Marginal and Interaction Effects in Ordered Response Models

Mallick, Debdulal

School of Accounting, Economics and Finance, Deakin University, Burwood, Victoria, Australia

July 2008

Marginal and Interaction Effects in Ordered Response Models

Debdulal Mallick*

July 2008

*The author would like to acknowledge the helpful comments and assistance offered by Robert Chirinko, Chris Doucouliagos, Randy Silvers and Mehmet Ulubasoglu. All errors and omissions are the sole responsibility of the author.

Marginal and Interaction Effects in Ordered Response Models

Abstract

In discrete choice models the marginal effect of a variable of interest that is interacted with another variable differs from the marginal effect of a variable that is not interacted with any variable. The magnitude of the interaction effect is also not equal to the marginal effect of the interaction term. I present consistent estimators of both marginal and interaction effects in ordered response models. This procedure is general and can easily be extended to other discrete choice models.

Keywords: Marginal effect, interaction effect, ordered probit

JEL Classification: C12, C25

Correspondence:

Debdulal Mallick
School of Accounting, Economics and Finance
Deakin University
221 Burwood Hwy, Burwood
Victoria 3125, Australia
Phone: (+ 61 3) 92517808
Fax: (+ 61 3) 92446283
Email: dmallic@deakin.edu.au

1. Introduction

Marginal and interaction effects of variables are of immense interest in applied economics and other branches of social sciences. Inference on interaction terms in non-linear models is different from that in linear models. This difference is particularly evident in the estimation of discrete choice models. Standard software (such as STATA® 10) incorrectly estimates the magnitude and standard error of the interaction term in nonlinear models. Ai and Norton (2003, p. 123; this journal) reviewed 13 economics journals listed on JSTOR and found that none of the 72 articles published between 1980 and 1999 that used interaction terms in nonlinear models interpreted the coefficient correctly. They also presented consistent estimators of the magnitude and standard error of the interaction effect in logit and probit models.

This paper extends Ai and Norton (2003) in two ways. First, we consider ordered choice models and provide consistent estimators of the magnitude of the interaction term and its standard error. The approach here is similar to that in Ai and Norton. The second extension is more fundamental; it shows that the marginal effects of the variables that are interacted are different from the marginal effects of the variables that are not interacted in that the former also involves the coefficient of the interaction term. For example, suppose three independent variables, x_1, x_2 and x_3 appear in an ordered probit (logit) model, and x_2 and x_3 are interacted (i.e. $x_2 * x_3$ is included as an additional independent variable). The formula for the marginal effect of x_2 (or x_3) will be different from that of x_1 because of the interaction effect. Standard software does not also account for this effect and therefore incorrectly estimates the marginal effect and standard error of x_2 (and x_3). The second extension also applies to other discrete choice models.

2. Estimation

Suppose, we have the following regression: $y^* = \beta'\mathbf{x} + \varepsilon$, where y* is the dependent variable but is unobserved. What is observed is the respondent's answer y which is related to y* as:

$$y = j \quad \text{if} \quad \kappa_{j\text{-}1} < y^* \leq \kappa_j, \qquad \text{--- (1)}$$

where j = 1, 2, …. J are the responses that are ordered in nature, and κ's are (J -1) unknown parameters known as *cut points* or *threshold parameters*. An example can be the responses when people are asked about their happiness. Assume, for simplicity and without loss of generality, that there are only three covariates (x_1, x_2 and x_3) in the **x** vector, and all are continuous. Only x_2 and x_3 are interacted while x_1 is not; therefore, $\beta'\mathbf{x} = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{23}(x_2 * x_3)$. The κ's and $\beta' = (\beta_1, \beta_2, \beta_3, \beta_{23})$ are jointly estimated by the Maximum Likelihood method.

Assuming ε ~ N(0,1), the probability for the j-th outcome is given by

$$\text{Prob}(\text{y} = j) = \Phi(\kappa_j - \beta'\mathbf{x}) - \Phi(\kappa_{j-1} - \beta'\mathbf{x}) \qquad \text{--- (2)}$$

where Φ is the cumulative standard normal (or logistic) distribution, which is continuous and twice differentiable.

2.1 Marginal effect

The marginal effect of x_1 for the j-th response is given by

$$\delta_{1,j} = \frac{\partial \text{Prob}[\text{y} = j \mid \mathbf{x}]}{\partial x_1} = \left[\phi\left(\kappa_{j-1} - \beta'\mathbf{x}\right) - \phi\left(\kappa_j - \beta'\mathbf{x}\right)\right]\beta_1 = \left[\phi_j(\cdot) - \phi_{j-1}(\cdot)\right]\beta_1, \qquad \text{---(3)}$$

where $\phi(.)$ is the standard normal (logistic) density function. It determines how a change in x_1 changes the distribution of the outcome variable, i.e. all outcome probabilities (Boes and Winkelmann, 2006, p. 169).[1]

However, the marginal effect of x_2 for the j-th response will be different from that in equation (3) and is given by

---

[1] If x_1 is a dummy variable such as gender then the marginal effect is computed as $\Delta\text{Prob}[\text{y} = j \mid \mathbf{x}] = \text{Prob}[\text{y} = j \mid \mathbf{x} + \Delta x_1] - \text{Prob}[\text{y} = j \mid \mathbf{x}]$.

$$\delta_{2,j} = \frac{\partial \text{Prob}[\text{y} = j|\mathbf{x}]}{\partial x_2} = \phi_{j-1}(\cdot)\left[\beta_2 + \beta_{23}x_3\right] - \phi_j(\cdot)\left[\beta_2 + \beta_{23}x_3\right]. \qquad \text{--- (4)}$$

We obtain a similar expression for the marginal effect of x_3. The difference between the formulas in equations (3) and (4) is that the marginal effect of x_1 in equation (3) is zero if the coefficient on x_1 (β_1) is zero, whereas the marginal effect of x_2 (or x_3) may be nonzero even if its coefficient is zero. This arises because the latter depends not only on x_2 but also on the combined effect of x_2 and x_3. To obtain the correct marginal effect of x_2 (or x_3), the formula in equation (4) must be estimated. However, standard software estimates equation (3) to obtain marginal effects of all variables entering the model, which is clearly wrong.

2.2 Interaction effect

The magnitude of the interaction effect for the *j*-th response is obtained by computing the cross derivative of equation (2) or partial derivative of equation (4) with respect to x_3:

$$\delta_{23,j} = \frac{\partial^2 \text{Prob}[\text{y} = j|\mathbf{x}]}{\partial x_2 \partial x_3} = \left[\phi_{j-1}(\cdot) - \phi_j(\cdot)\right]\beta_{23} - \left[\beta_2 + \beta_{23}x_3\right]\left[\beta_3 + \beta_{23}x_2\right]\left[\phi'_{j-1}(\cdot) - \phi'_j(\cdot)\right], \qquad \text{--- (5)}$$

where $\phi'(\cdot)$ is the first derivative of the density function with respect to its argument. The right hand side of equation (5) shows that, even if the coefficient on the interaction term, β_{23}, is zero, the magnitude of the interaction effect can be nonzero because it also depends on the individual coefficients on both x_2 and x_3. Again, standard software estimates the marginal effect of the interaction term,

$$\frac{\partial \text{Prob}[\text{y} = j|\mathbf{x}]}{\partial (x_2 * x_3)} = \left[\phi_j(\cdot) - \phi_{j-1}(\cdot)\right]\beta_{23} \qquad \text{----(6)}$$

which is different from the expression in equation (5). For a linear regression, these two terms will be the same.

To show the asymptomatic properties of the marginal and interaction effects, rewrite equation (2) as $\text{Prob}(y = j) = F_j(x, \beta)$. Then the estimated values of marginal effects of x_1 and x_2, and the interaction effect of x_2 and x_3 can be computed respectively as

$$\hat{\delta}_{1,j} = \frac{\partial F_j(x, \hat{\beta})}{\partial x_1}, \qquad \text{--- (7)}$$

$$\hat{\delta}_{2,j} = \frac{\partial F_j(x, \hat{\beta})}{\partial x_2}, \qquad \text{--- (8)}$$

$$\hat{\delta}_{23,j} = \frac{\partial^2 F_j(x, \hat{\beta})}{\partial x_2 \partial x_3}, \qquad \text{--- (9)}$$

where $\hat{\beta}$ is consistent estimator of β computed by the Maximum Likelihood. The consistencies of $\hat{\delta}_{1,j}$, $\hat{\delta}_{2,j}$ and $\hat{\delta}_{23,j}$ are ensured by the continuity of F_j and the consistency of $\hat{\beta}$. The asymptotic variances of $\hat{\delta}_{1,j}$, $\hat{\delta}_{2,j}$ and $\hat{\delta}_{23,j}$ are consistently estimated by the "delta method",[2]

$$\hat{\sigma}^2_{1,j} = \frac{\partial}{\partial \beta'}\left[\frac{\partial F_j(x, \hat{\beta})}{\partial x_1}\right]\hat{\Omega}_\beta \frac{\partial}{\partial \beta}\left[\frac{\partial F_j(x, \hat{\beta})}{\partial x_1}\right], \qquad \text{--- (10)}$$

$$\hat{\sigma}^2_{2,j} = \frac{\partial}{\partial \beta'}\left[\frac{\partial F_j(x, \hat{\beta})}{\partial x_2}\right]\hat{\Omega}_\beta \frac{\partial}{\partial \beta}\left[\frac{\partial F_j(x, \hat{\beta})}{\partial x_2}\right], \qquad \text{--- (11)}$$

[2] "Delta method" estimates the variance using a first-order Taylor approximation. It may provide poor approximation in non-linear functions. In such cases, a second-order Taylor approximation is suggested, and normal distribution is then replaced by a chi-square distribution. For details, see Spanos (1999, p. 493-494).

$$\text{and } \hat{\sigma}^2_{23,j} = \frac{\partial}{\partial \beta'}\left[\frac{\partial^2 \mathrm{F}_j(\mathrm{x}, \hat{\beta})}{\partial x_2 \partial x_3}\right]\hat{\Omega}_\beta \frac{\partial}{\partial \beta}\left[\frac{\partial^2 \mathrm{F}_j(\mathrm{x}, \hat{\beta})}{\partial x_2 \partial x_3}\right] \quad \text{--- (12)}$$

respectively, where $\hat{\Omega}_\beta$ is consistent covariance estimator of $\hat{\beta}$, and $\hat{\delta}_{m,j} \sim N(\delta_{m,j}, \sigma^2_{m,j})$, $\forall$ m = 1, 2, and 23, and j = 1, 2, …. J. The corresponding t-statistics are $\hat{\delta}_{1,j} / \hat{\sigma}_{1,j}$, $\hat{\delta}_{2,j} / \hat{\sigma}_{2,j}$ and $\hat{\delta}_{23,j} / \hat{\sigma}_{23,j}$ respectively. Under some regularity conditions, these t-statistics have standard normal distributions. Individual hypothesis that marginal or interaction effect is zero can be tested using these t-statistics.

It is important to note that the formulas for the variances in equations (10-12) involve square of the derivative of the coefficients estimated in equations (7-9). Therefore, the variances will be estimated with larger error than the corresponding marginal effects of the interacted variables if they are estimated using the incorrect formula.

The marginal and interaction effects have different signs for different observations, but for the present purpose this issue can be avoided by assuming that the effects are evaluated at the mean value of **x**. Ai and Norton (2003) provide an elegant discussion on this issue.

3. An example

In the following we estimate an ordered probit model using household and village level survey data on food security in Bangladesh. Based on food production, availability, purchasing power and access to common resources, the respondents were asked to define the food security status of their households in any of the four categories—severe (chronic) food shortage, occasional (transitory) food shortage, breakeven, and food surplus. The independent variables are i) amount of land cultivated in decimal (LAND), ii) percentage of household members engaged in income generating activities (IGA), iii) village level physical infrastructure calculated from several other variables using principal component analysis[3] (INFRA), and iv) interaction of the first two variables (LAND*IGA). Both the correct and incorrect marginal and interaction effects and their standard errors are reported in Table 1. For simplicity, we report the statistics only for transitory food insecurity category.

Table 1: Marginal and interaction effects for the transitory food insecurity category (Dependent variable: 1 = chronic food insecurity, 2 = transitory food insecurity, 3 = breakeven, and 4 = food surplus)

Independent variables	Coefficient	Marginal effect	
		Incorrect	Correct
LAND	1.222 (0.509)	0.009 (0.011) [a]	0.010 (0.004) [b]
IGA	16.708 (15.945)	0.132 (0.198) [a]	0.138 (0.132) [b]
INFRA [e]	0.034 (0.0234)	0.0002 (0.0003)	0.0002 (0.0003)
		Magnitude of the interaction effect	
		Incorrect	Correct
LAND*IGA	0.0002 (0.0009)	0.000 (0.0001) [c]	8.105 (9.961) [d]
Sample size = 2517			

Figures in the parentheses are robust standard errors.

a. using the incorrect formula in equation (3), b. using the correct formula in equation (4), c. using the incorrect formula in equation (6), d. using the correct formula in equation (5), e. using the correct formula in equation (3) because INFRA variable is not interacted.

We see from the results that magnitudes of the marginal effects of the variables that are interacted are slightly different for correct and incorrect formulas. However, the correct standard errors for Land and IGA are about 64% and 33% lower than those estimated with incorrect formula. The magnitude of the interaction effect is measured with very large error. It is close to zero when estimated using the wrong formula, while it is around eight when estimated using the correct formula.

[3] These variables are distance of the village from the Thana (lowest administrative unit) headquarters, nearest bazaar, all weather pucca road and bus stand. Since longer distance implies poor infrastructure, we first calculate the reciprocal of each variable and then use principal component analysis to calculate a score.

4. Conclusion

The marginal effect in discrete choice models is complicated especially when variables are interacted. I present a consistent estimator of the marginal effect of a variable that is interacted with another variable in ordered response models. This estimator differs from the marginal effect of a variable that is not interacted. Standard software incorrectly estimates the latter marginal effect for an interacted variable. A consistent estimator of the interaction effect is also presented. The procedure is general and can easily be extended to other discrete choice models.

Reference

Ai, C., Norton, E.C., 2003. Interaction Terms in Logit and Probit Models. Economics Letters 80 (1, July), 123-129.

Boes, S., Winkelmann, R., 2006. Ordered Response Models. Allgemeines Statistisches Archiv 90 (1, March), 167-181.

Spanos, Aris, 1999. Probability Theory and Statistical Inference: Econometric Modeling with Observational Data. Cambridge University Press, New York.