A Specification Test for Ordered Probit Models Against Multinomial Probit Models

Masahito Kobayashi, Yokohama National University

Abstract This paper proposes a specification test for ordered probit models against multinomial probit models. It is shown that the ordered probit model is a limiting case of the multinomial probit model where the disturbances follow a degenerate distribution when the number of alternatives is three. The conventional test statistics are unavailable, because the derivative of the log likelihood with respect to correlation coefficient is 0 identically under the null. A feasible test statistic is proposed by modifying the Lagrange multiplier test and the performance is examined by a Monte Carlo experiment.

1. INTRODUCTION

The ordered probit and logit models are natural extensions of the binary probit and logit models when more than two alternatives are ordered. However, the ordered alternatives are sometimes no self-evident, as is illustrated by Ii and Ohkusa (1999). Suppose that a patient with a common cold chooses from the three alternatives: Doing nothing, Buying medicines over the counter, and Consulting a doctor. Normally, the patient does nothing for a slight cold, then buys aspirin when headache and fever are unbearable, and finally consults a doctor when the symptom is serious. The ordered model is appropriate in this situation. However, if health insurance makes the cost of consulting a doctor lower than the price of medicines, one would prefer consulting a doctor to buying medicines even in the case of a slightest cold. Then the alternatives are no longer ordered.

In considering unordered alternatives, the multinomial logit analysis is a popular tool, for example, in analyzing the modal choice of commuting. This model has an unrealistic property that, when an alternative is dropped, the probabilities of choosing the remaining alternatives increase proportionately. The unreality of this assumption can be easily recognized by the well-known example: if Red Bus is dropped from the modes of commuting consisting of Red Bus, Blue Bus, and Train, the actual probability of choosing Blue Bus would be higher than the predicted value by the multinomial logit model. The color of the bus is of little significance to commuters and hence most of the commuters by Red Bus would choose Blue Bus.

This independence of irrelevant alternatives (IIA) is a consequence of the assumption that the random utilities are distributed independently and identically. We can make the model more realistic by assuming that similar alternatives have stronger correlation than unrelated alternatives. The nested logit model of McFadden (1978) generalized the multinomial logit model by introducing correlated random utilities. The multinomial probit model is less popular in the literature, with an exception of Hausman and Wise (1978), although it is more flexible than the nested logit model. However, the computational burden, to which the unpopularity has been attributed, poses no longer difficulties in the case of three alternatives, because it requires only two-dimensional numerical integration.

Tests for the nested logit model against the multinomial logit model have been proposed by Hausman and McFadden(1984) and McFadden(1987), and have been discussed in connection with the IIA assumption. However, 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 test for the ordered models against less restricted models has been proposed, although it would be very useful for checking whether or not the ordered alternatives are appropriate in practice.

We show that the ordered probit model is the limiting case of the multinomial probit model when the correlation coefficient of the disturbances, say ρ, converges to -1 in the case of three alternatives. The null and alternative models are locally unidentifiable and hence the conventional tests such as the Lagrange multiplier test cannot be defined, because the derivative of the log likelihood with respect to ρ converges to 0 when ρ approaches -1. Our case can be handled neither by the reparameterization suggested by Cox and Hinkley (1974, pp.117-118) nor by the use of higher-order derivatives suggested by Lee and

Chesher (1986). Then we propose a feasible test by modifying the Lagrange multiplier test for the ordered probit model against the multinomial probit model. The performance of the test is examined by Monte Carlo experiments.

2. MODELS

Assume that the alternative that gives the highest utility is chosen from the alternatives A, B, and C, and that their utilities are given as follows:

General Definition:

Utility of A : $\alpha + \beta' x_i + e_{Ai}$

Utility of B: 0,

Utility of C: $\gamma + \eta' x_i + e_{Ci}$

where x_i, $(i=1,\ldots,n)$ is a $k\times 1$ vector of regression variables, β and η are $k\times 1$ coefficent vectors, and (e_{Ai}, e_{Ci}), $(i=1,\ldots,n)$, 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 utility for B is set at 0, and the variance of the utility for A at 1. The suffixes A, B, and C correspond to the alternatives A, B, and C, respectively. The suffix i, which corresponds to the i–th individual, is dropped where there is no fear of ambiguity.

Assume that the dependent variable y_A takes 1 when the alternative A is chosen, and 0 otherwise, Then the multinomial probit model can be defined as follows:

Definition (Multinomial Probit Model):

$y_{Ai} = 1$ if $\alpha + \beta' x_i + e_{Ai} > 0$ and $\gamma + \eta' x_i + e_C < \alpha + \beta' x_i + e_{Ai}$,
$= 0$ otherwise,

$y_{Bi} = 1$ if $\gamma + \eta' x_i + e_{Ci} < 0$ and $\alpha + \beta' x_i + e_{Ai} < 0$,
$= 0$ otherwise,

$y_{Ci} = 1$ if $\gamma + \eta' x_i + e_{Ci} > 0$ and $\gamma + \eta' x_i + e_{Ci} > \alpha + \beta' x_i + e_{Ai.}$
$= 0$ otherwise.

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

$$\lambda_i = y_{Ai}\log P_{Ai} + y_{Bi}\log P_{Bi} + y_{Ci}\log P_{Ci},$$

where

$P_A = \Pr(y_{Ai}=1) = \Pr(\alpha + \beta' x_i + e_{Ai} > 0,\ \gamma + \eta' x_i + e_{Ci} < \alpha + \beta' x_i + e_{Ai})$,

$P_{Bi} = \Pr(y_{Bi}=1) = \Pr(\gamma + \eta' x_i + e_{Ci} < 0,\ \alpha + \beta' x_i + e_{Ai} < 0)$,

$P_{Ci} = \Pr(y_{Ci}=1)$
$= \Pr(\gamma + \eta' x_i + e_{Ci} > 0,\ \gamma + \eta' x_i + e_{Ci} > \alpha + \beta' x_i + e_{Ai})$.

In Figure 1 the alternative A is chosen when (e_A, e_C) is in the upper left–hand corner, namely when e_{Ai} is sufficiently large. The alternative B is chosen when (e_A, e_C) is in the lower left–hand corner, namely when either of the disturbances are not sufficiently large. The alternative C is chosen when (e_A, e_C) is in the lower right–hand corner, namely when e_C is sufficiently large.

If $\rho = -1$ and $\sigma = 1$, then the distribution of (e_A, e_C) degenerates and the probability mass concentrates on the line angled at 45 degrees from the upper left–hand corner to the lower right–hand corner. The conditions $\beta = -\eta$ and $\alpha < -\gamma$ ensure that the line is divided by the three regions for any x. Then, because the values of α, γ, and $\beta' x_i + e_{Ai} = -\eta' x_i - e_{Ci}$ determine the choice completely, we have the ordered probit model. The next proposition shows formally that the multinomial probit model is reduced to the ordered probit model under these conditions.

[Figure 1]

Assumption 1:

$$\sigma = 1,\ \eta = -\beta,\ \alpha < -\gamma, \text{ and } \rho = -1.$$

Proposition 1: Under Assumption 1 the nested model is reduced to the next ordered probit model:

$y_A = 1$ if $\alpha + \beta' x + e_A > 0$,
$= 0$ otherwise,

$y_B = 1$
if $-\gamma + \beta' x + e_A > 0$
and $\alpha + \beta' x + e_A < 0$,
$= 0$ otherwise,

$y_C = 1$ if $-\gamma + \beta' x + e_A < 0$
$= 0$ otherwise.

Proof:

It follows that $\gamma + \eta' x + e_C < \alpha + \beta' x + e_A$ from $\alpha + \beta' x + e_A > 0$, because we have $(\alpha + \beta' x + e_A) - (\gamma + \eta' x + e_C) > 2\alpha + 2\beta' x + 2e_A$ under Assumption 1. 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 $(\gamma + \eta' x + e_C) - (\alpha + \beta' x + e_A) > 2\gamma + 2\eta' x + 2e_C$. 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 $e_A = -e_C$.
Q.E.D.

3. TEST STATISTIC

We here propose a test for the ordered probit model against the multinomial probit model, assuming the null hypothesis $\rho = -1$, $\sigma = 1$, and $\eta =$

$-\beta$. The proposed test is based on the derivatives of the log likelihood given below. The algebraic detail is given in Appendix.

Proposition 2:

Let us denote

$\mu = -\alpha - \beta' x$ and $\tau = -\gamma + (\beta - \Delta)' x$,

where $\Delta = \eta + \beta$. Then, under the null hypothesis, the derivatives of the log likelihood for the i–th observation are given as follows

$\lambda_\sigma = \partial\lambda/\partial(1/\sigma) =$

$(y_A/P_A - y_C/P_C)(\alpha-\gamma)(1/4)\phi((\alpha-\gamma)/2) + (y_B/P_B - y_C/P_C)\phi(\tau)\tau$,

$\lambda_\rho = \partial\lambda/\partial\rho$

$= \lim_{\rho\to-1} (1/2)(2\pi)^{-1}(1-\rho^2)^{-1/2}\exp(-(1-\rho^2)^{-1}(\mu^2 + \tau^2 - 2\rho\tau\mu)/2)(-y_A/P_A + 2y_B/P_B - y_C/P_C) = 0$,

$\lambda_\Delta = \partial\lambda/\partial\Delta = \phi(\tau)(-y_B/P_B + y_C/P_C)x$,

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

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

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

This proposition shows that the null and alternative hypotheses are locally unidentifiable near the null hypothesis because the derivative of the log likelihood with respect to ρ is zero identically under the null hypothesis, and hence the conventional Lagrange multiplier test statistic cannot be defined for this problem. Lee and Chesher (1986) proposed the general principle of extremum tests by using the higher-order derivatives when the first-order derivative is identically zero. In our case, however, higher-order derivatives also converges to zero exponentially. The reparameterization proposed by Cox and Hinkley (1976) is inapplicable to our problem, because $\partial\lambda/\partial\rho$ is a complicate function of parameters and regressor variables.

Then we now propose a simpler feasible test statistic by replacing the coefficients on y_A/P_A, y_B/P_B, and y_C/P_C in λ_ρ by their relative ratios, 1, –2, and 1, namely by using

$\lambda_0 = y_A/P_A - 2y_B/P_B + y_C/P_C$.

instead of λ_ρ. This substitute, which would weaken the power, might be justifiable because the conventional test statistics are unavailable for our problem. The proposed test statistic is

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

where $U = n^{-1/2}(\Sigma_{i=1}^n\Lambda_{0i}, \Sigma_{i=1}^n\Lambda_{\Delta i}, \Sigma_{i=1}^n\Lambda_{\sigma i})'$, and Λ_0, Λ_Δ, and Λ_σ are the estimator of λ_0, λ_Δ, and λ_σ obtained by substituting the maximum likelihood estimators of α, β, and γ. Because λ_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

$v_1 = n^{-1/2}(\Sigma_{i=1}^n\lambda_{0i}, \Sigma_{i=1}^n\lambda_{\Delta i}, \Sigma_{i=1}^n\lambda_{\sigma i})'$,

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

$\Sigma_{11} = E(v_1v_1')$,

$\Sigma_{22} = E(v_2v_2')$,

$\Sigma_{12} = E(v_1v_2') - E(v_1)E(v_2')$,

and hence $U'\Sigma^{11}U$ follows the chi-square distribution with the degrees of freedom 3 asymptotically, where $\Sigma^{11} = (\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12}')^{-1}$. The derivation is similar to that of the Lagrange multiplier test, so that it is not given here. Note that Σ^{11} is the 3×3 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

$H = n^{-1}\Sigma_{i=1}^n \Lambda_i \Lambda_i'$,

where $\Lambda_i = (\Lambda_{0i}, \Lambda_{\Delta i}, \Lambda_{\sigma i})'$. We then define the test statistics by

$T = v_1' H^{11} v_1$,

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

4. MONTE CARLO EXPERIMENTS

We here examine the powers of the feasible test using Monte Carlo experiments. The independent variable in our experiment is a realization of random variables that follow NID(0,1), and is fixed throughout the experiment. The sample size is 400 and the number of replications is 2000. The regression coefficients are set at $\alpha = -0.40$, $\gamma = -0.50$, $\beta = -0.8$, and $\eta = 0.8$ under the null hypothesis. Table 1 shows that the actual sizes, 0.1305 and 0.0665, are larger than the nominal values 0.10 and 0.05. The test has sufficient power when $\sigma^2 \neq 1$, as well as when the regression coefficients differ, namely $\beta \neq -\eta$. However, the power of the test is not very high even when ρ differs from the null hypothesis substantially, namely when $\rho=0$ and $\rho=0.5$. This is no surprising to some degree because $\partial\lambda/\partial\rho = 0$ means that the difference in ρ is difficult to tell only by the likelihood, at least, locally.

[Table 1]

5. CONCLUDING REMARKS

It should be noted that the result of this paper applies only to the case of three alternatives, which is of most importance practically. It might be possible to construct a test for models with more than three alternatives by assuming additional structures. It is also suspected that the substitutions in $\partial\lambda/\partial\rho$ is responsible for the low power of the test when $\rho \neq -1$. Searching for alternative tests with higher power is an interesting problem yet to be examined. However, the both problems are far beyond the scope of this paper, and are left to further research.

6. REFERENCES

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

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

Hausman, J. A., and D. A. Wise, 1978, "A Conditional Probit for Qualitative Choice: Discrete Decisions Recognizing Interdependence and Heterogeneous Preferences," *Econometrica*, 46, 403-426.

Hausman, J. A., and D. McFadden, 1984, "Specification Tests for the Multinomial Logit Model, " *Econometrica*, 52, 1219-40.

Ii, M., and Y. Ohkusa,1999, "Is There Any Substitution Between Medical Services and Over-the-Counter Medications in the Case of the Common Cold?", Discussion Paper No. 478, The Institute of Social and Economic Research, Osaka University.

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

McFadden, D. 1987, "Regression-Based Specification Tests for the Multinomial Logit Model," *Journal of Econometrics*, 34, 63-82.

7. APPENDIX

The derivatives of the log likelihood are given here under the null, namely when $\rho = -1$, $\sigma = 1$, and $\eta = -\beta$. Let us denote

$$\mu = -\alpha-\beta'x,\ \tau = -\gamma-\eta'x,$$

for the sake of notational convenience. The correlated random variables e_A and e_C can be expressed as

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

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

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

where

$$\vartheta(u)= (1-\rho^2)^{-1/2}[(u+\tau-\mu)/\sigma-\rho u],$$
$$\Psi(u)=(1-\rho^2)^{-1/2}(\tau/\sigma- \rho u),$$
$$\Omega(u)=(1-\rho^2)^{-1/2}(u\sigma+\mu-\tau -\rho u).$$

The derivatives of the log likelihood can be obtained from the derivatives of P_A, P_B, and P_C by using the formula

$$\partial\lambda/\partial\mu = y_A(\partial P_A/\partial\mu)/P_A + y_B(\partial P_B/\partial\mu)/P_B + y_C(\partial Pc/\partial s\mu)/P_C,$$

for example.

(1) Derivatives with respect to $1/\sigma$

We here obtain the derivatives of the log likelihood with respect to $s=1/\sigma$, instead of σ, for the sake of convenience. When $\sigma=1$, namely when $s=1$, we have

$$\partial P_A/\partial s= \int_\mu^\infty(\partial\vartheta(u)/\partial s)\phi(\vartheta(u))\phi(u)du$$
$$= \int_\mu^\infty(u+\tau-\mu) [2\pi\omega(\rho)]^{-1/2}\exp(-(1/2)\omega(\rho)^{-1}[u + (\tau-\mu)/2]^2)du$$
$$2^{-1/2}(1-\rho)^{-1/2}(2\pi)^{-1/2}\exp(-(1/4)(1-\rho)^{-1}(\tau-\mu)^2),$$

where $\omega(\rho)=(1+\rho)/2$. Noting that, when ρ converges to -1, the density function $[2\pi\omega(\rho)]^{-1/2}\exp(-(1/2)\omega(\rho)^{-1}[u +(\tau-\mu)/2]^2)$, has the whole mass at $-(\tau-\mu)/2$. Then we have only to substitute u with $-(\tau-\mu)/2$, because the assumption $\tau+\mu > 0$ ensures that $-(\tau-\mu)/2$ is included in the integral region (μ, ∞). Then, under the null, we have

$$\partial P_A/\partial s = (\tau-\mu)(1/4)(2\pi)^{-1/2}\exp(-(1/8)(\tau-\mu)^2)$$
$$= (\tau-\mu)(1/4)\phi((\tau-\mu)/2).$$

We also have that

$$\partial P_B/\partial s$$
$$= \int_{-\infty}^{\mu} (\partial\Psi(u)/\partial s)\ \phi(u)\phi(\Psi(u))du$$
$$= \int_{-\infty}^{\mu}\phi(u)\phi((1-\rho^2)^{-1/2}(\tau s- \rho u))du$$
$$(1-\rho^2)^{-1/2}\tau$$
$$= (2\pi)^{-1}(1-\rho^2)^{-1/2}\tau\int_{-\infty}^{\mu}$$
$$\exp(-(1/2)(1-\rho^2)^{-1}$$
$$[(u-\tau\rho)^2+(1-\rho^2)\tau^2])du$$
$$= \tau(2\pi)^{-1/2}\exp(-(1/2)\tau^2)$$
$$(2\pi)^{-1/2} (1-\rho^2)^{-1/2}\int_{-\infty}^{\mu}$$
$$\exp((-(1/2)(1-\rho^2)^{-1}(u-\tau\rho)^2)du.$$

Because

$$(2\pi)^{-1/2} (1-\rho^2)^{-1/2}\int_{-\infty}^{\mu}$$
$$\exp((-(1/2)(1-\rho^2)^{-1}(u-\tau\rho)^2)du$$

converges to 1 as ρ approaches -1, we have, under the null,

$$\partial P_B/\partial s = \phi(\tau)\tau.$$

We can easily see that

$$\partial P_C/\partial s = -\phi(\tau)\tau - (\tau -\mu))(1/4)\phi((\tau -\mu)/2)$$

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

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

The derivatives of P_A with respect to μ is expressed as

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

$$\int_\mu^\infty \phi(\vartheta(u))\phi(u)\partial\vartheta(u)/\partial\mu du$$
$$= -\phi(\mu)\Phi(\vartheta(\mu))$$
$$-(1-\rho^2)^{-1/2}\int_\mu^\infty \phi(\vartheta(u))\phi(u)du$$
$$= -\Phi(\vartheta(\mu))\phi(\mu)$$
$$-(2\pi)^{-1/2}[2(1-\rho)]^{-1/2}2^{1/2}$$
$$\exp(-(1/4)(1-\rho)^{-1}(\tau-\mu)^2)$$
$$(2\pi)^{-1/2}\omega(\rho)^{-1/2}\int_\mu^\infty \exp((-1/2)\omega(\rho)^{-1}$$
$$[u+(\tau-\mu)/2]^2)du,$$

where $\omega(\rho)=(1+\rho)/2$. Note that the second term is 0, because, as ρ approaches -1, $\omega(\rho)$ converges to 0 and hence the density

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

has the whole mass at $-(\tau-\mu)/2$, which is not included in the integral region (μ, ∞). Also note that $\Phi(\vartheta(\mu))$ converges to 1 as ρ approaches -1, because $\vartheta(\mu)=(1-\rho^2)^{-1/2}(\tau+\mu)$ increases infinitely. Then we have that

$$\partial P_A/\partial\mu = -\phi(\mu).$$

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 $\sigma=1$ 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 have that

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

because, for example,

$$(\partial P_A/\partial\mu)(\partial\mu/\partial\alpha) = (-1)\partial P_A/\partial\mu,$$

from

$$\mu = -\alpha - \beta'x \quad\text{and}\quad \tau = -\gamma + (\beta-\Delta)'x.$$

Analogously, we also have that

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

(3) Derivative with respect to ρ

It follows that

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

from

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

Then, after some algebra, we have that

$$\partial P_A/\partial\rho =$$
$$(2\pi)^{-1}(1-\rho^2)^{-1/2}C\int_\mu^\infty A(u)\exp(B(u))du,$$

where

$$A(u) = (1-\rho^2)^{-1}[\rho(-\mu+\tau) + (\rho-1)u],$$
$$B(u) = -(1-\rho^2)^{-1}(1-\rho)[u + (\tau-\mu)/2]^2,$$
$$C = \exp((-1/4)(1-\rho^2)^{-1}(1+\rho)(\tau-\mu)^2).$$

This expression can be written as

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

using the equality

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

The first integral in the bracket is expressed as

$$-(1/2)(2\pi)^{-1}(1-\rho^2)^{-1/2}\exp(B(\mu)),$$

by means of integration by parts, where $B(\mu) = -(1-\rho^2)^{-1}(1-\rho)(\tau+\mu)^2/4$; the value of the second integral is negligible in comparison with the value of the first integral, because we have that

$$|(1-\rho^2)^{-1}\int_\mu^\infty(1+\rho)\exp(B(u))du$$
$$/\int_\mu^\infty B'(u)\exp(B(u))du\,|$$
$$< (1-\rho^2)^{-1}(1+\rho)\int_\mu^\infty \exp(B(u))du$$
$$/|B'(\mu)|\int_\mu^\infty \exp(B(u))du$$
$$= (1-\rho^2)^{-1}(1+\rho)/[(1-\rho^2)^{-1}(1-\rho)(\tau+\mu)] \to 0,$$

as ρ approaches -1. Thus, we have the expression

$$\partial P_A/\partial\rho$$
$$= -(1/2)(2\pi)^{-1}(1-\rho^2)^{-1/2}\exp((-1/2)(1-\rho^2)^{-1}(\mu^2$$
$$+ \tau^2 - 2\rho\tau\mu))(1 + o(1)),$$

as ρ approaches -1. From the symmetry of τ and μ in this expression we also have

$$\partial P_C/\partial\rho = \partial P_A/\partial\rho.$$

Table 1: Empirical Sizes and Powers of the Test for Nominal Sizes 0.10 and 0.05.

Parameter Values				
ρ	σ	$\beta + \eta$	$T>\chi^2_{0.90}$	$T>\chi^2_{0.95}$
Null				
-1.0	1.0	0.0	0.1305	0.0665
Alter-native				
-1.0	1.0	**0.1**	0.2430	0.1465
-1.0	1.0	**0.2**	0.6200	0.4900
-1.0	1.0	**0.3**	0.8305	0.7450
-1.0	**0.8**	0.0	0.5280	0.4000
-1.0	**1.4**	0.0	0.8030	0.6965
-0.5	1.0	0.0	0.1655	0.0910
0.0	1.0	0.0	0.2470	0.1680
0.5	1.0	0.0	0.6255	0.5210

Note: The number of iterations is 2,000 and the sample size is 400.

Fig. 1: Disturbance Distribution and Choice of Alternative