Testing Hypotheses About Interaction Terms in Nonlinear Models*

Abstract

We further examine the interaction effect in nonlinear models that has recently been discussed by Ai and Norton (2003). Statistical tests about partial effects and interaction terms are not necessarily informative in the context of the estimated model. We suggest more useful ways that do not involve statistical testing to examine the interaction effect in binary choice models.

Keywords: Interaction effect; Interaction term; Partial effect; Probit; Logit; Nonlinear Models

JEL classification: C12; C25 ; C51

----

1. Introduction

A recent, widely discussed contribution to econometric practice by Ai and Norton (2003) has proposed an approach to analyzing interaction effects of variables in nonlinear models. The authors focus attention on a binary choice (logit) model, though their results are easily extended to other nonlinear models. The main result of the study applies to a model that contains an interaction term, such as

$$E[y|x_1,x_2,z] = F(\beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \delta z).$$

The authors argue that the common computation of the *partial effect of the interaction term*, $\gamma_{12} = \beta_{12} F'(.)$, is misleading, and provides no useful information about the *interaction effect* in the model, $\Delta_{12} = \partial^2 F(.)/\partial x_1 \partial x_2$. They then provide results for examining the magnitude and statistical significance of estimates of Δ_{12}. This note argues that the proposals made Ai by and Norton are likewise uninformative about interaction effects in the model. The magnitude, itself, has no economic interpretation, even when examined one observation at a time, as suggested, and proposed plots of, e.g., t statistics for individual specific estimates of Δ_{12} against corresponding estimates of $F(.)$ are also uninformative about the relationships among the variables embedded in the model. We argue that the indicated relationships are inherently difficult to describe numerically by simple summary statistics, but graphical devices are much more informative.

As a corollary to this argument, we suggest that the common practice of testing hypotheses about partial effects, as opposed to about structural parameters, is less informative than one might hope, and could usefully be omitted from empirical analyses. The paper proceeds to a summary of the Ai and Norton (2003) results in Section 2, some discussion of the results in Section 3 and an application in Section 4. Conclusions are drawn in Section 5.

* Department of Economics, Stern School of Business, New York University, 44 West 4th St., 7-78, New York, NY, 10012, http://pages.stern.nyu.edu/~wgreene.

2. Estimation and Inference for Interaction Effects

Ai and Norton (2003) point out an ambiguity that arises in many recent applications that report partial effects for nonlinear models with interaction terms. Consider a model of the form

$$E[y|x_1,x_2,z] = F(\beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \delta z), \tag{1}$$

where $F(.)$ is a nonlinear conditional mean function such as the normal or logistic cdf in a binary choice model, x_1 and x_2 are variables of interest, either or both of which may be binary or continuous, and z is a related variable or set of variables, including the constant term if there is one. From this point forward, we will specialize the discussion to the probabilities in a probit model, for which

$$\begin{aligned} E[y|x_1,x_2,z] &= \text{Prob}(y = 1|x_1,x_2,z) \\ &= \Phi(\beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \delta z) \\ &= \Phi(A), \end{aligned} \tag{2}$$

where $\Phi(A)$ is the standard normal pdf. The results will generalize to other models with only minor modification.[1] Partial effects in the model when x_1 and x_2 are continuous are

$$\begin{aligned} \partial E[y|x_1,x_2,z]/\partial\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &= \Phi'(A) \times \partial A / \partial\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &= \phi(A) \times \begin{pmatrix} \beta_1 + \beta_{12} x_2 \\ \beta_2 + \beta_{12} x_1 \end{pmatrix}, \end{aligned} \tag{3}$$

where $\phi(A)$ is the standard normal pdf. When the variable of interest is binary, differentiation is replaced with first differencing. Thus, if x_1 is a dummy variable, the partial effect is

$$\begin{aligned} \Delta E[y|x_1,x_2,z]/\Delta \text{x}_1 &= E[y|x_1 = 1,x_2,z] - E[y|x_1 = 0,x_2,z] \\ &= \Phi(\beta_1 + \beta_2 x_2 + \beta_{12} x_2 + \delta z) - \Phi(\beta_2 x_2 + \delta z). \end{aligned} \tag{4}$$

The *interaction effect* is identified by the authors as the effect of a change in one of the variables on the partial effect on $E[y|x_1,x_2,z]$ of the other variable;

$$\begin{aligned} \frac{\partial^2 E[y \mid x_1, x_2, z]}{\partial x_1 \partial x_2} &= \beta_{12}\Phi'(A) + (\beta_1 + \beta_{12} x_2)(\beta_2 + \beta_{12} x_1)\Phi''(A) \\ &= \beta_{12}\phi(A) + (\beta_1 + \beta_{12} x_2)(\beta_2 + \beta_{12} x_1)[-A\phi(A)]. \end{aligned} \tag{5a}$$

Once again, differentiation is replaced with differencing when the variables are binary;

---

[1] Ai and Norton analyze a logit model, rather than a probit model as done here. A trivial modification of the notation will accommodate their case. In the probit model, $\Phi(.)$ and $\phi(.)$ are familiar notations, while in the logit model, the counterparts are $F(.) = \Lambda(.)$ and $F\,'(.) = \Lambda(.)[1\text{-}\Lambda(.)]$. There is little substantive difference between the models and the results found here would be identical for a logit model.

$$\frac{\partial(\Delta E[y\,|\,x_1,x_2,z]/\Delta x_1)}{\partial x_2} = (\beta_2+\beta_{12})\phi(\beta_1+\beta_2 x_2+\beta_{12}x_2+\delta z) - \beta_2\phi(\beta_2 x_2+\delta z) \quad (5b)$$

or, when both variables are binary,

$$\frac{\Delta^2 E[y\,|\,x_1,x_2,z]}{\Delta x_1 \Delta x_2} = [\Phi(\beta_1+\beta_2+\beta_{12}+\delta z) - \Phi(\beta_2+\delta z)] - [\Phi(\beta_1+\delta z) - \Phi(\delta z)]\,. \quad (5c)$$

The notation $\Delta_{12}(A) = \Delta_{12}$ will be used for all three cases.

The coefficient on the interaction term, β_{12} does not provide the change in the partial effect of either variable on the conditional mean function if the function is nonlinear. Even after scaling by $\phi(A)$ as in (3), the mismeasured interaction effect, $\phi(A)\beta_{12}$, which is what is likely to be reported by software that reports partial effects in the form of scaled coefficients, does not provide a useful measure of any interesting quantity. Ai and Norton (2003) suggest, "[However,] most applied economists instead compute the marginal effect of the interaction term, which is $\partial\Phi(.)/\partial(x_1x_2) = \beta_{12}\Phi'(.)$". A distinction is needed between the *marginal effect of the interaction term*, a first derivative which is meaningless – x_1x_2 cannot change *partially* independently of x_1 and x_2 – and the *interaction effect*, which is the second derivative in (5a,b,c) and which is the main subject of their and this paper. Ai and Norton's discussion is motivated by the fact that statistical software (in their case, Stata® 7) typically mechanically computes a separate "partial effect" for each variable that appears in the model. The product variable would naturally appear as a separate variable in the model specification, and the software would have no way of discerning that it is a product of two variables.

Asymptotic standard errors for the partial and interaction effects in any of these cases may be computed using the delta method, as suggested by the authors in their equations (4) and (5).[2] The argument made in the paper is that for inference about interaction effects, the appropriate path to take is inference about the quantities in (5a,b,c), not about $\beta_{12}\Phi'(A)$. In an application, they demonstrate with plots of observation specific results from (5a) and associated *t* ratios against the predicted probabilities for models without then with second order terms, $\beta_{12}x_1x_2$.

3. Inference About the Interaction Effect

While the computations suggested by Ai and Norton (2003) are correct, the indicated results are not as informative as might be hoped. For example, the applied researcher looking for guidance for "best practice" (their p. 124) might wonder what inference should be drawn from a plot of the "*t* statistics" associated with individual observations on (5a) against the fitted probabilities in (1), as is done in their Figure 1b (and our Figure 3).

Consider, first, a case in which x_1 and x_2 are both continuous variables. Even if β_{12} in (1) equals zero, the interaction effect in (5a) is likely to be nonzero;

---

[2] A computation needed for the calculation will be the third derivative of the normal cdf; $\Phi'''(A) = (A^2-1)\phi(A)$. The counterparts for the logit model would be $\Lambda''(A) = \Lambda(A)[1-\Lambda(A)][1-2\Lambda(A)]$ and $\Lambda'''(A) = \Lambda(A)[1-\Lambda(A)][1-6\Lambda(A)+6\Lambda^2(A)]$. Authors differ on whether the computation of partial effects is better done at the means of the data or averaged over the observations. In the latter case, some further derivation is required to obtain an appropriate standard error for the average of *N correlated* individual partial effects. (They all use the same estimated parameters.) See Greene (2008, p. 780-785). In their applications, Ai and Norton advocate doing the computations for each individual observation separately and plotting some of the results. We return to this suggestion below.

$$\{[\partial^2\Phi(A)/\partial x_1\partial x_2] \mid \beta_{12} = 0\} \;=\; \beta_1\beta_2\Phi''(\mathrm{A}) \;=\; \beta_1\beta_2 \times [-A\phi(A)]. \tag{6}$$

For a probit model, if the coefficients β_1 and β_2 are nonzero, since the density must be positive by construction, the result in (6) can be zero if and only if $(\beta_1 x_1 + \beta_2 x_2 + \delta z)$ is zero. (In the example developed below, each of the two coefficients is assessed as highly significantly different from zero.) Thus, the *interaction effect* is zero when and only when the index function is zero, which, in turn, means that the probability is one half. (This same result will apply in a logit model.) The economic content of the hypothesis is equivalent to that the probability for the individual outcome equaling one is one half.

In a model in which there is a second order term, the result is less transparent. Then, the desired effect is that in (5a). As suggested in the article, one cannot assess the statistical significance of this interaction effect with a simple *t* test on the coefficient on the interaction term β_{12} or even $\beta_{12}F(A)$. All coefficients and data could be positive; but $\Phi''(A) = -A\phi(A)$ might still tip the result toward zero. Even the sign of the interaction term depends on the data and, therefore, it could be close to zero numerically and insignificantly different from zero statistically. In this case, $\Phi = .5$ is neither necessary nor sufficient for the interaction effect to equal zero. On the other hand, if $\Phi = .5$, the interaction effect cannot be nonzero unless β_{12} is also nonzero. Once again, it is unclear how one is to interpret this result in economic terms.

One might answer the preceding by suggesting that the hypothesis, such as it is, should be taken at face value – we are interested in when or whether the interaction effect is zero or not. How it becomes so is merely a mathematical property of the model. We would agree. The fault in this methodology is in the *statistical testing about these partial effects*. The computations are correctly prescribed. But, the *hypothesis* can be meaningless in the context of the original model. The issue becomes clear in the case of a simple partial effect. In the probit model with no second order term, for example,

$$\Delta_1 = \partial E[y|x_1,x_2,z]/\partial x_1 = \beta_1 \times \phi(A). \tag{7}$$

The density is positive if the index is finite. Suppose one has (as in our example below), fit the model and found the estimate of β_1 to be highly significantly different from zero. In testing the hypothesis that Δ_1 equals zero, one is in the contradictory position of testing the hypothesis that the product of two nonzero terms is zero. In practice, in this setting, it rarely occurs that an estimated coefficient is statistically nonzero while the corresponding partial effect is not. But it could – the sampling variance of the estimate of a highly nonlinear finction such as $\phi(A)$ could be very large. If it should occur, the two outcomes are directly contradictory. Inuition, if not hard coded methodology, would suggest that the test based on Δ_1 is the one that should be ignored. If the hypothesis that β_1 equals zero has already been rejected, it is pointless to test the hypothesis that Δ_1 equals zero.

An interesting case is that in which x_1 is a dummy variable and x_2 is continuous. In this instance, the *interaction effect* shows how the partial effect of the continuous x_2 varies with a regime switch in x_1;

$$\frac{\Delta(\partial E[y \mid x_1, x_2, z]/\partial x_2)}{\Delta x_1} = (\beta_2 + \beta_{12})\phi(\beta_1 + \beta_2 x_2 + \beta_{12} x_2 + \delta z) - \beta_2\phi(\beta_2 x_2 + \delta z). \tag{8}$$

As the authors note, the interaction effect depends on all variables in the model. A test of zero interaction effect could now be carried out based on $\beta_2 = \beta_{12} = 0$, which is sufficient, but not necessary. But, this is likely to be stronger than desired, since one might not have in mind to eliminate the interaction effect by eliminating the second variable. The hypothesis that the combination of terms in (8) equals zero at a specific data point could be tested using the methods suggested earlier, as there are configurations of the data that will equate (8) to zero (statistically) without imposing $\beta_2 = \beta_{12} = 0$. When the model contains numerous variables however, interpreting the outcome would be difficult at best. In our application, the model contains, in addition to x_1 = gender and x_2 = age, z = (income, education, marital status, and presence of children in the household). Any number of different combinations of these variables could interact with age and gender to equate (8) to zero statistically. It is unclear what meaning one should attach to this.

The preceding suggests that one can test the hypothesis that the interaction effect is zero for a particular individual, or for the average individual in the data set. It is unclear, however, what the hypothesis means. It does seem that direct examination of $\Phi(A)$ given x_1=1 and x_1=0, or its derivative with respect to x_2, evaluated at interesting combinations of the other variables can, absent statistical testing, be very revealing as we consider in an example.

4. Application

Riphahn, Wambach and Million (RWM) (2003) constructed count data models for physician and hospital visits by individuals in the German Socioeconomic Panel (GSOEP). The data are an unbalanced panel of 7,293 families, with group sizes ranging from one to seven, for a total of 27,326 family-year observations.[3] To illustrate the computations, we fit pooled probit models for *Doctor*, defined to equal one if the individual reports at least one physician visit in the family-year observation and zero otherwise.[4] In the full sample, 62.9% of the individuals reported at least one visit. We have fit binomial probit models based on binary variables for marital status, gender and presence of young children in the household and continuous variables income, age and education. Probit estimates for several specifications are given in Table 1. *Female* and *Married* are binary variables while *Age* and *Income* are continuous, so the last three variables in the table present the different types of interactions discussed above. All estimated coefficients in all models are statistically significant at the 1% level save for that on *Income* in Models 1, 3 and 4, which is significant at the 5% level.

Table 2 reports the estimated partial effects for *Age*, *Female* and *Income* based on Model 0. The highly significant relationship between the binary gender variable and the probabilities is evident in the tabulated partial effects. On average, holding everything else constant, the probability that an individual reports at least one visit to the doctor is .13 larger for women than men. Since the overall proportion is .629, this effect is extremely large. The sample splits roughly equally between men and women, so these values suggest that the average probabilities are about .629 - .13/2 = .564 for men and about .694 for women, for a difference of roughly 23%. The partial effect of *Age* on the probability appears to be about +.004 per year. This is statistically highly significant for the average individual and for every individual in the sample. Over a 40 year observation period, if everything else were held fixed, this would translate to an increase in the probability of about 0.16. If nothing else changed, by age 65, the probability of a doctor visit in any given year would increase from .694 at age 45 for women to .774, and from .564 to .624 for men. However, this computation neglects the interaction between these two

[3]The data are described in detail in RWM (2003) and in Greene (2008, p. 1088).

[4]All computations were done with *NLOGIT 4.0* (http://www.nlogit.com). The raw data may be downloaded from the data archive of the *Journal of Applied Econometrics* at http://qed.econ.queensu.ca/jae/2003-v18.4/riphahn-wambach-million/. The data in the form of an *NLOGIT* project file may be downloaded from http://pages.stern.nyu.edu/~wgreene/healthcare.lpj. Program commands for replicating the results may be downloaded from. http://pages.stern.nyu.edu/~wgreene/InteractionEffects.lim.

effects suggested in Figure 5, where we find that the gap between women and men appears to diminish as they age.

Table 1 Probit Estimates for *Doctor*. (Absolute asymptotic *t* ratios in parentheses)

Variable	Mean	Std.Dev. (Range)	Model 0	Model 1	Model 2	Model 3	Model 4
Constant			-.1243 (2.138)	-.1423 (2.44)	-.2510 (4.03)	-.3058 (3.56)	-.4664 (5.18)
Female*	0.4788	.4996	.3559 (22.22)	.4552 (13.94)	.7082 (11.16)	.3453 (22.11)	.7647 (11.58)
Age	43.53	11.33 (25 - 64)	.01189 (14.95)	.01137 (14.05)	.01559 (15.20)	.01589 (9.89)	.01963 (10.96)
Income	.3521	.1769 (0.0 - 3.1)	-.1324 (2.85)	-.1197 (2.56)	-.1371 (2.94)	.4060 (2.09)	.4885 (2.51)
Married*	.7586	.4279	.07352 (3.56)	.1387 (4.99)	.06241 (3.01)	.07877 (3.80)	.1168 (4.11)
Young Kids*	.4027	.4905	-.1521 (8.30)	-.1613 (8.71)	-.1588 (8.64)	-.1525 (8.32)	-.1658 (8.94)
Education	11.32	2.325 (7 - 18)	-.01497 (4.19)	-.01587 (4.43)	-.01641 (4.578)	-.01452 (4.06)	-.0165 (4.59)
Female× Married				-.131 (3.49)			-.09607 (2.52)
Female× Age					-.00820 (5.74)		-.00787 (5.40)
Income× Age						-.01241 (2.86)	-.01418 (3.27)

* Binary Variable

Table 2 Estimated Partial Effects for Age and Female Based on Model 0

Variable	Coefficient	T ratio	Partial Effect*	Minimum Effect**	Maximum Effect**	Minimum t ratio**	Maximum t ratio**
Age	0.01189	14.95	0.00447	0.00308	0.00474	12.84	19.04
Female	0.3559	22.22	0.130	0.107	0.141	19.81	22.85
Income	-0.1324	-2.85	-0.0498	-0.0528	-0.0343	-3.06	-2.56

* Computed at the sample means
** Based on values computed for each observation

Based on the regression results in Table 1, conclusions based on Table 2 about statistical significance of the partial effects of *Age*, *Female* and *Income* are foregone. The economic content of the results is shown in Figure 1 which traces the impact of *Age* for a particular demographic group, married women under 46 with children, average income and 16 years of education.

Among the problems of partial effects for continuous variables such as *Age* and *Income* is accommodating the units of measurement. For *Age* in particular, it is informative to examine the impact graphically as in Figure 1. The partial effect per year as well as the range of variation are evident in the figure. To underscore the point, consider the income variable, which ranges from 0.0 to about 3.1 in the sample and has a measured partial effect in Model 0 of -0.0498 with a *t* ratio of -2.85. A change of one unit in *Income* is larger than five sample standard deviations, so the partial effect (per unit change) could be quite misleading. Once again, a graphical device that accommodates this scaling issue, such as Figure 2, is likely to be more informative than a simple report of the partial derivative, even if averaged over the sample observations.

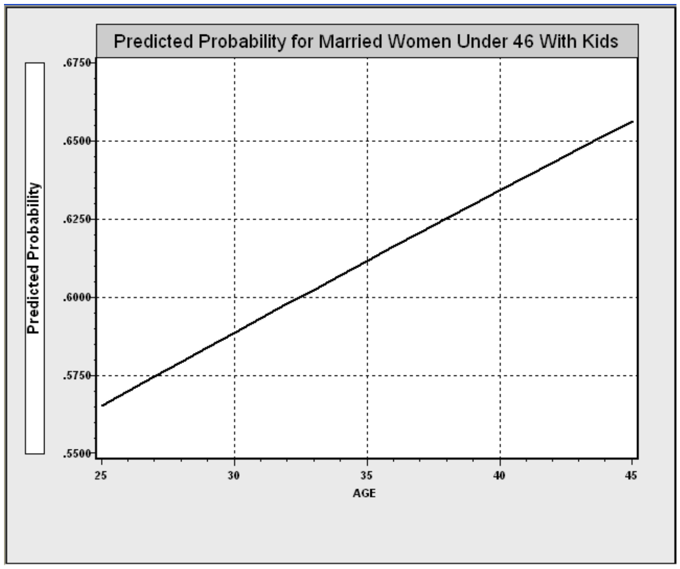

Figure 1 Relationship between Age and Probability of Doctor Visit

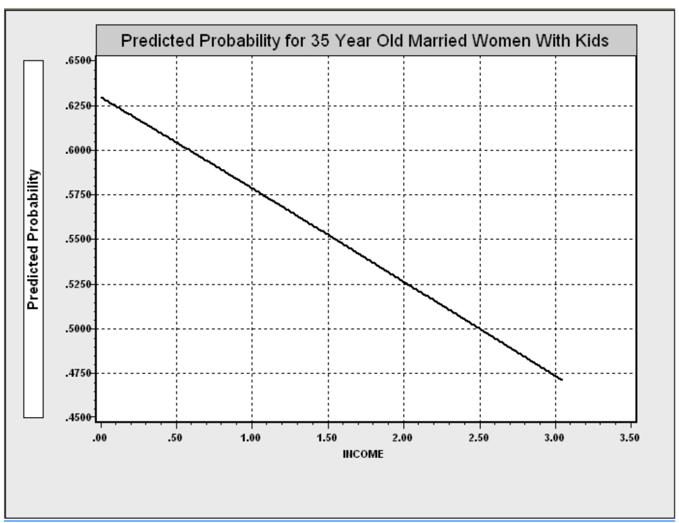

Figure 2 Relationship between Income and Probability of Doctor Visit

Consider, then, the *interaction effects* implied by Model 0, where we examine the interaction between *Age* and *Income*. Figure 3 shows the counterpart to Ai and Norton's (2003) Figure 1(b). It plots the *t* ratio for the hypothesis that the estimated effect in the second derivative in (5a|β_{12}=0) is zero against the estimated probabilities for a random 10% of the observations in the sample. The effect predicted earlier is obvious. The spray of points with *t* ratios close to zero corresponds precisely to those observations which have estimated probability near to 0.5. Since the average predicted probability in the sample matches the sample proportion of ones, 0.629, with a sample standard deviation of 0.09, most of the observations are in the range of probabilities that correspond to statistically insignificant interaction. This would be less pronounced in an unbalanced sample such as Ai and Norton's in which the sample proportion is only about 12% - it is clear in their figures that only a small fraction of the predicted probabilities are near 0.5 and the *t* ratios are correspondingly near to zero. It remains uncertain how this relationship of the probability to .5 relates to the economic content of the interaction of *Age* and *Income* in the model.

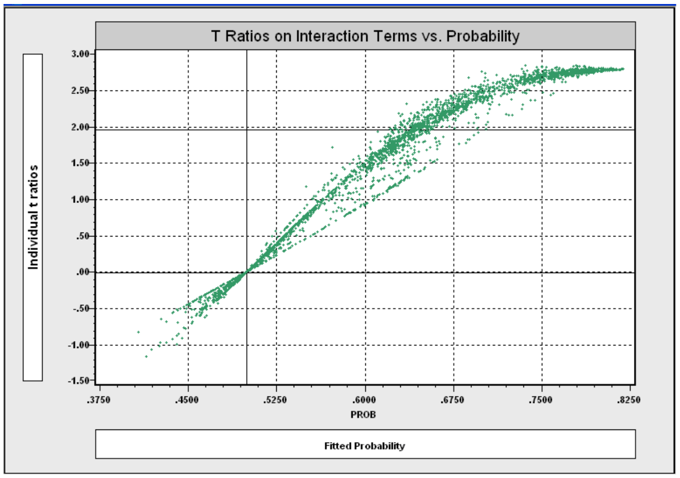

Figure 3 Relationship Between t Statistics and Probabilities In a Model with No Interaction Term

It seems less than obvious what the economic content of Figure 3 might be, however, or how the reader can relate the information it contains to the variation in the variable of interest, Prob(*Doctor*=1|x_1,x_2,z). We can examine the interaction effect more directly, as is done in Figure 4. Figure 4 shows the interaction effect in Model 0, specifically for *Age* and *Female*. In the figure, the partial effect, not the probability is plotted on the vertical axis. Figure 4 is a plot of (3) with *Income*, *Kids*, *Married* and *Education* held constant. The upper curve is for men; the lower is for women. The interaction effect is the change in the partial effect of *Age* with respect to change in gender, which is the distance between the two curves. The distance is given by (5b), $\Delta\{\partial(E[y\,|\,Female, Age, z]/\partial Age)\}/\Delta Female$. (The values of the two functions in Figure 4 are the slopes of the two functions in Figure 5 to follow. Although the functions in Figure 5 appear to be linear, there is actually a very small amount of curvature, as shown in Figure 4.) The interaction effect is of second order, as suggested in the figure, where we have multiplied the scale by 100 to enhance the visibility.

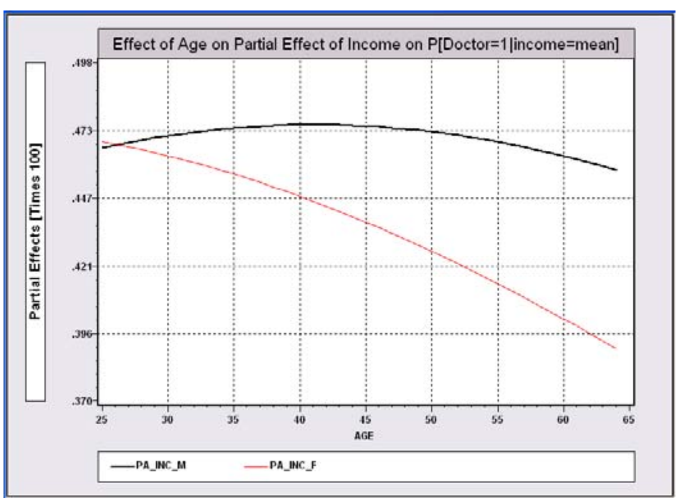

Figure 4 Interaction Effect Between Age and Gender

We might propose something like Figure 4 as an alternative to Figure 3 as an analysis of the interaction effect. However, the derivatives in Figure 4 remain unintuitive. Figure 5 is based on Models 0 and 2. The left panel plots the predicted probabilities for men and women using Model 0 which has no second order term.. The predictors are essentially parallel. (There is a very small, indiscernible difference in the slopes.) In the second model, there is an embedded second order term in the equation. We could interpret *an* "interaction effect" in this model as the change in the distance between the two sets of predicted probabilities. Perhaps confirming expectations, we see that in the expanded model, this interaction effect shows up as a narrowing of the distance between the two predictors. In economic terms, the impact of the interaction is to narrow the gap between predicted probabilities for men and women as age increases. None of Table 2, Figure 3 or Figure 4 would reveal this superficially.

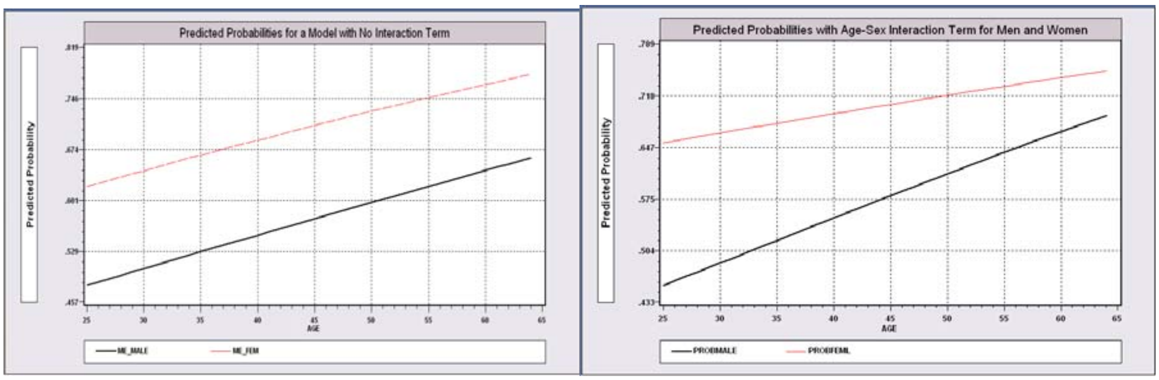

Figure 5 Predicted Probabilities for Models 0 and 2, By Gender

We now consider the interaction of two continuous variables. Table 3 displays the estimated probabilities and interaction effects based on (5a) in Model 3, where there is a second order, interaction term involving *Age* and *Income*. We observe, in contrast to Figure 2, that the interact effect is now statistically significant for every observation in the sample. Figure 6 is the counterpart to Ai and Norton's 2(b). In these data, we find no observations for which the interaction effect is even close to zero, at least statistically. The numerical values average between -0.005 and -0.001. This seems economically trivial, though there is no obvious metric on which to base an evaluation. The measured value is a second derivative of the probability.

Table 3 Estimated Interaction Effects Between Age and Income in Model 3

	Mean	Standard Deviation	Minimum	Maximum
Probability	0.6291	0.1002	0.4015	0.8366
Interaction Effect	-0.004244	0.0007766	-0.005085	-0.001669
t Ratio	-2.73	0.12	-3.18	-2.17

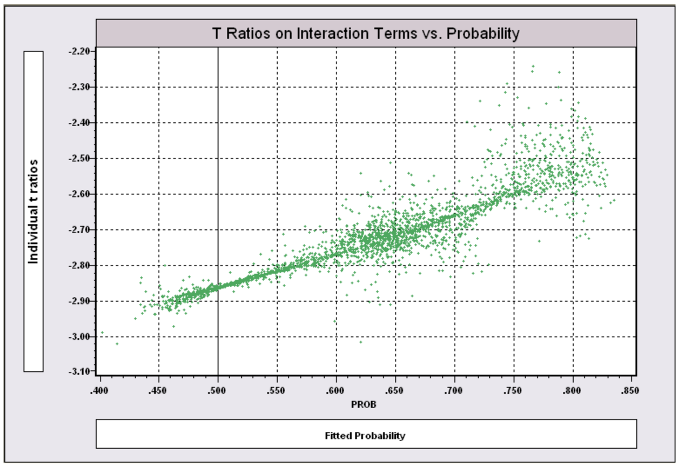

Figure 6 Relationship Between t Statistics and Probabilities In a Model with Interaction Between Age and Income

As in Figure 3, Figure 6 does not suggest obviously how the probability of interest varies with *Age* or *Income*. We do note, based on Model 0 in Table 1, that *Age* and *Income* appear to act in opposite directions. Thus, the striking result in Figure 7 might not be unexpected. Figure 7 displays the relationship between *Income* and the fitted probability for four ages, 25, 35, 45 and 55. The interaction effect between *Age* and *Income* acts to reverse the sign of the partial effect of *Income* at about age 34. Once again, this result would not have been anticipated by the analysis in Table 3 or Figure 6.

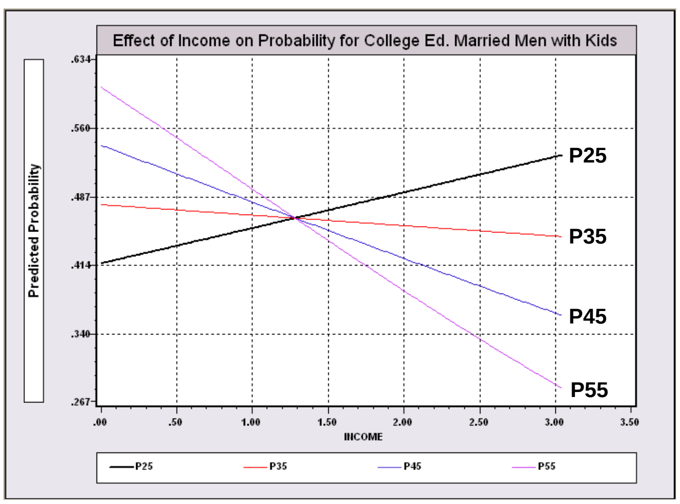

Figure 7 Relationships between Age and Income and Probability of Doctor Visit

5. Conclusions

The preceding does not fault Ai and Norton's (2003) suggested calculations. Rather, we argue that the process of statistical testing about partial effects, and interaction terms in particular, produces generally uninformative and sometimes contradictory and misleading results. The mechanical reliance on statistical measures of significance obscures the economic, numerical content of the estimated model. We conclude, on the basis of the preceding, in the words of the authors, that "to improve best practice by applied econometricians," a useful way to proceed in the analysis is a two step approach:

1. Build the model based on appropriate statistical procedures and principles. Statistical testing about the model specification is done at this step Hypothesis tests are about model coefficients and about the structural aspects of the model specifications. Partial effects are neither coefficients nor elements of the specification of the model. They are *implications* of the specified and estimated model.

2. Once the model is in place, inform the reader with analysis of model implications such as coefficient values, predictions, partial effects and interactions. We find that graphical presentations are a very informative adjunct to numerical statistical results for this purpose. Hypothesis testing need not be done at this point. Even where the partial effects are the ultimate target of estimation, it seems it would be rare for a model builder to build a structural model by hypothesizing (statistically) about partial effects and/or predictions that would be made by that model.

This prescription conflicts with common practice. Widely used software packages such as *Stata*, *NLOGIT*, *EViews,* and so on all produce standard errors and *t* statistics for estimated partial effects, and it has become commonplace to report them among statistical results. The computations detailed in the example above were also simple to apply. See footnote 4. Ai and Norton have also made generally available a set of *Stata* code, *INTEFF*, for this purpose. In this note, we suggest that in spite of the availability of off the shelf software which facilitates the computations, the most informative point in the analysis at which to do hypothesis testing is at the model building step, not at the analysis step.

References

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

Greene, W., *Econometric Analysis, 6th Ed.*, Prentice Hall, Englewood Cliffs, NJ, 2008.

Riphahn, R., A. Wambach and A. Million, "Incentive Effects in the Demand for Health Care: A Bivariate Panel Count Data Estimation," *Journal of Applied Econometrics*, 18, 4, 2003, pp. 387-405.