The Estimation of Interaction Effects in Probit Models with more than one Interaction Term

Richard Seymour[a]

[a]*School of Economics and Finance, Curtin University, GPO Box U1987, Perth, 6845 Western Australia*
Email: Richard.seymour@cbs.curtin.edu.au

Abstract: This paper examines the method developed by Norton, Wang and Ai (2004) to compute the interaction effects, and their standard errors, in probit and logit models. It finds that their method is not suitable for probit or logit models that contain multiple interaction terms that include the same variable. The contribution of the paper is that it develops a set of matrix algebra formulas that will estimate the interaction effects in probit models that contain multiple interaction terms that include the same variable. It also develops a set of matrix algebra formulas that will allow the standard errors for the interaction effects to be derived using the Delta method.

Keywords: *Interaction effect, Interaction term, Probit, Nonlinear models*.

1. INTRODUCTION

Ai and Norton (2003) found that despite the common use of interaction terms, most applied researchers incorrectly interpret the coefficient on the interaction terms in nonlinear models. They conducted a review of 13 economic journals listed in JSTOR (www.jstor.org) and identified 72 articles published between 1980 and 1999 that used interaction terms in nonlinear models. Of these 72 articles, they found that none had correctly interpreted the coefficient on the interaction term.

Norton, Wang and Ai (2004) note that most applied economists estimated the marginal effect of an interaction term instead of the interaction effect of an interaction term. To facilitate the correct computation of interaction effects for interaction terms in probit and logit models, they developed three formulas to compute the interaction effects for interaction terms based on: 1) the interaction of two continuous variables; 2) the interaction of a continuous variable and a discrete variable; or 3) the interaction of two discrete variables. The formulas are based on a probit or logit model that only contains one interaction term. To derive the standard error for the interaction effect in probit and logit models, they use the Delta method.

One limitation of the three formulas developed by Norton, Wang and Ai (2004) is that they do not accommodate the computation of interaction effects when there is more than one interaction term with the same independent variable. In addition, they do not accommodate the computation of the standard error for the interaction effect when there is more than one interaction term with the same independent variable; this is because the derivatives of some of the coefficients with respect to the interaction effect cannot be computed correctly.

2. THE ESTIMATION OF INTERACTION EFFECTS IN PROBIT MODELS WHERE MULTIPLE INTERACTION TERMS INCLUDE THE SAME VARIABLE

This section of the paper develops a set of matrix algebra formulas that will overcome the limitations of the formulas developed by Norton, Wang and Ai (2004). All the propositions in this section are based on the following probit model:

$$P[y = 1 \mid x] = \Phi(x'\beta + x'Ax) \tag{1}$$

where Φ is the standard normal cumulative distribution, $x = \begin{bmatrix} x_1 & \ldots & x_n \end{bmatrix}$ is a vector of independent variables, without any interaction terms, $\beta = \begin{bmatrix} \beta_1 & \ldots & \beta_n \end{bmatrix}$ is a vector of coefficients for the variables in x, and $A = \begin{bmatrix} 0 & \beta_{12} & \beta_{1n} \\ \beta_{21} & 0 & \beta_{2n} \\ \beta_{n1} & \beta_{n2} & 0 \end{bmatrix}$ is matrix of coefficients for the interaction terms in the model.

In the A matrix, $\beta_{ij} = \beta_{ji}$, where $i \neq j$. It is necessary to set β_{ij} equal to β_{ji} in order to correctly derive the formulas for the interaction effects. However, this means that the A matrix must be multiplied by ½ when it is used in the index for the standard normal cumulative distribution function, Φ. Based on the structure of the interaction term coefficients matrix, A, the $x'Ax$ term in 1 is not a quadratic. Consequently, the derivate of the $x'Ax$ term with respect to x will be equal to Ax, and not $2Ax$.

The constant coefficient value is also contained in the variable coefficients vector, β, with the corresponding variable value set to 1 in the variable vector, x. The position of the constant coefficient value in the variable coefficients vector, β, and 1 in the variable vector, x, depend on the position of the

constant in the probit model. The position of the variable values in the variable vector, x, and the variable coefficient values in the variable coefficients vector, β, are also based on their positions within the probit model. The position of the interaction term coefficients in the interaction term coefficients matrix, A, are based on the position in the model of the two variables used to form the interaction term. The row of the interaction term coefficient is determined by the position in the model of the first variable in the interaction term, and the column of the interaction term coefficient is determined by the position in the model of the second variable in the interaction term.

Proposition A.1 (*The Computation of Continuous * Continuous Interaction Effects*)

Let y be a random variable which follows the probability law as defined in 1, then the interaction effects of continuous variables in the x vector is:

$$\frac{\partial \Phi(u)}{\partial x \partial x'} = \Phi'(x'\beta + x'\tfrac{1}{2}Ax)A + \Phi''(x'\beta + x'\tfrac{1}{2}Ax)[\beta + Ax][\beta' + x'A]$$

The preceding formula computes the interaction effects for interaction terms based on the interaction of two continuous variables, where $u = x'\beta + x'Ax$, Φ' is the first derivative of the standard normal cumulative distribution, and Φ'' is the second derivative of the standard normal cumulative distribution. The computation of the formula results in a $n \times n$ matrix. It should be noted that only the interaction effects for continuous * continuous interaction terms specified in the probit model will be valid.

Proposition A.2 (*The Computation of the Derivatives of the Independent Variable Coefficients in the β Vector with respect to the Continuous * Continuous Interaction Effects*)

The derivatives of the independent variable coefficients in the β vector with respect to the interaction effects of continuous variables from proposition A.1 is:

$$\frac{\partial vec\left(\frac{\partial \Phi(u)}{\partial x \partial x'}\right)}{\partial \beta'} = \Phi''(x'\beta + x'\tfrac{1}{2}Ax)vec(A)x' +$$

$$\Phi'''(x'\beta + x'\tfrac{1}{2}Ax)\left((I_n \otimes [\beta + Ax])vec([\beta' + x'A])\right)x' +$$

$$\Phi''(x'\beta + x'\tfrac{1}{2}Ax)\left((I_n \otimes [\beta + Ax])I_n\right) +$$

$$\Phi''(x'\beta + x'\tfrac{1}{2}Ax)\left(([\beta' + x'A]' \otimes I_n)I_n\right)$$

The preceding formula computes the derivatives of each of the variable coefficients in the β vector with respect to each of the continuous * continuous interaction effects, where $u = x'\beta + x'Ax$, Φ'' is the second derivative of the standard normal cumulative distribution, and Φ''' is the third derivative of the standard normal cumulative distribution. The computation of the formula results in a $n \times n$ matrix. It should be noted that the results of the formula will need to be unveced. It should also be noted that only the derivatives of each of the independent variable coefficients in the β vector with respect to each of the interaction effects for continuous * continuous interaction terms specified in the probit model will be valid.

Proposition A.3 (*The Computation of the Derivatives of the Interaction Term Coefficients in the Matrix, A, with respect to the Continuous * Continuous Interaction Effects*)

The derivatives of the interaction term coefficients in the A matrix with respect to the interaction effects of continuous variables from proposition A.1 is:

$$\frac{\partial vec\left(\frac{\partial \Phi(u)}{\partial x \partial x'}\right)}{\partial (vec(A))'} = \Phi'(x'\beta + x' \tfrac{1}{2} Ax)I_{n^2} +$$

$$\Phi''(x'\beta + x' \tfrac{1}{2} Ax)vec(A)(vec(xx'))' +$$
$$\Phi'''(x'\beta + x' \tfrac{1}{2} Ax)vec([\beta + Ax][\beta' + x'A])(vec(xx'))' +$$
$$\Phi''(x'\beta + x' \tfrac{1}{2} Ax)((I_n \otimes [\beta + Ax])((I_n \otimes x')I_{n^2})) +$$
$$\Phi''(x'\beta + x' \tfrac{1}{2} Ax)(([\beta' + x'A]' \otimes I_n)((x' \otimes I_n)I_{n^2}))$$

The preceding formula computes the derivatives of each of the interaction term coefficients in the interaction term coefficients matrix, A, with respect to each of the continuous * continuous interaction effects, where $u = x'\beta + x'Ax$, Φ' is the first derivative of the standard normal cumulative distribution, Φ'' is the second derivative of the standard normal cumulative distribution, and Φ''' is the third derivative of the standard normal cumulative distribution. The computation of the formula results in a $n^2 \times n^2$ matrix. It should be noted that the results of the formula will need to be unveced. It should also be noted that only the derivatives of each of the interaction term coefficients in the A matrix with respect to each of the interaction effects for continuous * continuous interaction terms specified in the probit model will be valid.

Proposition B.1 (*The Computation of Continuous * Discrete Interaction Effects*)

Let y be a random variable which follows the probability law as defined in 1, then the interaction effects of a discrete variable, x_i, and the continuous variables in the x vector is:

$$\frac{\Delta \frac{\partial \Phi(u)}{\partial x}}{\Delta x_i} = \left(\Phi'(x'\beta + x' \tfrac{1}{2} Ax)[\beta + Ax]\right) - \left(\Phi'(x0'\beta + x0' \tfrac{1}{2} Ax0)[\beta + Ax0]\right)$$

The preceding formula computes the interaction effects for interaction terms that are based on the interaction of a continuous and a discrete variable, where $u = x'\beta + x'Ax$, x_i is a discrete independent variable, Φ' is the first derivative of the standard normal cumulative distribution, $x = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}$ is a vector of independent variables with the discrete variable $x_i = 1$, and $x0 = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}$ is a vector of independent variables with the discrete variable $x_i = 0$. The computation of the formula results in a $n \times 1$ vector. It should be noted that only the interaction effects for continuous * discrete (x_i) interaction terms specified in the probit model will be valid.

Proposition B.2 (*The Computation of the Derivatives of the Variable Coefficients in the β Vector with respect to the Continuous * Discrete Interaction Effects*)

The derivatives of the independent variable coefficients in the β vector with respect to the interaction effects of the discrete variable, x_i, and the continuous variables from proposition B.1 is:

$$\frac{\partial\left(\dfrac{\Delta\dfrac{\partial\Phi(u)}{\partial x}}{\Delta x_i}\right)}{\partial\beta'} = \left(\Phi''(x'\beta + x'\tfrac{1}{2}Ax)[\beta + Ax]x' + \Phi'(x'\beta + x'\tfrac{1}{2}Ax)I_n\right) -$$
$$\left(\Phi''(x0'\beta + x0'\tfrac{1}{2}Ax0)[\beta + Ax0]x0' + \Phi'(x0'\beta + x0'\tfrac{1}{2}Ax0)I_n\right)$$

The preceding formula computes the derivatives of each of the variable coefficients in the β vector with respect to each of the continuous * discrete (x_i) interaction effects, where $u = x'\beta + x'Ax$, x_i is a discrete independent variable used in the model, Φ' is the first derivative of the standard normal cumulative distribution and Φ'' is the second derivative of the standard normal cumulative distribution. The computation of the formula results in a $n \times n$ matrix. It should be noted that only the derivatives of each of the independent variable coefficients in the β vector with respect to each of the interaction effects for continuous * discrete x_i interaction terms specified in the probit model will be valid.

Proposition B.3 (*The Computation of the Derivatives of the Interaction Term Coefficients in the Matrix, A, with respect to the Continuous * Discrete Interaction Effects*)

The derivatives of the interaction term coefficients in the *A* matrix with respect to the interaction effects of the discrete variable, x_i, and the continuous variables from proposition B.1 is:

$$\frac{\partial vec\left(\dfrac{\Delta\dfrac{\partial\Phi(u)}{\partial x}}{\Delta x_i}\right)}{\partial(vec(A))'} = \left(\Phi''(x'\beta + x'\tfrac{1}{2}Ax)vec([\beta + Ax])vec(xx')' +\right.$$
$$\Phi'(x'\beta + x'\tfrac{1}{2}Ax)\left((x'\otimes I_n)I_{n^2}\right)\right) -$$
$$\left(\Phi''(x'\beta + x0'\tfrac{1}{2}Ax0)vec([\beta + Ax0])vec(x0x0')' +\right.$$
$$\Phi'(x0'\beta + x0'\tfrac{1}{2}Ax0)\left((x0'\otimes I_n)I_{n^2}\right)\right)$$

The preceding formula computes the derivatives of each of the interaction term coefficients in the A matrix with respect to each of the continuous * discrete (x_i) interaction effects, where $u = x'\beta + x'Ax$, x_i is a discrete independent variable used in the model, Φ' is the first derivative of the standard normal cumulative distribution and Φ'' is the second derivative of the standard normal cumulative distribution. The computation of the formula results in a $n^2 \times n$ matrix. It should be noted that the results of the formula will need to be unveced. It should also be noted that only the derivatives of each of the interaction term coefficients in the A matrix with respect to each of the interaction effects for discrete x_i * continuous interaction terms specified in the probit model will be valid.

Proposition C.1 (*The Computation of Discrete * Discrete Interaction Effects*)

Let *y* be a random variable which follows the probability law as defined in 1, then the interaction effects of the discrete variables, x_i and x_j in the *x* vector is:

$$\frac{\Delta^2\Phi(u)}{\Delta x_i \Delta x_j} = \left(\Phi(j'\beta + j'\tfrac{1}{2}Aj) - \Phi(k'\beta + k'\tfrac{1}{2}Ak)\right) -$$
$$\left(\Phi(l'\beta + l'\tfrac{1}{2}Al) - \Phi(m'\beta + m'\tfrac{1}{2}Am)\right)$$

The preceding formula computes the interaction effects for interaction terms based on the interaction of two discrete variables, where $u = x'\beta + x'Ax$, x_i and x_j are discrete independent variables used in the model, Φ is the standard normal cumulative distribution, $j = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}$ is a vector of independent variables with the discrete variables $x_i = 1, x_j = 1$; $k = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}$ is a vector of independent variables with the discrete variables $x_i = 0, x_j = 1$; $l = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}$ is a vector of independent variables with the discrete variables $x_i = 1, x_j = 0$; and $m = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}$ is a vector of independent variables with the discrete variables $x_i = 0, x_j = 0$. The computation of the formula results in a scalar.

Proposition C.2 (*The Computation of the Derivatives of the Variable Coefficients in the Vector, β, with respect to the Discrete * Discrete Interaction Effects*)

The derivatives of the independent variable coefficients in the β vector with respect to the interaction effect of the discrete variables from proposition C.1 is:

$$\frac{\partial \left(\dfrac{\Delta^2\Phi(u)}{\Delta x_i \Delta x_j} \right)}{\partial \beta} = \left(\Phi(j'\beta + j'\tfrac{1}{2}Aj)j - \Phi(k'\beta + k'\tfrac{1}{2}Ak)k\right) -$$
$$\left(\Phi(l'\beta + l'\tfrac{1}{2}Al)l - \Phi(m'\beta + m'\tfrac{1}{2}Am)m\right)$$

The preceding formula computes the derivatives of each of the variable coefficients in the β vector with respect to the discrete * discrete interaction effect, where $u = x'\beta + x'Ax$, x_i and x_j are discrete independent variables used in the model, Φ is the standard normal cumulative distribution, $j = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}$ is a vector of independent variables with the discrete variables $x_i = 1, x_j = 1$; $k = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}$ is a vector of independent variables with the discrete variables $x_i = 0, x_j = 1$; $l = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}$ is a vector of independent variables with the discrete variables $x_i = 1, x_j = 0$; and $m = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}$ is a vector of independent variables with the discrete variables $x_i = 0, x_j = 0$. The computation of the formula results in a $n \times 1$ vector.

Proposition C.3 (*The Computation of the Derivatives of the Interaction Term Coefficients in the Matrix, A, with respect to the Discrete * Discrete Interaction Effects*)

The derivatives of the interaction term coefficients in the A matrix with respect to the interaction effect of the discrete variables from proposition C.1 is:

$$\frac{\partial\left(\dfrac{\Delta^2\Phi(u)}{\Delta x_i \Delta x_j}\right)}{\partial A} = \left(\Phi(j'\beta + j'\,\tfrac{1}{2}\,Aj)jj' - \Phi(k'\beta + k'\,\tfrac{1}{2}\,Ak)kk'\right) - \left(\Phi(l'\beta + l'\,\tfrac{1}{2}\,Al)ll' - \Phi(m'\beta + m'\,\tfrac{1}{2}\,Am)mm'\right)$$

The preceding formula computes the derivatives of each of the interaction term coefficients in the interaction term coefficients matrix, A, with respect to the discrete * discrete interaction effect, where $u = x'\beta + x'Ax$, x_i and x_j are discrete independent variables used in the model, Φ is the standard normal cumulative distribution, $j = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}$ is a vector of independent variables with the discrete variables $x_i = 1, x_j = 1$; $k = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}$ is a vector of independent variables with the discrete variables $x_i = 0, x_j = 1$; $l = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}$ is a vector of independent variables with the discrete variables $x_i = 1, x_j = 0$; and $m = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}$ is a vector of independent variables with the discrete variables $x_i = 0, x_j = 0$. The computation of the formula results in a $n \times n$ matrix.

3. CONCLUSION

In conclusion, the paper has found that the formulas and method developed by Norton, Wang and Ai (2004) to compute the interaction effects, and their standard errors, in probit and logit models are not suitable for computing the interaction effects and standard errors for probit or logit models which contain multiple interaction terms that include the same variable. To overcome the limitation of the formulas and method developed by Norton, Wang and Ai (2004), the paper develops a set of matrix algebra formulas that will compute the interaction effects for probit models that contain multiple interaction terms, including where two or more interaction terms include the same variable. It also develops a set of matrix algebra formulas that compute the derivatives of each of the coefficients with respect to the interaction effects. The matrix algebra formulas developed in the paper are divided into three subsets. The first subset of formulas compute the interaction effects for interaction terms based on the interaction of two continuous variables; they also compute the derivatives of each of the coefficients with respect to the continuous * continuous interaction effects. The second subset of formulas compute the interaction effects for interaction terms based on the interaction of a continuous variable and discrete variable; and also the derivatives of each of the coefficients with respect to the continuous * discrete interaction effects. The final subset of formulas compute the interaction effects for interaction terms based on the interaction of two discrete variables, as well as, the derivatives of each of the coefficients with respect to the discrete * discrete interaction effects.

References

Ai, C., and E.C. Norton. 2003. Interaction Terms in Logit and Probit Models. *Economic Letters*. 80: 123-129.

Norton, E.C., H. Wang, and C. Ai. 2004. Computing Interaction Effects and Standard Errors in Logit and Probit Models. *The Stat Journal* 4 (2): 154-167.