Available online at www.sciencedirect.com
SCIENCE DIRECT
Economics Letters 80 (2003) 123–129
economics letters
www.elsevier.com/locate/econbase

Interaction terms in logit and probit models

Chunrong Ai[a], Edward C. Norton[b,*]

[a]University of Florida, Gainesville, FL, USA
[b]Department of Health Policy and Administration, University of North Carolina, CB#7411 McGarvan-Greenberg Building, Chapel Hill, NC 27599-7411, USA

Received 4 December 2002; accepted 14 January 2003

Abstract

The magnitude of the interaction effect in nonlinear models does not equal the marginal effect of the interaction term, can be of opposite sign, and its statistical significance is not calculated by standard software. We present the correct way to estimate the magnitude and standard errors of the interaction effect in nonlinear models.

Keywords: Interaction effect; Interaction term; Logit; Probit; Nonlinear models

JEL classification: C12; C25; C51

1. Introduction

Applied economists often estimate interaction terms to infer how the effect of one independent variable on the dependent variable depends on the magnitude of another independent variable. Difference-in-difference models, which measure the difference in outcome over time for the treatment group compared to the difference in outcome over time for the control group, are examples of models with interaction terms. Although interaction terms are used widely in applied econometrics, and the correct way to interpret them is known by many econometricians and statisticians, most applied researchers misinterpret the coefficient of the interaction term in nonlinear models. A review of the 13 economics journals listed on JSTOR found 72 articles published between 1980 and 1999 that used interaction terms in nonlinear models. None of the studies interpreted the coefficient on the interaction term correctly. The recent paper by DeLeire (2000) is a welcome exception.

In linear models the interpretation of the coefficient of the interaction between two variables is

*Corresponding author. Tel.: +1-919-966-8930; fax: +1-919-966-6961.
E-mail address: edward_norton@unc.edu (E.C. Norton).

0165-1765/03/$ – see front matter
doi:10.1016/S0165-1765(03)00032-6

straightforward. Let the continuous dependent variable y depend on two independent variables x_1 and x_2, their interaction, a vector of additional independent variables $\mathbf{X}$ including the constant term independent of x_1 and x_2, and βs are unknown parameters. If x_1 and x_2 are continuous, the interaction effect of the independent variables x_1 and x_2 is the cross-derivative of the expected value of y

$$\frac{\partial^2 \mathrm{E}[y|x_1, x_2, \mathbf{X}]}{\partial x_1 \partial x_2} = \beta_{12}.$$

If x_1 and x_2 are dichotomous, then the interaction effect of a change in both x_1 and x_2 from zero to one is found by taking discrete differences

$$\frac{\Delta^2 \mathrm{E}[y|x_1, x_2, \mathbf{X}]}{\Delta x_1 \Delta x_2} = \beta_{12}.$$

The statistical significance of the interaction effect can be tested with a single t-test on the coefficient β_{12}.

The intuition from linear models, however, does not extend to nonlinear models. To illustrate, consider a probit model similar to the previous example, except that the dependent variable y is a dummy variable. The conditional mean of the dependent variable is

$$\mathrm{E}[y|x_1, x_2, \mathbf{X}] = \Phi(\beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \mathbf{X}\beta) = \Phi(\cdot), \tag{1}$$

where Φ is the standard normal cumulative distribution. Suppose x_1 and x_2 are continuous. The *interaction effect* is the cross derivative of the expected value of y

$$\frac{\partial^2 \Phi(\cdot)}{\partial x_1 \partial x_2} = \beta_{12}\Phi'(\cdot) + (\beta_1 + \beta_{12}x_2)(\beta_2 + \beta_{12}x_1)\Phi''(\cdot). \tag{2}$$

However, most applied economists instead compute the marginal effect of the interaction term, which is $\partial\Phi(\cdot)/\partial(x_1 x_2) = \beta_{12}\Phi'(\cdot)$. Perhaps this is because statistical software packages, such as Stata® 7, compute the marginal effect for any explanatory variable. However, Eq. (2) shows clearly that the interaction effect is not equal to $\beta_{12}\Phi'(\cdot)$.

There are four important implications of Eq. (2) for nonlinear models. Firstly, the interaction effect could be nonzero, even if $\beta_{12} = 0$. For the probit model with $\beta_{12} = 0$, the interaction effect is

$$\left.\frac{\partial^2 \Phi(\cdot)}{\partial x_1 \partial x_2}\right|_{\beta_{12}=0} = \beta_1 \beta_2 \Phi''(\cdot).$$

Secondly, the statistical significance of the interaction effect cannot be tested with a simple t-test on the coefficient of the interaction term β_{12}. Thirdly, the interaction effect is conditional on the independent variables, unlike the interaction effect in linear models. (It is well known that the marginal effect of a single uninteracted variable in a nonlinear model is conditional on the independent variables.) Fourthly, the interaction effect may have different signs for different values of covariates. Therefore, the sign of β_{12} does not necessarily indicate the sign of the interaction effect.

In order to improve best practice by applied econometricians, we derive the formulas for the magnitude and standard errors of the estimated interaction effect in general nonlinear models. The formulas apply easily to logit, probit, and other nonlinear models. We illustrate our points with an example.

2. Estimation

We begin by introducing notation for general nonlinear models. Let y denote the raw dependent variable. Let the vector $\mathbf{x}$ be a $k \times 1$ vector of independent variables, so $\mathbf{x}' = (x_1 \ldots x_k)$. The expected value of y given x is

$$\mathrm{E}[y|\mathbf{x}] = F(x, \beta), \tag{3}$$

where the function F is known up to β and is twice continuously differentiable. Let Δ denote either the difference or the derivative operator, depending on whether the regressors are discrete or continuous. For example, $\Delta F(\mathbf{x}, \beta)/\Delta x_1$ denotes the derivative if x_1 is continuous and the difference if x_1 is continuous. The key point of this paper is that the interaction effect is found by computing cross derivatives (or differences), not by just looking at the coefficient on the interaction term. The *interaction effect* of x_1 and x_2 on y is

$$\mu_{12} = \frac{\Delta^2 F(\mathbf{x}, \beta)}{\Delta x_1 \Delta x_2}.$$

The interaction effect is estimated by

$$\hat{\mu}_{12} = \frac{\Delta^2 F(\mathbf{x}, \hat{\beta})}{\Delta x_1 \Delta x_2}, \tag{4}$$

where $\hat{\beta}$ is a consistent estimator of β. The continuity of F and consistency of $\hat{\beta}$ ensures the consistency of $\hat{\mu}_{12}$ to μ_{12}.

The standard error of the estimated interaction effect $\hat{\mu}_{12}$ is found by applying the Delta method

$$\hat{\mu}_{12} \sim N\left(\mu_{12}, \frac{\partial}{\partial \beta'}\left[\frac{\Delta^2 F(\mathbf{x}, \beta)}{\Delta x_1 \Delta x_2}\right] \Omega_\beta \frac{\partial}{\partial \beta}\left[\frac{\Delta^2 F(\mathbf{x}, \beta)}{\Delta x_1 \Delta x_2}\right]\right).$$

The asymptotic variance of $\hat{\mu}_{12}$ is estimated consistently by

$$\hat{\sigma}_{12}^2 = \frac{\partial}{\partial \beta'}\left[\frac{\Delta^2 F(\mathbf{x}, \hat{\beta})}{\Delta x_1 \Delta x_2}\right] \hat{\Omega}_\beta \frac{\partial}{\partial \beta}\left[\frac{\Delta^2 F(\mathbf{x}, \hat{\beta})}{\Delta x_1 \Delta x_2}\right], \tag{5}$$

where $\hat{\Omega}_\beta$ is a consistent covariance estimator of $\hat{\beta}$. The t statistic is $t = \hat{\mu}_{12}/\hat{\sigma}_{12}$, which has an asymptotic standard normal distribution under some regularity conditions. Use the t statistic to test the hypothesis that the interaction effect equals zero, for given $\mathbf{x}$.

Eq. (3) encompasses many commonly used models, including logit, probit, tobit, censored regression models, log transformation models with normal errors, count models, and duration models. Interaction terms between three or more variables are found in an analogous way.

3. Empirical example

To illustrate our points, we estimated a logit model to predict HMO enrolment as a function of three continuous variables—age, number of activities of daily living (a count from 0 to 6 of the number of basic physical activities a person has trouble performing), and the percent of the county population enrolled in a HMO—and their interactions (Mello et al., 2002). The data are primarily from the 1993–1996 Medicare Current Beneficiary Survey, a longitudinal survey of Medicare eligibles. There are 38,185 observations at the person–year level, after excluding persons who lived in counties not served by a Medicare HMO. The dependent variable is a dummy variable indicating whether the individual is enrolled in a HMO. About 12% are in HMOs. The average age is 77 years, 65% have no limitations in ADLs, the average number of limitations is one, and the average market penetration is 9% with a range from 0.0001 to 0.52. Data on market penetration are from the Medicare Market Penetration File.

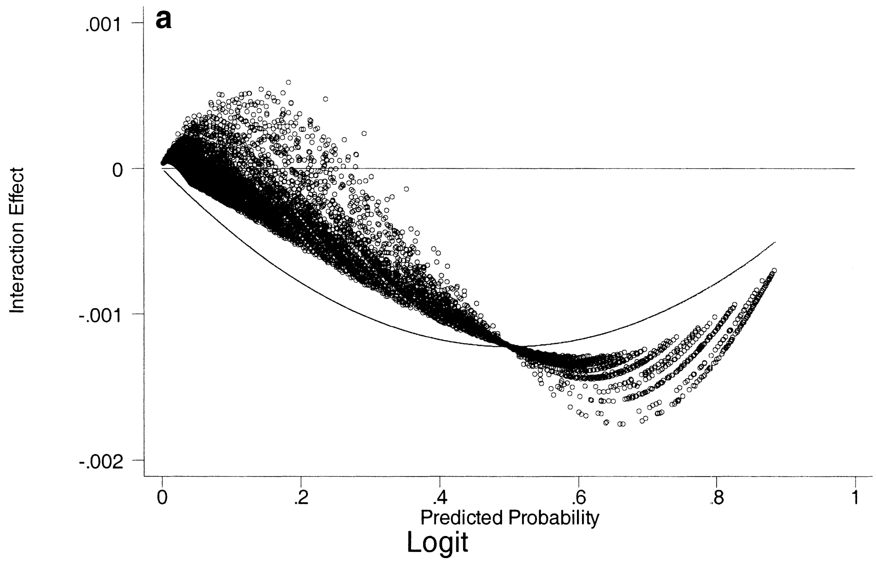
The model was run twice, once with an interaction between age and ADLs, and once with an interaction between age and market penetration. A person is more likely to join a HMO if they are younger, have fewer ADLs, and live in a county with high HMO market penetration.

The coefficient on the interaction term between age and ADLs is negative and statistically significant (see Table 1). However, the magnitude and statistical significance of the interaction effect varies by observation. For many observations with a predicted value of being enrolled in a HMO less than 0.2, the interaction effect is positive, not negative (see Fig. 1A). The concave line drawn for reference is the marginal effect of the interaction term computed by $\beta_{12}F'(\cdot)$. The statistical significance of the interaction effect is often stronger when the interaction effect is positive than when negative, with t-statistics as high as 10 (see Fig. 1B).

Table 1
Logit estimates with interaction terms

Variable	Mean	Min	Max	Model 1	Model 2
Dependent variable					
HMO enrolment	0.122	0	1		
Independent variables					
Constant				3.107**	3.129**
				(0.045)	(0.058)
Age—65	12.49	0	47	0.0233**	0.0214**
				(0.0030)	(0.0045)
Activities of daily living	0.973	0	6	0.105**	0.179**
				(0.030)	(0.015)
HMO market penetration	0.090	0.0001	0.516	10.01**	10.33**
				(0.13)	(0.23)
Interaction terms					
Age × ADLs	16.9	0	270	0.0049**	
				(0.0018)	
Age × market penetration	1.09	0	17.4		0.028
					(0.017)

The sample size is 38,185. Standard errors are in parentheses. * and ** indicate statistical significance at the 5% and 1% levels.



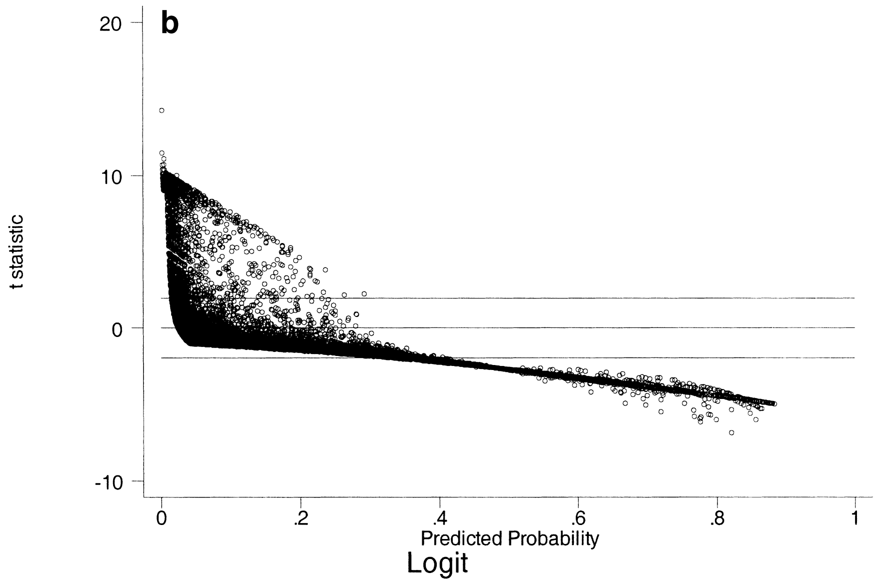


Fig. 1. (a) Interaction effect as a function of the predicted probability, model 1. (b) t-Statistic as a function of the predicted probability, model 1.

In the second model the interaction term between age and market penetration is negative but not statistically significant. Again, the interaction effect varies widely, and is positive for many observations (see Fig. 2A). Even though the interaction term is itself not statistically significant, the interaction effect is significant for most observations (see Fig. 2B).

Having plotted the interaction effect for many logit and probit models with different data sets, we can say that these two examples are typical. The interaction effect always follows an S-shaped pattern

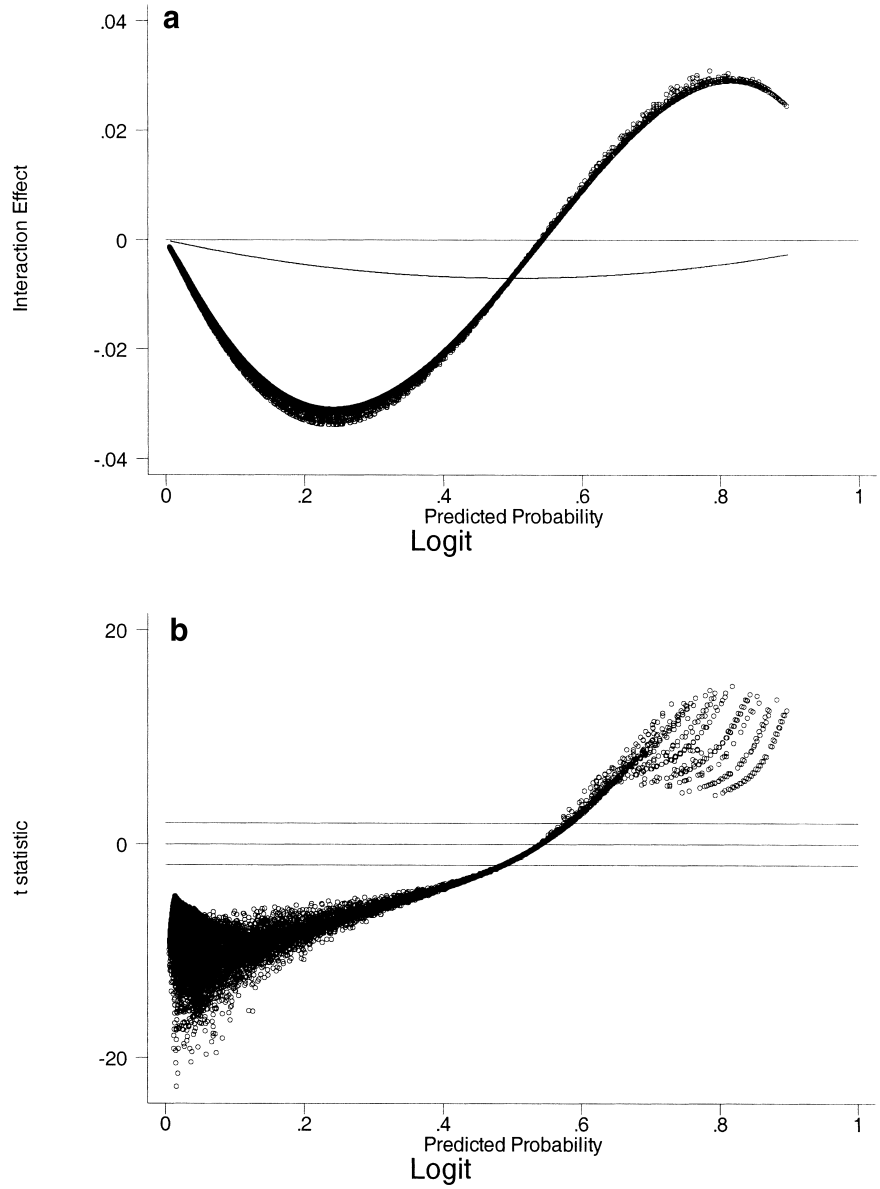


Fig. 2. (a) Interaction effect as a function of the predicted probability, model 2. (b) t-Statistic as a function of the predicted probability, model 2.

when plotted against predicted probability. It crosses the (incorrect) reference line $\beta_{12}F'(\cdot)$ close to $F(\cdot) = 0.5$. The interaction effect is always positive for some observations and negative for others. Unlike the marginal effect of a single variable, the results are strongest, in both magnitude and statistical significance, for values of predicated probability not near 0.5. The results are virtually identical for logit and probit models run on the same data.

4. Conclusion

The interaction effect, which is often the variable of interest in applied econometrics, cannot be evaluated simply by looking at the sign, magnitude, or statistical significance of the coefficient on the interaction term when the model is nonlinear. Instead, the interaction effect requires computing the cross derivative or cross difference. Like the marginal effect of a single variable, the magnitude of the interaction effect depends on all the covariates in the model. In addition, it can have different signs for different observations, making simple summary measures of the interaction effect difficult.

We present a consistent estimator for the interaction effect (cross-difference) for nonlinear model, and for the asymptotic variance of the estimated interaction effect. An example shows that not calculating the correct interaction effect would lead to wrong inference in a substantial percentage of the sample. Sample programs are available from the authors upon request.

References

DeLeire, T., 2000. The wage and employment effects of the Americans with Disabilities Act. Journal of Human Resources 35 (4), 693–715.

Mello, M.M., Stearns, S.C., Norton, E.C., 2002. Do Medicare HMOs still reduce health services use after controlling for selection bias? Health Economics 11 (4), 323–340.