SAGE Research Methods Foundations

Marginal Effects and Adjusted Predictions

Contributors: Dustin S. Stoltz, Richard A. Williams
Pub. Date: 2020
Product: SAGE Research Methods Foundations
DOI: https://dx.doi.org/10.4135/9781526421036939917
Methods: Independent variables, Prediction, Dependent variables
Keywords: age and aging
Disciplines: Anthropology, Business and Management, Communication and Media Studies, Computer Science, Counseling and Psychotherapy, Criminology and Criminal Justice, Economics, Education, Engineering, Geography, Health, History, Marketing, Mathematics, Medicine, Nursing, Political Science and International Relations, Psychology, Social Policy and Public Policy, Science, Social Work, Sociology, Technology
Access Date: November 8, 2022
Publishing Company: SAGE Publications Ltd
City: London
Online ISBN: 9781529750874

Marginal effects and adjusted predictions are means for providing insights into how important effects really are. Adjusted predictions are expected values of a dependent variable computed from the results of a regression, where all independent variables are held at specified values. A marginal effect is the change in the predicted value of a dependent variable after changing one independent variable—either a discrete change in categorical variables or an instantaneous change in continuous variables—while all other variables are held at specified values. Comparing predicted values and marginal effects is a tool for summarizing, interpreting, and testing the significance of independent variables.

While the coefficients of the simplest linear models tend to be easy to understand in substantive terms, the models' underlying assumptions are often not met by the nature of the data. The usual techniques for relaxing these assumptions, however, often incur costs in interpretability. For example, when independent variables are polynomials like powers or interactions or when the dependent variable is categorical, calculating, and comparing predicted values can illuminate the practical importance of each coefficient.

In addition to helping understand the practical significance of complex models, marginal effects can aid in testing the statistical significance of interaction terms in nonlinear models where it is often tempting, but erroneous, to rely on the *p* value of the interaction term in regression outputs (Mize, 2019). Furthermore, marginal effects, as they are calculated in terms of the natural (i.e., untransformed) metric of the dependent variable, can be used to compare groups in nonlinear models (Long & Mustillo, 2018).

After providing an overview of marginal effects and adjusted predictions, this entry provides computational details and discusses statistical significance, using the example of voter turnout.

Overview

Consider, as a simple example, predicting the values of an outcome variable at different ages. Take the following regression equation:

$$y = \beta_0 + \beta_{age} age + \varepsilon$$

Here, age is a continuous variable, and ε is the error term. If age has a linear effect, we would expect each additional year to matter the same when someone is 21 as when they are 81. To calculate the marginal effect of an additional year is straightforward: Calculate *y* when age is zero, and then subtract it from *y* when age is one. If all variables are continuous and linear, the marginal effect is the partial derivative of the regression function with respect to the independent variable of interest, which would be equivalent to the slope coefficient

(Cameron & Trivedi 2010).

What if, however, the effect of age is nonlinear?

$$y = \beta_0 + \beta_{age}age + \beta_{age^2}age^2 + \varepsilon$$

In the aforementioned equation, the effect of age is allowed to differ when someone is 21 as when they are 81. This makes calculating and interpreting the effects of the coefficients a little trickier. We cannot know from the coefficients alone what the predicted value for a 21-year-old would be but must instead plug “21” into the equation and solve it.

To make matters more complicated, age might interact with class: The effect of age for lower or working-class person of a certain age might be different than for middle- or upper-class person of the same age:

$$y = \beta_0 + \beta_{age}age + \beta_{age^2}age^2 + \beta_{class}class + \beta_{(class * age)}(class * age) + \varepsilon$$

Finally, if we are predicting whether someone voted ($y = 1$) or did not vote ($y = 0$) in a presidential election (i.e., a discrete choice), the model might be

$$\Pr(vote = 1) = F\left(\beta_0 + \beta_{age}age + \beta_{age^2}age^2 + \beta_{class}class + \beta_{(class * age)}(class * age)\right)$$

As the outcome is binary, and thus the model is nonlinear regardless of the independent variables, we would likely use logistic regression (and the inverse logit function) or else probit. Here the coefficients are log odds, which are difficult to understand intuitively. While authors often exponentiate the coefficients (i.e., convert them to “odds ratios” for better interpretation), odds ratios can be misleading because it is tempting, but incorrect, to treat them as relative risk. In each of these cases, understanding the practical significance of coefficients is difficult, and thus calculating adjusted predictions and marginal effects using statistical software is often necessary.

Example 1: Voter Turnout by Age and Class

Using the 1972– 2018 General Social Survey from the United States, and whether a respondent said they voted in a presidential election as the outcome variable, we first estimate a logistic regression for the model (shown as both log odds and odds ratios—that is, the exponentiated log odds—in Table 1).

The .62 positive coefficient for class (coded 0 = *lower/working class*, 1 = *middle/upper class*) indicates that those in the upper classes are more likely to vote than those in the lower social classes—but it does not make clear how much more likely they are. The positive .12 coefficient for age has the same limitations. The presence of squared terms and interactions further muddles the ease of interpretation.

As we are interested in the more intuitive *probability* that a respondent would vote (rather than, e.g., the log odds of voting), we need to calculate predicted probabilities. We can then easily visualize the predicted probability of voting, *P(voting)*, by age (Figure 1, left panel). Then, we show the same but disaggregated by whether the respondent self-identifies as lower/working class or middle/upper class (Figure 1, right panel). The plot shows that the effect of age is nonlinear, and the effect of class is roughly consistent across age—the difference between the predicted values for lower/working class and those for middle/upper class (right panel) is the marginal effect for class at different ages.

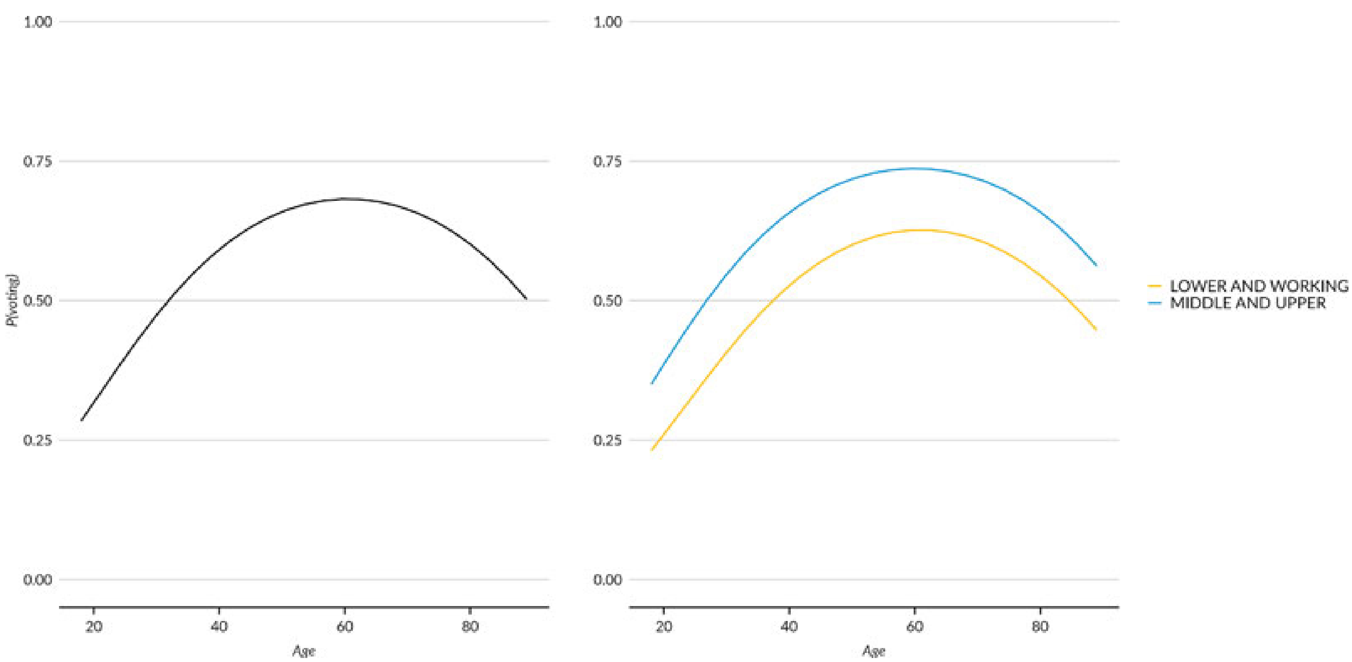
Table 1. Logit results for voting in presidential elections by age and class.

	Log odds	**Odds ratio**
Age	0.1153	1.1222
Class	0.6152	1.8501
Age^2	−0.0009	0.9991
Age × class	−0.0017	0.9983
(Intercept)	−3.5604	0.0284

N = 60,999

More specifically, the left-hand side of Figure 1 shows that 20-year-olds have 31.7% predicted probability of voting. The probability of voting then rises with age, peaking at 68.2% at age 60, and then gradually declines after that. The right-hand side of Figure 1 further shows that middle- and upper-class voters are consistently 10–14 percentage points more likely to vote than their similarly aged lower- and working-class counterparts. The predicted probabilities in Figure 1 provide a far clearer picture of the effects of age and social class on voting than the logistic regression coefficients did.

Figure 1. Predicted probabilities of voter turnout in presidential elections.



We could also summarize the impact of each predictor variable (age and class), on the outcome variable (voting), by calculating their respective average marginal effect (AME). Here, the marginal effect of being "middle or upper class" as opposed to "lower or working class" is 0.1227. That is, on average, middle- or upper-class individuals are 12.27 percentage points more likely to vote than are members of the lower or working class. Again, the original coefficients indicated there were class differences in the likelihood of voting, but the AME made it much clearer how great those differences were.

With the categorical variable class, which has only two possible values, the AME is easy to interpret. It is simply the difference in the predicted values for the two groups. For continuous variables like age, which can potentially take on an infinite number of values, AMEs may be less useful (Cameron & Trivedi, 2010). Other methods, such as adjusted predictions and marginal effects at representative values, may be better choices.

Computational Details

The way marginal effects and adjusted predicted are computed depends on whether the independent variable is discrete or continuous, and also on how other covariates are held constant.

Discrete Changes Versus Instantaneous Rate of Change

As stated previously, a marginal effect is what happens to a dependent variable when one independent variable changes, while the others do not. There are two ways a variable can change. An *instantaneous change* is an infinitely small change in the independent variable, which is often used when the variable of interest is continuous. A *discrete change* (or first difference) is a finite change in the independent variable, usually used for binary or categorical predictors but can be used for any kind of variable, for any amount.

As some consider "marginal" to be synonymous with "instantaneous," marginal effects obtained using discrete change are sometimes referred to as "partial effects," reserving "marginal effects" for only those predictions obtained using instantaneous rate of change. However, to make matters more confusing, as an instantaneous change in a continuous variable corresponds to an (infinitely small) portion of the unit of the variable, this has also been referred to as "partial change" in contrast to "discrete change" (see e.g., Long 1997).

In the example provided, if we use the instantaneous rate of change, we are presuming that age is infinitely divisible. However, age could also be measured using the discrete change method, presuming it is divided into finite units like days, months, years, or decades. More generally, an instantaneous change is a discrete change as the amount of discrete change approaches zero. How linear the predicted probability curve is determines how similar the discrete change and the instantaneous change will be. Discrete changes tend to be easier to interpret. Nevertheless, the default for most margins packages for both Stata and R statistical software is to use instantaneous change for continuous variables and discrete change for categorical variables.

Holding Covariates Constant

As mentioned, marginal effects of an independent variable on the dependent variable and predicted values are determined by the value of the independent variable (in the literature, this is usually denoted as x_k) and the values of all other variables held at some specified level (often denoted as x^* or a bold $\mathbf{x}$). In other words, we are looking at what happens when one variable changes while the others do not.

In the first panel of Figure 1, the simplicity of the model allows us to visualize predicted probabilities of voting for all values of age. But, class is also in the model, therefore at what value is "class" being held constant? This is a problem for nonlinear models, as the value of the marginal effect will depend on the specific values of all of the independent variables (Long & Freese, 2006). There are generally three methods to determine these values (Cameron & Trivedi, 2010; Williams, 2012). Similar approaches are used for computing different types of adjusted predictions.

Marginal effect at the mean (MEM) is perhaps the most common method of determining these values, or at least it was before advances in computing software made other approaches more accessible. This entails holding covariates at their mean values (in the case of binary variables like gender, they will be held at their sample proportions). Although the MEM is simple to compute and understand, it is sometimes criticized for using a "typical" case that does not exist among the observations and perhaps cannot exist (Long, 1997; Williams, 2012), for example a person who is 42.3 years old and 55% middle/upper class and 42% male.

An increasingly popular alternative is the AME. An adjusted prediction and marginal effect are computed using the observed values *for each case*. These values are averaged across cases, giving the average adjusted prediction and the average marginal effect. Richard Williams (2012) detailed how the calculations are done and argued that AMEs are superior to MEMs because they use *all* of the information available on the variables, not just their means.

Still, like MEM, AME is a summary measure where the average size of the marginal effect may not be close to any actual observation's marginal effect (Long, 2014). Furthermore, AME tends to be more complex to calculate and, depending on the data, MEM is sometimes considered an adequate approximation of AME.

While promoting the advantages of AMEs and MEMs, Williams (2012) also noted

> The biggest problem with both [AMEs and MEMS] may be that they only produce a single estimate of the ME. No matter how "average" is defined, averages can obscure differences in effects across cases. In reality, the effect that variables like race have on the probability of success varies with the characteristics of the person; for example, racial differences could be much greater for older people than for younger. (p. 326)

To deal with these concerns, Williams (2012) recommended calculating *marginal effects at representative values* (MERs), MERs are similar to MEMs, but as the name implies, the analyst manually chooses (usually substantively or theoretically interesting) values for each covariate, rather than the mean. Williams (2012) said "With MERs, you choose ranges of values for one or more independent variables and then see how the MEs differ across that range. MERs can be intuitively meaningful, while showing how the effects of variables vary by other characteristics of the individual" (p. 326). Williams (2012) presented a logistic regression where the dependent variable, diabetes, is positively affected by age: Older people are more likely to have diabetes than are younger people. In this example, he showed that the AME for race was .04; that is, on average, Blacks were 4 percentage points more likely to have diabetes than were Whites. However, this average number obscured a great deal of racial variability by age. At age 20, the racial Black/White difference in the likelihood of having diabetes was less than 1 percentage point. This is not surprising because, in the 1970s sample used, before diabetes in the United States rates started skyrocketing, hardly any 20-year-olds in the study had diabetes, regardless of their race. But, the racial gap gradually widened with age, with the gap being nearly nine percentage points by age 70.

Example 2: Voter Turnout by Age, Class, and Gender

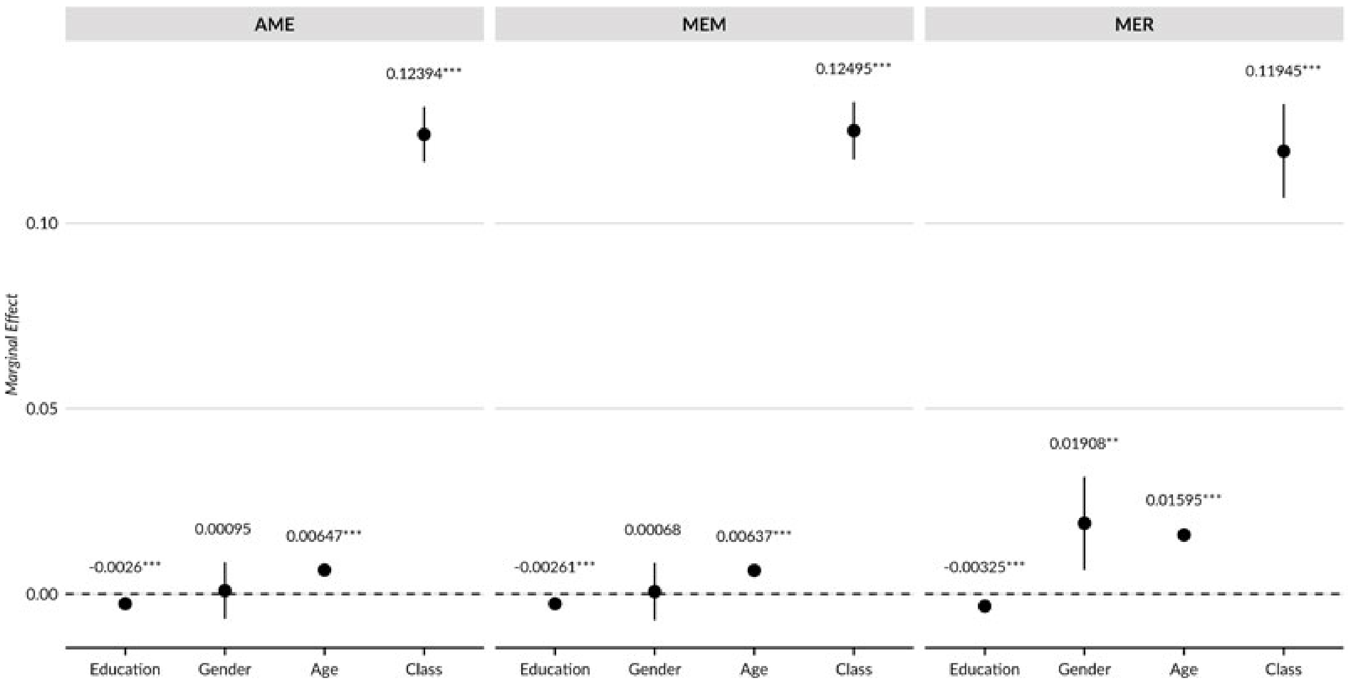
Table 2. Logit results for voting in presidential elections by age, class, gender, and education.

	Log odds	**Odds ratio**
Age	0.1170	1.1241
Class	0.6773	1.9685
Gender	0.2691	1.3088
Education	−0.0196	0.9806
Age^2	−0.0009	0.9991
Age × class	−0.0019	0.9981

Age × gender	−0.0031	0.9969
Age × education	0.0002	1.0002
Class × gender	−0.0840	0.9194
(Intercept)	−3.6288	0.0265
N = 60,999		

Again using the 1972–2018 General Social Survey and whether a respondent said they voted in a presidential election as the outcome variable, we estimate a more complicated logistic regression using age, class, gender, and education (highest year of school completed) as predictor variables and allowing all variables to vary with age, and interacting class and gender (shown as both log odds and odds ratios in Table 2).

Figure 2. Marginal effects using AMEs, MEMs, and MERs.

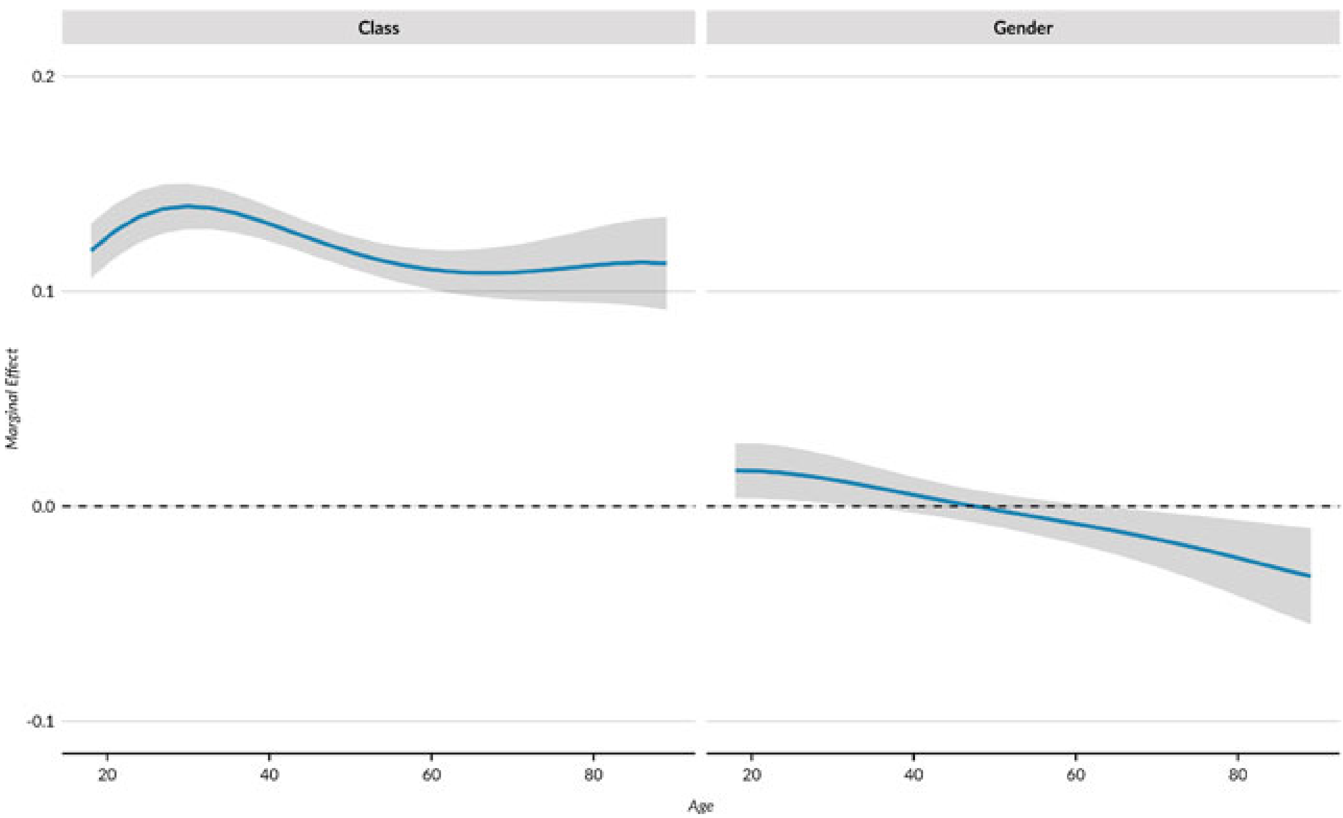


Note: 95% confidence intervals. Representative value in the MER is age at 18.

We then calculate MEMs, AMEs, and MERs for each variable (Figure 2). We see that, even in this more complicated model, with different ways of computing the marginal effects, the upper social classes are still about 12 percentage points more likely to vote than the lower social classes. We also see that, for these data, the difference between MEMs and AMEs for every variable in this model are very small. For the MER, we set age to 18 and calculate the average marginal effects for all variables, revealing that the marginal effect of gender is small (less than 2 percentage points) but statistically significant for this age-group.

Next, we plot the predicted probabilities at representative values; specifically, we show the marginal effects of class and gender for different values of age (Figure 3). We can see that, for all ages, the marginal effect of class is always above 10 percentage points and significantly different from zero, but this is not so for gender. The marginal effect of gender is much smaller than the marginal effect for class and is not distinguishable from zero from roughly 40–60 years old.

Figure 3. Marginal effect of class and gender by age.



Note: 95% confidence intervals.

Statistical Significance

This entry has focused primarily on ways to assess the substantive significance of results; but statistical significance is usually assessed too. The following subsections discuss how that is done with marginal effects and adjusted predictions, and how group differences and interaction effects can then be tested.

Confidence Intervals

The previous examples and discussion covered adjusted predictions and marginal effects to assess the practical significance of coefficients in regression models (especially nonlinear models). However, adjusted predictions can also be used to assess statistical significance of group differences and interaction terms in nonlinear models, the first step of which is to estimate standard errors of predicted values and marginal effects. This is fairly easy to see in the visualization of the marginal effect of gender (Figure 3, right panel). Where the confidence interval contains zero, the difference in the marginal effect of women compared to men is not statistically significant.

The delta method is the most common means of estimating standard errors and the default method used by both the Stata margins package (Pitblado, 2014), and most marginal effects packages in R (Leeper, 2017). Roughly, this process entails finding a linear approximation of the more complex nonlinear function of the model (Long & Freese, 2006). The variance of this approximation is then used to construct confidence intervals for the predicted values.

One other method deserves consideration: bootstrapping (Efron & Tibshirani, 1994; Long & Freese, 2006). By repeatedly sampling from the data used in the model, we can estimate the standard deviation of the sampling distribution that would occur if we took repeated samples from the population (Long & Freese, 2006). While bootstrapping may produce more accurate estimates of variance, the delta method is generally much less computationally intensive to calculate and the two are otherwise very similar (Dowd et al., 2014).

Significance of Group Differences and Interaction Terms

Once we have estimated standard errors for marginal effects, we can easily test group differences using a Wald test (Long & Mustillo, 2018). This entails taking the difference of the marginal effects between two groups, divided by the square root of the sum of the variance of each group's marginal effects minus their covariance. For example, following the work of Trenton D. Mize (2019), if we are interested in whether the effect of age on the probability of voting differs for men and women: Formally, $\Delta_{age_{women}}$ and $\Delta_{age_{men}}$ are the AME of age for women and men, respectively. Similarly, $\sigma^2_{age_{women}}$ and $\sigma^2_{age_{women}}$ are the estimated variance of each marginal effect. Therefore, to conduct a Wald test to determine whether the AME of age differs between men and women, we would use the following:

$$z = \frac{\hat{\Delta}_{age_{women}} - \hat{\Delta}_{age_{men}}}{\sqrt{\hat{\sigma}^2_{age_{women}} + \hat{\sigma}^2_{age_{men}} + \sigma^2_{age_{women}, age_{men}}}}$$

The same procedure can be used with MEM. However, both AME and MEM are summaries of marginal effects. As Figure 3 shows, the effect of age for men and women is nonlinear: At some points, the confidence interval contains zero. Therefore, the analyst may also use MERs, for example, whether the marginal effects for 20-year-old women significantly differ from 20-year-old men (for a comparison of using AMEs and MERs to test group differences, see Long & Mustillo, 2018).

For a variety of reasons, the editors of the *American Sociological Review* (Mustillo et al., 2018) stated that researchers should not "use the coefficient on the interaction term to draw conclusions about the significance of statistical interaction in categorical models" (p. 1282; see also Allison, 1999; Mood, 2010; Williams, 2009). Techniques such as those described in this entry provide one means of drawing such conclusions. Using the same procedure for testing group differences, marginal effects, and adjusted predictions can be used to test the statistical significance of interactions in nonlinear models (Mize, 2019).

Conclusion

When presenting results from regression analyses, researchers in many fields often focus on the signs and statistical significance of their models' coefficients. While commonly done, this approach is often far less informative than it could be. With nonlinear models like logistic regression, the actual effect of a variable on the probability of an event occurring is not easily ascertained from the model coefficients alone. Even in ordinary least squares regression, the effects of variables can become much more complicated to understand once squared terms and interactions are added to a model. Coefficients alone can also obscure differences across groups by indicating that an interaction term is or is not statistically significant, missing the fact that group differences may be significant across some values of the independent variables but not significant at others.

The prudent use of adjusted predictions and marginal effects can help alleviate these limitations. Instead of saying, for example, that the effect of social class is .62 in a logistic regression, one can show that this translates into the more intuitive finding that social classes differ by about 12 percentage points in their likelihood of voting. Instead of saying that, on average, Blacks are about 4 percentage points more likely to experience an event than are Whites, one can say that racial differences are very small at young ages, but the racial gap greatly widens as people get older. Instead of making a blanket statement that group differences involving gender are or are not statistically significant, one can examine whether gender differences are significant at some values of age but not others. With advances in statistical software making the computation of adjusted predictions and marginal effects far easier than it used to be, these methods may make the practical and substantive importance of researchers' work clearer and more forceful.

References

Allison, P. D. (1999). Comparing logit and probit coefficients across groups. Sociological Methods & Research, 28(2), 186–208. https://doi.org/10.1177/0049124199028002003

Cameron, A. C., & Pravin, K. T. (2010). *Microeconometrics using Stata* (Rev. ed.). Stata Press.

Dowd, B. E., William, H. G., & Edward, C. N. (2014). Computation of standard errors. Health Services Research, 49(2), 731–750. https://doi.org/10.1111/1475-6773.12122

Efron, B., & Tibshirani, R. J. (1994). *An introduction to the bootstrap*. CRC Press.

Leeper, T. J. (2017). *Interpreting regression results using average marginal effects with R's margins*. Comprehensive R Archive Network.

Long, J. S. (1997). *Regression models for categorical and limited dependent variables* (Advanced Quantitative Techniques in the Social Sciences, Vol. 7). SAGE Publications.

Long, J. S. (2014). Regression models for nominal and ordinal outcomes. In H.Best & C.Wolf (Eds.), The SAGE handbook of regression analysis and causal inference (pp. 173–204). SAGE Publications. https://doi.org/10.4135/9781446288146

Long, J. S., & JeremyF. (2006). *Regression models for categorical dependent variables using Stata* (2nd ed.). Stata Press.

Long, J. S., & Mustillo, S. A. (2018). Using predictions and marginal effects to compare groups in regression models for binary outcomes. Sociological Methods & Research. https://doi.org/10.1177/0049124118799374

Mize, T. D. (2019). Best practices for estimating, interpreting, and presenting nonlinear interaction effects. *So-*

ciological Science, 6, 81–117.

Mood, C. (2010). Logistic regression: Why we cannot do what we think we can do, and what we can do about it. European Sociological Review, 26(1), 67–82. https://doi.org/10.1093/esr/jcp006

Mustillo, S. A., Omar, A. L., & Rory, M. M. (2018). Editors' comment: A few guidelines for quantitative submissions. American Sociological Review, 83(6), 1281–1283. https://doi.org/10.1177/0003122418806282

Pitblado, J. (2014). How are average marginal effects and their standard errors computed by margins using the delta method?Stata FAQ. Retrieved October 21, 2019, from https://www.stata.com/support/faqs/statistics/compute-standard-errors-with-margins/

Williams, R. (2009). Using heterogeneous choice models to compare logit and probit coefficients across groups. Sociological Methods & Research, 37(4), 531–559. https://doi.org/10.1177/0049124109335735

Williams, R. (2012). Using the margins command to estimate and interpret adjusted predictions and marginal effects. The Stata Journal, 12(2), 308–331. https://doi.org/10.1177/1536867X1201200209

https://dx.doi.org/10.4135/9781526421036939917