## American Politics Research

http://apr.sagepub.com/

## Interpretation of Interaction Effects in Logit and Probit Analyses: Reconsidering the Relationship Between Registration Laws, Education, and Voter Turnout CHI HUANG and TODD G. SHIELDS <br> American Politics Research 2000 28: 80 <br> DOI: 10.1177/1532673X00028001005

The online version of this article can be found at:
http://apr.sagepub.com/content/28/1/80

Published by:<br>©SAGE<br>http://www.sagepublications.com

Additional services and information for American Politics Research can be found at:
Email Alerts: http://apr.sagepub.com/cgi/alerts
Subscriptions: http://apr.sagepub.com/subscriptions
Reprints: http://www.sagepub.com/journalsReprints.nav
Permissions: http://www.sagepub.com/journalsPermissions.nav
Citations: http://apr.sagepub.com/content/28/1/80.refs.html

> >> Version of Record - Jan 1, 2000
> What is This?

# INTERPRETATION OF INTERACTION EFFECTS IN LOGIT AND PROBIT ANALYSES Reconsidering the Relationship Between Registration Laws, Education, and Voter Turnout 

CHI HUANG<br>National Chung-Cheng University<br>TODD G. SHIELDS<br>University of Arkansas


#### Abstract

Scholars have argued that more restrictive registration laws most drastically deter the least educated citizens from political participation. Others, however, argue that the most educated, rather than the least educated, are most drastically impeded by restrictive registration requirements. These opposing conclusions have dramatically different implications concerning registration reform in the United States. In this analysis, we urge scholars to take the arguments made by Nagler more seriously, and we argue that past models have not fully considered the inherently nonlinear functional form of the logit and probit models. Using graphical displays, we show that citizens with moderate levels of education are actually those who are "hardest hit" by restrictive closing dates. Consequently, we moderate all prior conclusions and show evidence that it is neither the most nor the least educated who are the "hardest hit" by early closing dates.


In their seminal work, Wolfinger and Rosenstone (1980) concluded that " $[t]$ he barriers imposed by restrictive [registration] laws seem to make little difference to the well educated but are a fairly formidable impediment to people with less interest and bureaucratic skill" (p.80). Nagler (1991), however, explicitly tests for these hypothesized interactive relationships and concludes that "more-educated persons have a harder time than less-educated persons in dealing with the burden of registration" (italics in original, pp. 1400-1402) (see also Brians \& Grofman, 1994; Nagler, 1994). Although scholars of political participation continue to be extremely interested in the effects of registration requirements, investigators have not fully heeded the advice of Nagler ( 1991,1994 ) and continue to discuss interactive relationships between closing date and education without explicitly modeling such
multiplicative effects within standard logit and probit analyses. In fact, scholars continue to speak of the effects of registration requirements being greatest among the least educated without explicitly testing for such interactive relationships (Highton, 1997). We argue that if scholars explicitly test for hypothesized interactive effects and carefully consider the nonlinear functional form involved in probit and logit analyses, it is clear that registration requirements have their greatest effects among those with moderate levels of education.

## INTERACTIVE EFFECTS IN LOGIT AND PROBIT ANALYSIS

The $S$-shaped response curves in probit and logit analyses reflect a common and substantively meaningful outcome: A given change in probability is more difficult to obtain when the probability is closer to the limits of 0 (the floor) and 1 (the ceiling), and is easier to obtain when the indeterminacy is highest (usually at the midpoint of .5) (Cox \& Snell, 1989, pp. 7-9; Hanushek \& Jackson, 1977, pp. 183-184). Thus, the effect of a change in an independent variable on the probability of an event occurring is sensitive to the location on the curve where the assessed change takes place. For a model with multiple independent variables, the location of a unit, in turn, is determined not only by the concerned variable but also by all other explanatory variables. For example, suppose that two independent variables individually have a positive effect on the probability of an event occurring in a probit model. Then, without adding the product to the equation, a negative interactive effect is already "built in" above the midpoint of .5 so as to "suppress" the growth rate of the probability, whereas a positive interaction is built in below .5 to "compensate" for the decline rate of the probability. ${ }^{1}$ Built-in interactions implied by the $S$-shaped curve become part and parcel of the entire model. They reflect our presumption of how probability behaves and thus have important implications for substantive interpretation. Consequently, a "substantive interaction" (Nagler 1991, p. 1397) must be modeled by a statistically significant multiplicative term in the equation. In addition, scholars must be constantly aware of the nonlinearity of the probit and logit models because their link functions cannot be disentangled from their linear index functions. We thus argue that scholars should not evaluate inter-
active effects in logit and probit analysis based solely on the "underlying linear model" (Nagler, 1991, p. 1402) but explicitly incorporate the nonlinearity of these models into their analysis and interpretation. If the underlying data-generating process does not follow a nonlinear functional form, as assumed in logit or probit analyses (or other similar functional forms such as scobit), then such an assumption becomes a severe limitation, and a more theoretically appropriate functional form should be chosen. However, once a researcher chooses to conduct a probit or logit analysis, the nonlinear functional form becomes vital to understanding the full results of the analysis. Furthermore, we suggest the use of graphical techniques to display the full range of interactive relationships involved in logit or probit analyses (or any analysis assuming a nonlinear functional form). These graphs provide meaningful information because they reflect the derivatives of probability with respect to a specific independent variable while holding others constant. ${ }^{2}$ They also provide researchers with a visually appealing manner of interpretation that retains the information of the assumed nonlinear functional form.

For example, it is difficult to further increase the probability of those who already rank high on the curve (the ceiling effect) or decrease the probability of those who already rank near the bottom of the curve (the floor effect). Those whose fitted probabilities are located somewhere in the middle will be more sensitive to changes in variables. Everyone knows that it takes much more effort (in terms of increasing the value of the latent index by changing the value(s) of one or more variables) to raise a person's probability, say, from .8 to .9 than it takes to raise the probability from . 4 to .5. Yet, it is equally true that it will take more effort, in the opposite direction, to lower a person's probability from .9 to .8 than from .5 to $.4 .^{3}$ This point turns out to be deceptively simple, but crucial. That is, if we ignore the transformation function of the probit model, we also ignore the built-in ceiling and floor effects-and the same is true in analyses assuming a scobit or logit functional form. As will be shown below, even if early closing dates tend to cause a greater decrease in the estimated latent index among the better educated, it is still not daunting enough to lower the probability of voting in equal proportion among the well educated who already have a relatively high projected probability of voting.

With these methodological precautions in mind, we reexamine the relationships between registration laws and education. ${ }^{4}$ First, we look at Wolfinger and Rosenstone's (1980) original model, then Nagler's (1991) replication of their 1972 model, and then estimate Nagler's final model using 1992 data. Needless to say, all the following interpretations remain at the disaggregate level, and the assessment of marginal and interactive effects is strictly based on the probability of voting (mathematical derivations and their applications to Nagler's full model are presented in the appendix). Following the graphical methods suggested by Fox (1987) and King (1989, pp. 104-106) to calculate and plot the predicted probability of voting for the most "typical" individuals (shown in the following figures), the two variables of central concern, closing date and education, are permitted to vary over their empirical ranges, whereas the other independent variables are fixed at their most typical values. Hence, all the graphs below display the effects of closing date on turnout across different education levels of typical individuals, who are at the median age (among eligible voters) of 43 in 1972 and 45 in 1992, live in nonsouthern states that have regular registration office hours, allow for absentee registration, do not require evening/Saturday registration, and do not have a concurrent gubernatorial election. ${ }^{5}$

Figure 1 illustrates the level of predicted probabilities for typical individuals based on Wolfinger and Rosenstone's original model for 1972. Due to the positive marginal effect of education on voting, as shown by (A4) in the appendix, those with higher education also have a higher fitted probability of voting. Because the marginal effect of closing date is negative (see [A5] in the appendix), all eight curves are downward sloping: Early closing dates discourage voting for citizens across all education levels. However, visual inspection of the plot indicates that the downward slopes become less steep for those who gain more than 8 years of education. Figure 1 also shows that this 8 -year education level has a fitted probability of voting close to the midpoint of .5 , an area most sensitive to even a slight change in any variable. That is, for those who have fewer than 8 years of schooling, higher education does raise the level of their probabilities to vote. Yet, as this level continues to move toward the threshold of .5, the type of person we are dealing with also shifts from a "very unlikely voter" toward a


Figure 1: Predicted Probability of Voting (vertical axis) by Closing Date (horizontal axis)Wolfinger and Rosenstone's Final Model (1972 Current Population Survey)
" $50-50$ chance voter" and thus becomes more and more sensitive to the closing date. Those who have more than 8 years of education, however, tend to already have crossed this threshold. As a result, the slopes of the probability curves increase, that is, the effect of closing dates become less and less negative, as education levels continue to rise.

Consequently, Wolfinger and Rosenstone's "more educated, less deterred" conclusion is not a general one even for their own model specification and only applies to those whose combined values of all the independent variables are high enough to lift them beyond the threshold of a .5 predicted probability, but not to individuals who fall below this point. ${ }^{6}$ In other words, individuals with very low levels of education tend to have a predicted probability of voting so close to the floor that it is difficult to lower it much further even if the states they

TABLE 1
Estimates of Nagler's Replication of
Wolfinger and Rosenstone's Final Model and His Own Complete Model, 1972 and 1992

| Independent Variables | Current Population Studies: Voter Supplement 1972 |  |  |  | Current Population Studies: <br> Voter Supplement 1992 |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Wolfinger \& Rosenstone (1972) |  | Nagler's Full Model (1972) |  | Nagler's Final Model (for 1984) Estimated on the 1992 Presidential Election |  |
|  | Estimated Coefficient | $t$-Ratio | Estimated Coefficient | $t$-Ratio | Estimated Coefficient | $t$-Ratio |
| Intercept | -2.4928 | $-53.74 * *$ | -2.7597 | -26.51 ** | -2.491 | $-18.99^{* *}$ |
| Education | . 2635 | 20.52** | . 3544 | 7.90** | . 3182 | 4.90** |
| Education ${ }^{2}$ | . 0035 | 2.56** | -. 0029 | -. 59 | . 0147 | 1.80* |
| Age | . 0653 | 48.69** | . 0652 | 48.60** | . 0551 | 36.73** |
| Age ${ }^{2}$ | -. 0005 | -35.93** | -. 0005 | -35.79** | -. 0004 | -26.67** |
| South | -. 1935 | -15.03** | -. 1939 | -15.07** | -. 1429 | -13.49** |
| Gubernatorial election | . 0683 | 6.66** | . 0686 | 6.70** | -. 0199 | -1.51 |
| Irregular registration hours | -. 0155 | -. 97 | -. 0134 | -. 84 | - | - |
| Evening/Saturday registration | - 1009 | 10.10** | . 1010 | 10.11** | - | - |
| No absentee registration | -. 0291 | -1.92* | -. 0293 | -1.94* | - | - |
| Closing date | -. 0062 | $-10.27 * *$ | . 0032 | . 96 | . 0056 | 1.16 |
| Closing Date $\times$ Education |  |  | -. 0032 | $-2.07 * *$ | -. 0052 | -2.11** |
| Closing Date $\times$ Education ${ }^{2}$ |  |  | . 0002 | 1.30 | . 0004 | 1.45 |
| Number of cases | 90,279 |  | 90,279 |  | 81,051 |  |
| Percentage voting | 65.30 |  | 65.30 |  | 69.4 |  |
| Correctly predicted | 70.62 |  | 70.61 |  | 73.2 |  |
| Log-likelihood - | -51,915 |  | -51,909 |  | -43,828 |  |




Figure 2: Predicted Probability of Voting (vertical axis) by Closing Date (horizontal axis)—Nagler's Full Model (1972 Current Population Survey)
reside in make registration deadlines many days earlier. On the other hand, individuals with very high education levels are almost sure to vote, so that an early closing date of voter registration has only a mild effect on their likelihood of voting. In contrast, it is the type of person who has nearly a $50-50$ chance of turning out to vote that is hardest hit by an early closing date.

Of course, adding multiplicative terms to explicitly model this interactive relationship is necessary but complicates mathematical derivation of marginal and interactive effects (as shown in the appendix, [A7] through [A9]). Fortunately, graphical methods make interpretation easier. Presented in column 2 of Table 1 and graphically displayed in Figure 2 is Nagler's (1991) interactive model. Due to the positive effect of education, the level of predicted probability to vote in Figure 2 rises as education increases. With a slight exception of the
lowest education level, all other curves are downward sloping due to the negative effects of closing dates on voting. A visual inspection of the slopes of fitted probability curves reveals that the decline rate increases from the lowest education level up to high school graduates and then levels off at some college and beyond. This pattern-the negative marginal effects of closing dates first deepening and then attenuating as education increases-restricts the "better educated, more deterred" conclusion and shows that those who are "hardest hit" are actually those with moderate levels of education.

## A COMPARISON BETWEEN THE 1972 MODELS

Although both models in columns 1 and 2 of Table 1 indicate that the negative effect of closing dates on voting first deteriorates and then ameliorates across greater education levels, there are two important differences between them. First, they differ in terms of the education level of the typical individuals who are most sensitive to the closing date. Unlike Wolfinger and Rosenstone's model, which shows that people with 8 years of schooling are the hardest hit, Nagler's full model points to high school graduates as most likely to be adversely affected by an early closing date. This difference is important in that high school graduates account for the largest proportion of the eligible voters ( $36 \%$ according to the 1972 Current Population Studies [CPS] data). Second, the two models also differ in terms of the location of the threshold on fitted probability. According to Nagler's model, the education group that is hit hardest by closing date (i.e., high school graduates) is no longer the same group that is closest to the fitted probability of .5 (i.e., those with 8 years of education). This means that compared with Wolfinger and Rosenstone's model, Nagler's model "raises" the threshold at which education begins to diminish the negative effect of closing dates.

Both of these differences are the direct consequence of including two multiplicative interaction terms in Nagler's model. Because the first multiplicative term (closing date $\times$ education) has a large negative estimated coefficient of -.0032 compared with the much smaller positive coefficient estimate of .0002 attached to the second multiplicative term (closing date $\times$ education-squared), the generally positive effect
of education is now "pulled" downward by the closing date. This dragging effect of the closing date demands even higher education for a person to compensate for the additional cost imposed by the statelevel registration systems. However, it by no means overwhelms the beneficial effect of education on voting for all citizens.

Finally, Nagler (1991, p. 1401; 1994, p. 251) shows that these findings concerning closing date and education persist through the 1984 election, although he drops the variables for "irregular registration hours," "evening/Saturday registration," and "no absentee registration" from the model for the 1984 election because "closing [date] is by far the most important of the registration variables" (Nagler, 1991, p. 1404). To demonstrate that these relationships still persist, we estimate Nagler's 1984 restricted model on the more recent 1992 data (shown in column 3 of Table 1). As shown in Figure 3, the same patterns remain true: It is neither the least educated nor the most educated who are hardest hit by early closing dates but those with moderate levels of education. The predicted probabilities of voting for those at the highest and lowest levels of education change very little across closing dates. Those with moderate educational attainment, however, suffer the greatest drop in predicted probability of voting as registration requirements become more severe.

## CONCLUSION

We argue that citizens most deterred by an early closing date of voter registration are neither the best educated nor the least educated. Typically, the negative effect of closing date on the probability of voting deteriorates only up to the middle level of education. This downward trend is then reversed for those who obtain some higher education. As a result of the complexity involved in nonlinear analyses, it is imperative that researchers exercise great caution in both model specification and interpretation. ${ }^{7}$ Interactions assumed by an $S$-shaped curve are not an evil to avoid but an approximation of the underlying data-generation process of binary responses. Therefore, the marginal and interactive effects of independent variables should be evaluated against the probability of an event occurring. We urge scholars to not
Predicted Probability of Voting

Education Level

Closing Date
Figure 3: Predicted Probability of Voting (vertical axis) by Closing Date (horizontal axis)Nagler's Final Model (1992 Current Population Survey)
only heed the advice of $\operatorname{Nagler}(1991,1994)$ and include explicit interaction terms to test the hypothesized relationships but also to remember that such multiplicative terms, once included, are incorporated into the entire nonlinear model and, thus, affect the behavior of built-in interactions. As these interactive effects become too complex to trace, graphical displays are recommended to simplify interpretation. Finally, these results suggest that efforts aimed at easing registration requirements may do little to increase the probability of voter turnout among the least (or most) educated citizens. Instead, such reforms are likely to have their greatest effects on those citizens with moderate levels of educational attainment.

## APPENDIX

## General Forms of Marginal and Interactive <br> Effects in Probit and Logit Models

A dichotomous dependent variable is assumed to be the realization of a random variable $Y$ with the Bernoulli distribution, whose single unknown parameter is the probability for the $i$ th observation to take the value of 1 (King, 1989, pp. 98-99). A researcher's task then is to explain the variation of this probability $\pi_{i}$ with a set of independent variables $x_{i}$. However, since $\pi_{i}$ is bounded between 0 and 1 , whereas the independent variables can take any real values, we also need to choose a function $F$ that links the set of explanatory variables with $\pi_{i}$ appropriately. Thus, in a general probability model, the probability for the $i$ th unit to have an observed value of 1 is:

$$
\begin{equation*}
\operatorname{Pr}\left(Y_{i}=1\right) \equiv \pi_{i}=F\left(x_{i}^{\prime} \beta\right) \tag{A1}
\end{equation*}
$$

where the transformation function $F(\bullet)$ is nondecreasing and satisfies the constraint $0 \leq F\left(\boldsymbol{x}_{i}^{\prime} \beta\right) \leq 1$, a condition that characterizes the cumulative distribution function (cdf) of a continuous random variable. In practice, the most frequently used functions are the cdf of a standard normal, $\Phi$, and that of a logistic distribution, $\Lambda$. The argument of $F, \boldsymbol{x}_{i}^{\prime} \beta$, looks like the systematic component of a linear multiple regression and is assumed to be linear in parameter. In particular,
$\mathbf{x}_{i}^{\prime} \beta=\left\{\begin{array}{c}\beta_{0}+\sum_{l=1}^{p} x_{i l} \beta_{l} \text { if the linear predictor is also linear in all the variables; } \\ \beta_{0}+\sum_{l=1}^{p} x_{i l} \beta_{l}+\sum_{m=p+1}^{q} x_{i m}^{*} \beta_{m} \text { where } x_{i m}^{*} \text { refers to any term that enters } \\ \text { the equation in power and / or product. }\end{array}\right.$
Alternatively, the probability model for a binary dependent variable can be derived from the random utility theory (Ben-Akiva \& Lerman, 1985, pp. 42-58). In this case, $x^{\prime} \beta$ is often called the index function because it is used to model the net benefit (for an individual $i$ to make a certain choice) as an unobserved (or latent) index $y^{*}$ such that $y_{i}{ }^{*}=x_{i}^{\prime} \beta+u_{i}$. This last equation is what Nagler (1991, p. 1402) calls the "underlying linear model." Indeed, by assuming that the error term $u$ follows a standard normal or logistic distribution, one has the probit or logit model, respectively, with $E\left[y^{*}\right]=x_{i}^{\prime} \beta$. It should be emphasized, however, that $y^{*}$ serves only as a theoretical construct to derive a probability model in the general form of (A1). Therefore, the assessment of marginal and interactive effects of the explanatory variables should be based on the resulting probability model instead of the intermediate equation for the latent index.

The general form of the instantaneous effect of the $j$ th independent variable, $x_{j}$, on the probability $\pi_{i}$ is:

$$
\begin{align*}
\frac{\partial \pi_{i}}{\partial x_{i j}} & =f\left(\mathbf{x}_{i}^{\prime} \beta\right) \cdot \frac{\partial \mathbf{x}_{i}^{\prime} \beta}{\partial x_{i j}}  \tag{A2}\\
& =\left\{\begin{array}{l}
\phi\left(\mathbf{x}_{i}^{\prime} \beta\right) \cdot \frac{\partial \mathbf{x}_{i}^{\prime} \beta}{\partial x_{i j}} \text { for probit model, } \\
\lambda\left(\mathbf{x}_{i}^{\prime} \beta\right) \cdot \frac{\partial \mathbf{x}_{i}^{\prime} \beta}{\partial x_{i j}}=\pi_{i} \cdot\left(1-\pi_{i}\right) \cdot \frac{\partial \mathbf{x}_{i}^{\prime} \beta}{\partial x_{i j}} \text { for logit model, }
\end{array}\right.
\end{align*}
$$

Where $f(\bullet)$ is the corresponding probability density function (pdf) of the chosen cdf. Specifically, $\phi(\bullet)$ is the pdf of the standard normal, whereas $\lambda(\cdot)$ is the pdf of the logistic distribution. In the special case where the index function is linear in variable $x_{j}$, that is, $x_{j}$ involves no power or product term of any order, the partial derivative of $\boldsymbol{x}^{\prime} \beta$ B with respect to $x_{j}$ is just the latter's coefficient $\beta_{j}$ (Amemiya, 1981, p. 1488; King, 1989, p. 109; Maddala, 1983, p. 23). The presence in (A2) of $f\left(\mathbf{x}_{i} \beta\right.$ ), whose value "controls" the size of the marginal effect, is a vivid reminder of the importance of taking into account individual $i$ 's location on the response surface prior to the assessed change in variable $x_{j}$. The term $\phi\left(\mathbf{x}_{i} \beta\right.$ ) in (A2) for the probit model cannot be simplified into a closed algebraic form, as its counterpart in the logit model, and thus may look less familiar. It is actually the ordinate of the familiar standard normal distribution at the location point of $\mathbf{x}_{i} \beta$. The value of $\phi\left(x_{i}^{\prime} \beta\right)$ reaches its peak of .3989 at the center of the distribution (i.e., when $\mathbf{x}_{i}{ }^{\prime} \beta=0$ ), declines as the point departs from the center, and approaches 0 as the point approaches the extreme of $+\infty$ or $-\infty$. In other words, $0<\phi\left(\mathbf{x}_{i} \beta\right) \leq .3989$. To examine the interactive effect between two independent variables, $x_{j}$ and $x_{k}$, on the probability $\pi_{i}$, we need to further evaluate (A2) with respect to $x_{k}$. This procedure leads to the following general form of the cross-partial derivative of $\pi_{i}$ with respect to $x_{j}$ and $x_{k}$ :

$$
\begin{align*}
\frac{\partial^{2} \pi_{i}}{\partial x_{i j} \partial x_{i k}} & =f\left(\mathbf{x}_{i}^{\prime} \beta\right) \cdot \frac{\partial^{2} \mathbf{x}_{i}^{\prime} \beta}{\partial x_{i j} \partial x_{i k}}+f^{\prime}\left(\mathbf{x}_{i}^{\prime} \beta\right) \cdot \frac{\partial \mathbf{x}_{i}^{\prime} \beta}{\partial x_{i j}} \cdot \frac{\partial \mathbf{x}_{i}^{\prime} \beta}{\partial x_{i k}}  \tag{A3}\\
& =\left\{\begin{array}{l}
\phi\left(\mathbf{x}_{i}^{\prime} \beta\right) \cdot \frac{\partial^{2} \mathbf{x}_{i}^{\prime} \beta}{\partial x_{i j} \partial x_{i k}}-\phi\left(\mathbf{x}_{i}^{\prime} \beta\right) \cdot\left(\mathbf{x}_{i}^{\prime} \beta\right) \cdot \frac{\partial \mathbf{x}_{i}^{\prime} \beta}{\partial x_{i j}} \cdot \frac{\partial \mathbf{x}_{i j}^{\prime} \beta}{\partial x_{i k}} \text { for probit model } \\
\pi_{i} \cdot\left(1-\pi_{i}\right) \cdot \frac{\partial^{2} \mathbf{x}_{i}^{\prime} \beta}{\partial x_{i j} \partial x_{i k}}+\pi_{i} \cdot\left(1-\pi_{i}\right) \cdot\left(1-2 \pi_{i}\right) \cdot \frac{\partial \mathbf{x}_{i}^{\prime} \beta}{\partial x_{i j}} \cdot \frac{\partial \mathbf{x}_{i} \beta}{\partial x_{i k}} \text { for logit, }
\end{array}\right.
\end{align*}
$$

where $f^{\prime}(\bullet)$ denotes the first derivative of the pdf. In the special case where the linear predictor is linear in $x_{j}$ and $x_{k}$, the first term drops out of the expression (A3) because the cross-partial derivative of $\boldsymbol{x}_{i}^{\prime} \beta$ equals zero. Now, applying (A2) to Nagler's (1991, p. 1399) replication of Wolfinger and Rosenstone's final probit model, as reproduced in the first column of Table 1, we find that the marginal effect of education on the probability of voting is

$$
\begin{equation*}
\phi\left(\mathbf{x}_{i} \hat{\beta}\right) \cdot(.2635+.007 \cdot \text { education })>0 \tag{A4}
\end{equation*}
$$

where $\hat{\beta}$ is the maximum-likelihood estimate of the coefficient vector. This means that education in general has a positive effect on turnout. On the other hand, the marginal effect of closing date on the probability of voting is

$$
\begin{equation*}
\phi\left(\mathbf{x}_{i}, \hat{\beta}\right) \cdot(-.0062)<0, \tag{A5}
\end{equation*}
$$

which means that closing date in general has a negative effect on turnout. To examine the direction of the built-in interaction between closing date and education, we take the cross-partial derivative of $\pi_{i}$ with respect to closing date and education. This yields

$$
\begin{equation*}
-\phi\left(\mathbf{x}_{i}^{\prime} \hat{\beta}\right) \cdot\left(\mathbf{x}_{i}^{\prime} \hat{\beta}\right) \cdot(-.0062) \cdot(.2635+.007 \bullet \text { education }) \tag{A6}
\end{equation*}
$$

It should be obvious that the sign of (A6) is determined by the sign of the term ( $\mathbf{x}_{i} \hat{\beta}$ ). That is, (A6) is positive if $\mathbf{x}_{i}^{\prime} \hat{\beta}>0$ and negative if $\mathbf{x}_{i}^{\prime} \hat{\beta}<0$. When $\mathbf{x}_{i}^{\prime} \hat{\beta}=0$ (or when $\pi_{i}=$ .5), the built-in interaction vanishes. This means that education can ease the negative effect of an early closing date (i.e., make its effect less negative) only when individual $i$ crosses the threshold of $\pi_{i}=.5$.

Based on the probit estimates reproduced in the second column of Table 1, we obtain a more complicated marginal effect of education on the probability of voting:

$$
\begin{align*}
& \phi\left(\mathbf{x}_{i} \hat{\beta}\right) \cdot[.3544-.0058 \bullet \text { education }-.0032 \bullet \text { closing }+.0004  \tag{A7}\\
&\bullet(\text { closing } \bullet \text { education })]>0 .
\end{align*}
$$

On the other hand, the marginal effect of closing date on the probability of voting is:

$$
\begin{align*}
\phi\left(\mathbf{x}_{i} \hat{\beta}\right) & \cdot(.0032-.0032 \bullet \text { education }+.0002  \tag{A8}\\
& \left.\left.\cdot e^{2} e^{2}\right)^{2}\right)
\end{align*}<0, \text { for education }>1 .
$$

Multiplicative terms in the index function of this model make the cross-partial derivative of $\mathbf{x}_{i} \hat{\beta}$ nonzero in general but complicate the two partial derivatives of $\mathbf{x}_{i} \hat{\beta}$ in (A3). Thus, the cross-partial derivative of $\pi_{i}$ with respect to closing date and education becomes:

$$
\begin{align*}
\phi\left(\mathbf{x}_{i}^{\prime} \hat{\beta}\right) & \bullet(-.0032+.0004 \bullet \text { education })-\phi\left(\mathbf{x}_{i}^{\prime} \hat{\beta}\right) \bullet\left(\mathbf{x}_{i} \hat{\beta}\right)  \tag{A9}\\
& \bullet\left(.0032-.0032 \bullet \text { education }+.000 \bullet \text { education }{ }^{2}\right) \\
\bullet(.3544-.0058 & \bullet \text { education }-.0032 \bullet \text { closing }+.0004 \bullet \text { closing } \bullet \text { education }) .
\end{align*}
$$

The sign of (A9) is not as obvious as (A6). We can be certain, however, that it is not always negative. For example, constructing a "worst-case scenario" where the "more educated, more deterred" conclusion is most likely to apply. In this scenario, all the negative-effect variables are set at their highest value (e.g., closing date $=50$; southern state $=1$ ) and all the positive-effect variables, except education, are set at their lowest value (e.g., age $=23$, the youngest age that an ordinary person reaches 5 or more years of college education). In this case, we find that (A9) >0 for education $>7$. This means that even for this extreme case, gaining some graduate education can still reverse the negative effect of closing date on turnout. The farther the values of the independent
variables deviate from this worst scenario, the easier it becomes for (A9) to turn positive. For the typical individual, (A9) $>0$ for education $>5$, even when the variable of closing date is set at its maximum value of 50 .

## NOTES

1. The assumed interactive effect actually vanishes at the midpoint of .5 because there is no need to suppress or to compensate the change rate of probability at this location. Instead, what exercises the greatest impact on the probability at this point is the marginal effect of each independent variable (King, 1989, p. 109).
2. We are indebted to an anonymous reviewer for making these points clear.
3. For a detailed table of relations between the values of latent index and probability in the probit model, as well the values of the first to fourth derivatives of a standard normal distribution, see Beyer (1991, pp. 108-115). In probit analysis, the estimates of this latent index look like the familiar standardized scores (i.e., $z$-scores) given the assumed standard normal distribution. The reason we normalize the distribution is because we are interested in the probability of an event, not the scale of the latent index. Otherwise, the probit model will be underidentified (Maddala, 1983, p. 23). As Amemiya (1981) says, "Once [a researcher] sets up a probit or logit model, he is only concerned with how the probability of an event in question is related to the independent variables. . . . Thus, without loss of generality, one can adopt a normalization: for example, [zero mean and unit variance]" (p. 1489). After all, the term probit is defined as probability unit (Aldrich \& Nelson, 1984, p. 37; Bliss, 1934, pp. 38-39; Finney, 1971, p. 23).
4. These arguments also apply to other functional forms used in models of categorical response. If a researcher has theoretical reasons to employ an alternative functional form, such as scobit (Nagler, 1994), the interactive effects should still be evaluated on the basis of the nonlinear functional form. If they are not, the "built-in" interaction effects resulting from the nonlinear functional form are ignored. Even the scobit model assumes the S-shaped curve. It does add an additional parameter $\alpha$ to allow the curve to bend in different degrees. For all legitimate values of $\alpha>0$, however, the "built-in" interactions based on the nonlinear functional form (i.e., the "ceiling" and "floor" effects) remain. For example, when $\alpha=.42$ (as estimated in Nagler's [1994] scobit model), it only takes 6 unit of the latent index to raise the probability of alternative 1 from .4 to .5 and yet 1.7 such units to raise the probability from .8 to .9 . Actually, this is a central point of the underlying assumptions of the nonlinear models of logit, probit, and scobit. In addition, graphical displays of the interactive effects will greatly facilitate interpretation regardless of the chosen functional form. We focus this article on logit and probit because of their immense popularity and because Nagler's scobit estimates (1994, p. 251) lead to the same substantive conclusions as his earlier paper (1991).
5. The maximum closing date in 1992 was 30 compared to 50 in 1972. Also, the coding of education for the 1992 Current Population Study is somewhat different from that of the 1972, that is, the three lowest education levels in the 1972 data are recoded into two groups in order to reflect the substantial improvement of education in the United States during the two decades.
6. The aggregation method used by Wolfinger and Rosenstone (1980), called sample enumeration (Ben-Akiva \& Lerman, 1985, pp. 146-148; Train, 1986, pp. 99-101), is sound for estimating "the aggregate marginal effect of a single variable on the probability of a particular sub-
population voting" (p. 123, italic added). Nevertheless, dividing the entire sample into subsamples according to one variable creates a homogeneous subsample only in that particular variable but remains heterogeneous in all other variables. The combination of such within-group heterogeneity and the sensitivity of nonlinearity makes it practically impossible to disentangle disaggregate-level effects based on aggregates.
7. In addition, there exists a rich but more advanced literature dealing with semiparametric estimation of the derivatives that also demonstrates the inherent difficulties in disentangling the link functional form and the linear index functions (e.g., Lee, 1996; Powell, Stock, \& Stoker, 1989; Stoker, 1986).

## REFERENCES

Aldrich, J. H., \& Nelson, F. D. (1984). Linear probability, logit, and probit models (Sage University Paper Series on Quantitative Applications in the Social Sciences, Series No. 07-045). Beverly Hills, CA: Sage.
Amemiya, T. (1981). Qualitative response models: A survey. Journal of Economic Literature, 19(4), 1483-1536.
Ben-Akiva, M., \& Lerman, S. R. (1985). Discrete choice analysis: Theory and application to travel demand. Cambridge: MIT Press.
Beyer, W. H. (1991). CRC standard probability and statistics: Tables and formulae. Boca Raton, FL: CRC Press.
Bliss, C. I. (1934). The method of probits. Science, 79(2037), 38-39.
Brians, C., \& Grofman, B. (1994, September). Voter registration laws and turnout in the U.S. states: A longitudinal analysis. Paper presented at the annual meeting of the American Political Science Association, New York.
Cox, D. R., \& Snell, E. J. (1989). Analysis of binary data (2nd ed.). London: Chapman and Hall.
Finney, D. J. (1971). Probit analysis (3rd ed.). London: Cambridge University Press.
Fox, J. (1987). Effect displays for generalized linear models. In C. C. Clogg (Ed.), Sociological methodology (pp. 347-361). Washington, DC: American Sociological Association.
Hanushek, E. A., \& Jackson, J. E. (1977). Statistical methods for social scientists. New York: Academic Press.
Highton, B. (1997). Easy registration and voter turnout. Journal of Politics, 59(2), 565-576.
King, G. (1989). Unifying political methodology: The likelihood theory of statistical inference. Cambridge, UK: Cambridge University Press.
Lee, M. (1996). Methods of moments and semiparametric econometrics for limited dependent variables. New York: Springer.
Maddala, G. S. (1983). Limited-dependent and qualitative variables in econometrics. Cambridge, UK: Cambridge University Press.
Nagler, J. (1991). The effect of registration laws and education on U.S. voter turnout. American Political Science Review, 85(4), 1393-1405.
Nagler, J. (1994). Scobit: An alternative estimator to logit and probit. American Journal of Political Science, 38(1), 230-255.
Powell, J. L., Stock, J. H., \& Stoker, T. M. (1989). Semiparametric estimation of index coefficients. Econometrica, 57(6), 1403-1430.
Stoker, T. M. (1986). Consistent estimation of scaled coefficients. Econometrica, 54(6), 1461-1481.

Train, K. (1986). Qualitative choice analysis: Theory, econometrics, and an application to automobile demand. Cambridge: MIT Press.
Wolfinger, R. E., \& Rosenstone, S. J. (1980). Who votes? New Haven, CT: Yale University Press.

Chi Huang is a professor and Chairman of the Department of Political Science, National Chung-Cheng University, Taiwan, Republic of China.

Todd G. Shields is an associate professor of political science at the University of Arkansas and associate director of the Fulbright Institute of International Relations.

