-\infty}^{\mu}\int_{-\infty}^{\Psi(u)}\phi(u)\phi(v)dvdu = \phi(\mu)\Phi(\Psi(\mu)).$$

Then, under the null, we have

$$\partial P_B/\partial\mu = \phi(\mu),$$

because $\Phi(\Psi(\mu))$ converges to 1, when σ=1 and ρ approaches -1. Then we see that

$$\partial P_C/\partial\mu = 0$$

from the identity $\partial P_A/\partial\mu + \partial P_B/\partial\mu + \partial P_C/\partial\mu = 0$.

From the symmetry of μ and τ we see that

$$\partial P_A/\partial\tau = 0,\ \partial P_B/\partial\tau = \phi(\tau),\ \partial P_C/\partial\tau = -\phi(\tau).$$

Then we can have the derivatives with respect to α, β, γ, and δ from $\mu = -\alpha - \beta'x$ and $\tau = -\gamma + (\beta-\delta)'x$ as follows:

$$\partial P_A/\partial\alpha = \phi(\mu),\ \partial P_B/\partial\alpha = -\phi(\mu),\quad \partial P_C/\partial\alpha=0,$$

$$\partial P_A/\partial\gamma = 0,\ \partial P_B/\partial\gamma = -\phi(\tau),\quad \partial P_C/\partial\gamma = \phi(\tau),$$

$$\partial P_A/\partial\beta = (\partial P_A/\partial\mu - \partial P_A/\partial\tau)(-x) = \phi(\mu)x,\ \partial P_B/\partial\beta = (\partial P_B/\partial\mu - \partial P_B/\partial\tau)(-x) = -[\phi(\mu) - \phi(\tau)]x,$$

$$\partial P_C/\partial\beta = (\partial P_C/\partial\mu - \partial P_C/\partial\tau)(-x) = -\phi(\tau)x, \qquad \text{(A.6)}$$

$$\partial P_A/\partial\delta = (-x)\partial P_A/\partial\tau = 0,\ \partial P_B/\partial\delta = (-x)\partial P_B/\partial\tau = -\phi(\tau)x,\ \partial P_C/\partial\delta = (-x)\partial P_C/\partial\tau = \phi(\tau)x.$$

(3) Derivatives with respect to ρ

First, noting that

$$\partial\vartheta(u)/\partial\rho = (1-\rho^2)^{-3/2}[\rho(-\mu+\tau)+(\rho-1)u],$$

we have

$$\partial P_A/\partial\rho = (1-\rho^2)^{-3/2}\int_\mu^\infty \phi(u)\phi((1-\rho^2)^{-1/2}[-\mu+\tau+(1-\rho)u])\,[\rho(-\mu+\tau)+(\rho-1)u]du.$$

After some algebra, we have

$$\partial P_A/\partial\rho$$

$$= (1/2)(2\pi)^{-1}(1-\rho^2)^{-1/2}C[\int_\mu^\infty B'(u)\exp(B(u))du+(1-\rho^2)^{-1}(\tau-\mu)\int_\mu^\infty(1+\rho)\exp(B(u))du]$$

where

$$B(u) = -(1-\rho^2)^{-1}(1-\rho)[u+(\tau-\mu)/2]^2,$$

$$C = \exp((-1/4)(1-\rho^2)^{-1}(1+\rho)(\tau-\mu)^2).$$

The first integral can be expressed as

$$-(2\pi)^{-1}(1-\rho^2)^{-1/2}\exp(-(1-\rho^2)^{-1}(1-\rho)(\tau+\mu)^2/4).$$

The second integral can be neglected in comparison with the first integral, because

$$|(1-\rho^2)^{-1}\int_\mu^\infty(1+\rho)\exp(B(u))du\,/\int_\mu^\infty B'(u)\exp(B(u))du\,|$$

$$< \quad (1-\rho^2)^{-1}(1+\rho)\int_\mu^\infty\exp(B(u))du\,/|B'(\mu)|\int_\mu^\infty\exp(B(u))du$$

$$= \quad (1-\rho^2)^{-1}(1+\rho)/[(1-\rho^2)^{-1}(1-\rho)(\tau+\mu)] \to 0,$$

when ρ converges to -1, because

$$B'(u) = -2(1-\rho^2)^{-1}(1-\rho)[u+(\tau-\mu)/2]$$

is negative and decreasing for $u > \mu$. Then we have that

$$\partial P_A/\partial\rho = -(1/2)(2\pi)^{-1}(1-\rho^2)^{-1/2}\exp[(-1/2)(1-\rho^2)^{-1}(\mu^2+\tau^2-2\rho\tau\mu)](1+o(1)). \qquad (A.7)$$

Finally, we have

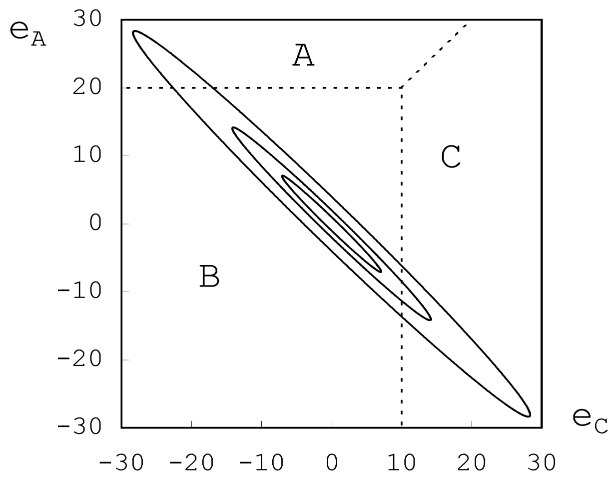
$$\partial P_C/\partial\rho = \partial P_A/\partial\rho,\ \partial P_B = -2\partial P_A/\partial\rho, \qquad (A.8)$$

from the symmetry of (A.7) and the identity $\partial P_A/\partial\mu + \partial P_B/\partial\mu + \partial P_C/\partial\mu = 0$.

Table 1: Estimation by the Ordered and Multinomial Probit Models

	ordered Probit	multinomial probit: first latent variable	multinomial probit: second latent variable	difference
constant	-21.9	-26.0	-19.3	6.7
	(-11.3)	(-6.08)	(-8.30)	(1.37)
cash flow (in natural log)	1.77	2.03	1.82	-0.21
	(11.3)	(6.18)	(8.75)	(-0.53)
debt / cash flow	0.070	0.100	0.071	-0.029
	(5.53)	(2.34)	(5.02)	(-0.65)
stockholders' equity ratio	0.072	0.106	0.068	-0.038
	(7.41)	(4.52)	(5.72)	(-1.46)
ordinary income / total assets	0.0849	0.299	0.020	-0.279
	(1.75)	(2.76)	(0.344)	(-2.26)
interest coverage	1.17	-3.62	1.37	4.99
	(0.157)	(-1.98)	(1.36)	(2.39)
ordinary transaction flow ratio	-4.35	-1.08	-0.00	1.08
	(-1.87)	(-2.34)	(-0.006)	(1.77)
threshold	1.82			
	(22.1)			

Note: For the ease of comparison the signs of the coefficients in the second latent variables are reverted. T-values are given in parentheses.



Figigure 1: Disturbance Distribution and Choices Under the Null Hypothesis

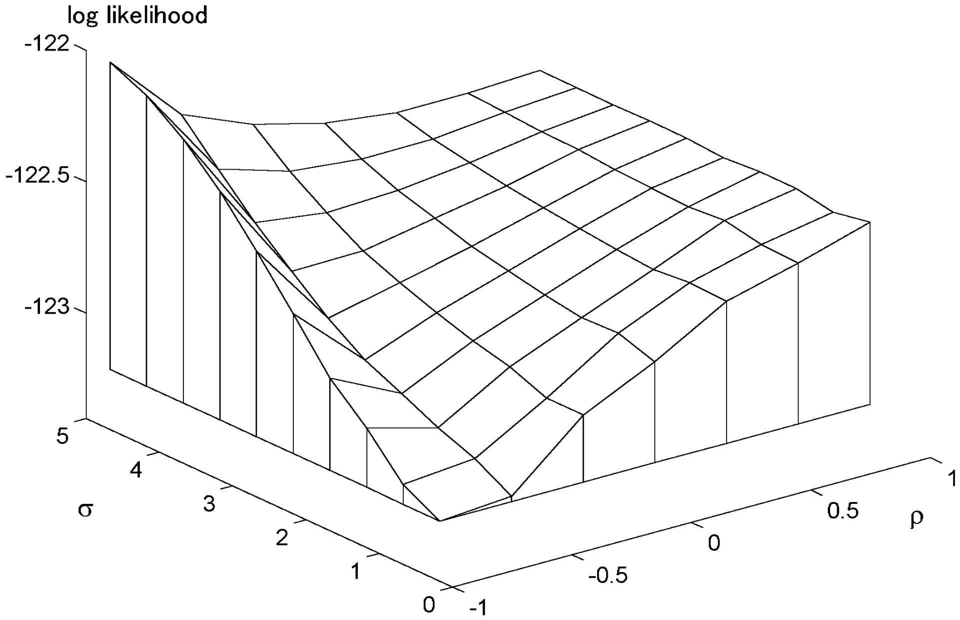


Figure 2: concnetratd log likelihood

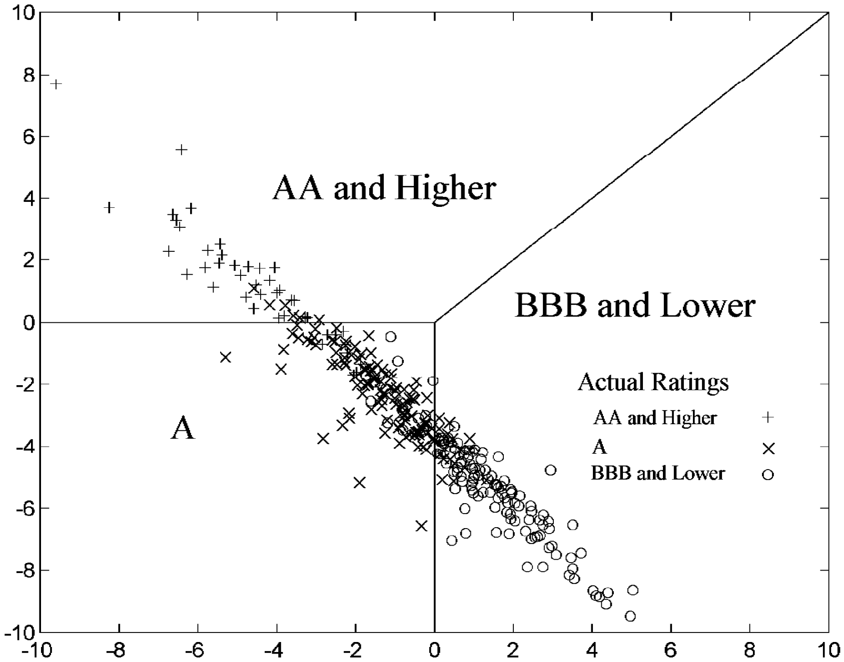


Figure 3: Latent Variables and Actual Ratings