



ginteff: A generalized command for computing interaction effects

Marius Radean
Department of Government
University of Essex
Colchester, U.K.
mradean@essex.ac.uk

Abstract. Interaction analyses are useful tools to examine complex socioeconomic outcomes in which the effect of one variable depends on the presence or values of another variable. Interaction effects capture simultaneous changes in two (or more) covariates, and their computation is especially challenging in nonlinear models. For such models, a statistically significant interaction-term coefficient does not necessarily indicate significant interactive effects. For analyses in which the interaction effect cannot be inferred from the model estimates, I introduce **ginteff**, a new command that automatically computes two- and three-way interaction effects. The command accommodates a large suite of estimation models and allows researchers to use either the partial derivative or the first difference to model the effect of the interacted variables.

Keywords: st0711, ginteff, ginteffplot, average interaction effect, individual-level interaction effects, two-way interactions, three-way interactions

1 Introduction

Interaction analyses are used to examine complex socioeconomic outcomes in which the effect of one variable depends on the presence or values of another variable.¹ For example, to assess whether the 2008 financial crisis had a more pronounced effect on minorities, in a regression analysis, we would interact (that is, multiply) the indicator of minority status with that of the before- and after-2008 periods. In linear models, the coefficient on the interaction term can be used to infer whether the effect of the treatment variable is statistically different at alternative values of the moderating variable. For instance, if the interaction coefficient were to reach the conventional levels of statistical significance, we would conclude that the 2008 crisis had a statistically different effect on majority and minority groups. In nonlinear models, however, the coefficient on the interaction term does not tell us the direction, magnitude, or significance of the interaction effect (Ai and Norton 2003). For analyses in which the interaction effect cannot be inferred from the model estimates, I introduce **ginteff**, a new command that automatically computes two- and three-way interaction effects. The command

1. This type of analysis is particularly popular in social sciences (for example, economics, sociology, political science, psychology). For instance, in political science, roughly a quarter of published studies contain interactive hypotheses (Kam and Franzese 2007, 7–8).

accommodates a large suite of estimation models and allows researchers to use either the partial derivative or the first difference to model the effect of the interacted variables.

2 What is an effect and how do we calculate it?

Broadly speaking, the interaction effect is the change in the effect of a given variable as another variable also changes. Before considering two simultaneous changes, I briefly review what an unconditional effect is and how it is computed. Let us say our dependent variable y depends on the independent variable of interest, x , and a vector of other covariates plus the constant term, \mathbf{X} . Conditional on the independent variables, the predicted value of y is

$$\Pr(y|x, \mathbf{X}) = F(\beta_x x + \beta \mathbf{X})$$

where $F(\cdot)$ is the (possibly nonlinear) link function of model predictors.² The effect of x on y is the change in $\Pr(y)$ attributable to a change in x . There are two general approaches to computing the effect of x , which I denote by $\Delta(x)$. One alternative is to calculate the first difference, which is the change in $\Pr(y)$ associated with an n -unit increase in x (frequently a one-unit increase), $\Delta(x) = \{\Pr(y|x+n) - \Pr(y|x)\}$. This is the default approach for factor variables, in which case the effect is the discrete change from the base level. For example, if x is a dummy variable, the discrete difference is $\Delta(x) = \Pr(y|x=1) - \Pr(y|x=0)$.

For continuous variables, researchers can alternatively compute the instantaneous rate of change, which is the partial derivative with respect to x , $\Delta(x) = \{\partial \Pr(y)\}/(\partial x)$. This estimate can then be used to calculate the impact on y of a very small increase, say, 0.001, in x .³ In this case, $\Pr(y)$ would increase by about $0.001 \times \Delta(x)$. In practice, however, many analysts extrapolate and interpret the value of $\Delta(x)$ as representing the change in y associated with a one-unit increase in x . In many (but not all) instances, this is a good approximation. In nonlinear models, there is no guarantee that a one-unit increase in x would lead to a change in y of $1 \times \Delta(x)$. In fact, substantive deviations are likely when x is measured in large units (Williams 2012, 2021).

Turning to interactions, let us say we have a multiplicative model where two independent variables, x_1 and x_2 , are interacted. In this case, the predicted value of y is

$$\Pr(y|x_1, x_2, \mathbf{X}) = F(\beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta \mathbf{X})$$

One can use either the partial derivative or the first difference to compute the simultaneous changes in x_1 and x_2 . Specifically, the interaction effect can be computed as the cross-partial derivative with respect to both variables, $\{\partial^2 \Pr(y)\}/(\partial x_1 \partial x_2)$, or as the discrete difference between two first differences, $\{\Pr(y|x_1+n_1; x_2+n_2) - \Pr(y|x_1; x_2+n_2)\} - \{\Pr(y|x_1+n_1; x_2) - \Pr(y|x_1; x_2)\}$. Technically, they are both valid approaches. From a purely practical perspective, the first-difference approach has several advantages.

2. If the link function is strictly linear (as in the case of the standard linear regression), the expression simplifies to $\Pr(y|x, \mathbf{X}) = \beta_x x + \beta \mathbf{X}$.

3. As a rule of thumb, Cameron and Trivedi (2022, 620–622) suggest using the standard deviation of the variable divided by 1,000 as a value for this “very small” amount.

First, it does not require complex math, because we need to compute predicted probabilities only for alternative values of x_1 and x_2 . By contrast, taking the cross-partial derivative is a challenging task even for relatively simple likelihood functions and may be intractable for complex ones. Second, it is easier to explain what the interaction effect represents in substantive terms when using clearly defined increments (for example, a one-unit increase). By contrast, partial effects reference an undefined “very small” amount.⁴

3 The ginteff command

3.1 Description

`ginteff` computes the average and individual-level interaction effects for two- and three-way interactions. The effect of the interacted variables can be computed via either the partial derivative or the first difference.

3.2 Syntax

```
ginteff [if] [in] [weight], effect_computation [options]
```

`fweights`, `awweights`, `iweights`, and `pweights` are allowed; see [U] 11.1.6 `weight`.

4. One question that remains open with the first-difference approach is the starting value for continuous variables. Unlike factor variables, these do not have a natural or predefined base level. Put differently, we evaluate an n -unit increase from what value? In practice, it is up to the researcher to choose a reasonable baseline. This can be the variable’s observed values, its mean, or another theoretically informed value. Often touted as a limitation of the first-difference approach, this modeling decision also affects partial derivative estimates. To make it more concrete, we would obtain different results if we were to type `margins, dydx(x)` or `margins, dydx(x) at((mean) x)`, in which x is set at its observed values in the first case and at its mean in the second case.

3.3 Options

Short descriptions of the `ginteff` options are presented in table 1. For the extended descriptions, see *The ginteff User's Manual*.

Table 1. The `ginteff` options

<i>effect_computation</i>	Description
* <code>dydxs(dxspec)</code>	specify the interacted variables for which to compute the effect via partial derivative
<code>fd(fdspec)</code>	shorthand for <code>firstdiff()</code>
* <code>firstdiff(fdspec)</code>	specify the interacted variables for which to compute the effect via first difference
* One of <code>dydxs()</code> or <code>firstdiff()</code> is required. A minimum of two and a maximum of three variables must be specified in <code>dydxs()</code> or <code>firstdiff()</code> .	
<i>options</i>	Description
<code>atdxs(atdxspec)</code>	fix the interacted variables in <code>dydxs()</code> to specified values
<code>nunit(# varlist)</code>	specify the unit increase for each variable in <code>firstdiff()</code>
<code>obseff(stub)</code>	create new variables with the interaction effect for each observation
<code>at(atspec)</code>	compute the interaction effect at specified values of covariates
<code>intequation(eqno)</code>	identify the interaction equation; default is <code>intequation(#1)</code>
<code>level(#)</code>	set confidence level; default is <code>level(95)</code>
<code>many</code>	report more than 100 results; maximum is 1,000
<code>nolegend</code>	suppress output legend
<code>noweights</code>	ignore weights specified in estimation
<code>post</code>	post interaction effects and their variance–covariance estimate as estimation results
<code>predict(pred_opt)</code>	compute the interaction effect for <code>predict</code> , <code>pred_opt</code>
<code>vce(vcetype)</code>	specify how the variance–covariance estimate and standard errors are calculated; default is <code>vce(delta)</code>

3.4 Stored results

`ginteff` stores the following in `r()`:

Scalars

<code>r(N)</code>	number of observations
<code>r(N_psu)</code>	number of sampled primary sampling units, survey data only
<code>r(N_strata)</code>	number of strata, survey data only
<code>r(df_r)</code>	variance degrees of freedom, survey data only
<code>r(level)</code>	confidence level of confidence intervals

Macros

<code>r(cmd)</code>	<code>ginteff</code>
<code>r(cmdline)</code>	command as typed
<code>r(est_cmd)</code>	<code>e(cmd)</code> from original estimation results
<code>r(est_cmdline)</code>	<code>e(cmdline)</code> from original estimation results
<code>r(fdstat)</code>	the <code>firstdiff()</code> specification
<code>r(model_vce)</code>	<code>vcetype</code> from estimation command
<code>r(obseff)</code>	the list of new variables created because of the <code>obseff()</code> option
<code>r(vce)</code>	<code>vcetype</code> specified in <code>vce()</code>
<code>r(atstat)</code>	the <code>at()</code> specification

Matrices

<code>r(at)</code>	matrix of values from the <code>at()</code> option
<code>r(b)</code>	the interaction effect estimates
<code>r(fd)</code>	matrix of values from the <code>firstdiff()</code> option
<code>r(ginteff)</code>	matrix containing the average interaction effects with their standard errors, test statistics, p -values, upper and lower confidence limits, and critical values
<code>r(nunit)</code>	matrix of values from the <code>nunit()</code> option
<code>r(V)</code>	variance-covariance matrix of the interaction effect estimates

`ginteff` with the `post` option also stores the following in `e()`:

Scalars

<code>e(N)</code>	number of observations
<code>e(N_psu)</code>	number of sampled primary sampling units, survey data only
<code>e(N_strata)</code>	number of strata, survey data only
<code>e(df_r)</code>	variance degrees of freedom, survey data only

Macros

<code>e(cmd)</code>	<code>ginteff</code>
<code>e(properties)</code>	<code>b V</code>

Matrices

<code>e(b)</code>	estimates
<code>e(V)</code>	variance-covariance matrix of the estimates

Functions

<code>e(sample)</code>	marks estimation sample
------------------------	-------------------------

4 The `ginteffplot` command

4.1 Description

`ginteffplot` graphs the results of the immediately preceding `ginteff` command.

4.2 Syntax

```
ginteffplot [ , options ]
```

4.3 Options

Short descriptions of the `ginteffplot` options are presented in table 2. For the extended descriptions, see *The ginteffplot User's Manual*.

Table 2. The `ginteffplot` options

<i>options</i>	Description
Main	
<code>aiepoint</code> ([#clockposstyle] ["text for label"] [, marker_label_options marker_options])	customize marker for the point estimate of the average interaction effect
<code>aierange</code> ([line_options] [msize(markersizestyle)])	customize the range plot of the average interaction effect using capped spikes
<code>obseff</code> [(obseff_options)]	plot observation-level interaction effects
<code>output</code> (#)	identify a specific <code>ginteff</code> output to be graphed; default is <code>output(1)</code>
<code>save</code> (newfile.suffix[, export_options])	export current graph
<code>xcommon</code> ([numlist] [*])	give <i>x</i> -axes common scale
<code>zeroline</code> [(line_options)]	add a vertical line at the 0 value
<i>x</i> and <i>y</i> axes	
<code>xtitle</code> (axis_title)	customize <i>x</i> -axis title
<code>ytitle</code> (axis_title)	specify <i>y</i> -axis title
<code>xlabel</code> (rule_or_values)	customize ticks and labels for <i>x</i> axis
<code>ylabel</code> (rule_or_values)	specify ticks and labels for <i>y</i> axis
<code>xscale</code> (axis_suboptions)	customize how <i>x</i> axis looks
<code>yscale</code> (axis_suboptions)	specify how <i>y</i> axis looks
Plot and graph areas	
<code>aspectratio</code> (# [, pos_option])	plot region aspect ratio
<code>graphregion</code> (suboptions)	customize attributes of graph region
<code>plotregion</code> (suboptions)	customize attributes of plot region
<code>scheme</code> (schemename)	customize the graphics scheme
<code>xsize</code> (#)	change width of graph
<code>ysize</code> (#)	change height of graph
Titles, legend, and notes	
<code>legend</code> ([contents] [location])	standard legend, contents, and location
<code>note</code> (tinfo)	note about graph
<code>title</code> (tinfo)	overall title
<code>subtitle</code> (tinfo)	subtitle of the graph

5 What is new or different with `ginteff`?

In this section, I first compare and contrast `ginteff` with existing community-contributed commands and then with Stata's official commands.

5.1 Comparing `ginteff` with community-contributed commands

There are two community-contributed commands for calculating interaction effects, that is, `inteff` and `inteff3`, for two- and three-way interactions, respectively (Cornelißen and Sonderhof 2009; Norton, Wang, and Ai 2004). These commands have significantly increased our understanding of interaction models and helped analysts correctly compute interaction effects. Yet they accommodate only logit and probit models, and if the interaction involves factor variables, they must be dummy variables. In fact, `inteff3` can handle solely binary variables and cannot be used with continuous variables. `inteff` can compute the effect of a continuous-by-continuous interaction via the cross-partial derivative, but alternative modeling options are excluded (for example, the first difference). Another restriction is that numeric variables cannot have higher-order terms (for example, squared terms).

`ginteff` has none of these limitations. Specifically, `ginteff` extends the computation of interaction effects to cases in which factor variables have more than two categories. It also allows researchers to choose whether to compute the effect of the interacted variables via the partial derivative or the first difference. Additionally, `ginteff` accepts weights (thus accommodating survey analyses), and the interacted variables can have higher-order terms. Last, `ginteff` can be used after a wide range of estimation models because it accommodates most models after which `margins` can be used. This includes regression models for panel data (for example, `xtreg`, `xtlogit`); models for censored or truncated data (for example, `truncreg`, `tobit`); and multiequation models (for example, `biprobit`, `sureg`). It can also be used after models with a polychotomous dependent variable, and the outcomes may be either ordered or not (for example, ordered logit, multinomial logit). In sum, `ginteff` is more general than its predecessors.

`ginteff` also features a simplified syntax and a larger set of options. For instance, users need not provide the names of the interacted variables in a particular sequence.⁵ Besides offering increased flexibility, many elective options have the same syntax as the homonym options of `margins` (for example, `at()`, `level()`, `noweights`, `post`, `predict()`). Because Stata users are already familiar with these options, this should facilitate the adoption of the new command.

This said, `ginteff` comes with one notable restriction. It can estimate interaction effects between continuous variables via the cross-partial derivative only for logistic, logit, and probit models. Even when considering the restricted set of models, computing three-way interaction effects for continuous variables is a novel feature that `inteff3` does not possess. This limitation, though, does not concern interactions involving factor variables or analyses in which the effect is computed via the first-difference approach.

5. `inteff` syntax requires users to list continuous variables first and factor variables second.

For clarity, there are no model restrictions when the interaction is between 1) one factor and one continuous variable, 2) two factor variables, 3) two factor variables and one continuous variable, 4) three factor variables, and, if one is using the first-difference approach, 5) two continuous variables, 6) one factor variable and two continuous variables, and 7) three continuous variables.

Technicalities aside, the substantive difference between `ginteff` and its predecessors is that the former computes the *average* interaction effect, whereas the latter compute *individual* effects for each case in the data.⁶ The decision to compute the average or case-level effects boils down to one's view about how to best report nonconstant effects. This goes beyond interaction effects because it concerns marginal effects in nonlinear models more generally. In nonlinear models, even the *unconditional* effect of a variable differs from one observation to the next because the magnitude of the effect is conditional on the values of other covariates, which are case specific. Thus, there is no one single effect.

Each approach comes with its own advantages and disadvantages. The case-level approach entails computing the interaction effect separately for each individual observation. Because it reveals the heterogeneity of individual effects, this approach can prevent gross generalizations. The alternative approach is to aggregate the individual effects and report the average. The advantage is that we can make inferences about a variable's effect in the population or specific subgroups. This approach is particularly useful when individual cases are anonymous and do not carry special meaning. For example, when examining the effect of an initiative to increase voter turnout, it is the average response that is of immediate interest. In fact, a report that focuses on individuals' idiosyncratic responses may be of little practical relevance to policymakers.

It is beyond the scope of this article to compare the two approaches. In practice, their usefulness depends on the research question and type of data (for example, the voting record of the nine U.S. Supreme Court justices or a large population survey). Importantly, `ginteff` users can still compute the interaction effect for each observation via option `obseff()`. Because it can estimate both the average and the individual-level effects, `ginteff` offers researchers more options to analyze the data.

5.2 Comparing `ginteff` with Stata's official commands

Most `ginteff` results can be reproduced via `margins` because, behind the scenes, it is this Stata command that does the heavy lifting when computing interaction effects. However, there are some notable exceptions. For example, `margins` cannot automatically calculate two- or three-way interaction effects between continuous variables via the partial derivative. Because `dydx()` can take just the first derivative, it cannot be used for second- and third-order cross-partial derivatives. To calculate two- and three-way

6. By taking the mean of `inteff`'s or `inteff3`'s individual interaction effects, one can calculate the average effect. However, one cannot obtain the correct standard error of the mean by averaging the individual standard errors. Thus, from the information provided by `inteff` and `inteff3`, one cannot assess whether the average effect in the population is statistically significant.

interaction effects, the user must write the correct vector notation formula in option `expression()`. Similarly, one cannot directly compute interaction effects via the first-difference approach. For instance, to compute a three-way effect, one needs to first correctly specify eight distinct `at()` options. These represent all combinations between the initial value and the counterfactual increase in the interacted variables. Then, the user must correctly subtract the respective predictive margins with the exact sequence varying based on the order in which the `at()` options were introduced. This second step also requires the use of a different command, `nlcom`. These intermediary steps can be taxing and error prone.⁷

`ginteff` also has extra features that provide additional options and checks. For instance, one cannot simultaneously calculate the average and individual-level interaction effects using `margins`.⁸ This task is straightforward with `ginteff` because option `obseff()` automatically computes the interaction effect for each observation. Furthermore, `ginteff` also checks that the specified variables are actually interacted and there are no missing terms. For example, the analyst is notified if only two terms of a three-way interaction are listed.

Ultimately, having a specialized command to compute interaction effects minimizes mistakes. `margins` is a general command that can compute many quantities of interest, whereas `ginteff` is a one-stop command for most interaction effect calculations. While being versatile is an appealing feature of `margins`, it can also overwhelm users. For instance, there is more than one alternative to compute a given quantity of interest via `margins`. Specifically, some show how to compute interaction effects for factor variables using the `contrast` option (for example, Wiggins [2004]), while others show how to do so using `pwcompare` (for example, Pinzon [2016]). Combining the reference manuals for 1) `margins`, 2) `margins, contrast`, and 3) `margins, pwcompare` amounts to over 80 pages of dense information detailing dozens of options.

6 Interpreting interaction effects

Before showing how `ginteff` can be used to compute interaction effects, I briefly review how to interpret them. Let us assume we have an interaction between two dummy variables, x_1 and x_2 , and the predicted value of y is a function of the interacted variables as well as other covariates,

$$\Pr(y|x_1, x_2, \mathbf{X}) = F(\beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta \mathbf{X})$$

For dummy variables, the marginal effect is the first difference, that is, the change in $\Pr(y)$ in the presence and absence of that variable. In practice, this means calculating $F(\beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta \mathbf{X})$ when the variable of interest is set to 0 and 1 and then subtracting the former from the latter. The interaction effect, in turn, is the discrete

7. For additional details and an actual illustration, see online appendix A.2.1.

8. Even advanced users who are aware of the undocumented `margins, generate()` option cannot compute the individual interaction effects without additional manipulations, because this option cannot be used in conjunction with `margins, contrast` or `margins, pwcompare`.

double difference with respect to both x_1 and x_2 or the difference between two first differences (Norton, Wang, and Ai 2004, 157):

$$\begin{aligned}
\frac{\Delta^2 \Pr(y)}{\Delta x_1 \Delta x_2} &= \frac{\Delta \frac{\Delta \Pr(y)}{\Delta x_1}}{\Delta x_2} \\
&= \frac{\Delta \frac{\Delta F(\beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta \mathbf{X})}{\Delta x_1}}{\Delta x_2} \\
&= \frac{\Delta \{F(\beta_1 \times 1 + \beta_2 x_2 + \beta_{12} \times 1 \times x_2 + \beta \mathbf{X}) - F(\beta_1 \times 0 + \beta_2 x_2 + \beta_{12} \times 0 \times x_2 + \beta \mathbf{X})\}}{\Delta x_2} \\
&= \frac{\Delta \{F(\beta_1 + \beta_2 x_2 + \beta_{12} x_2 + \beta \mathbf{X}) - F(\beta_2 x_2 + \beta \mathbf{X})\}}{\Delta x_2} \\
&= \{F(\beta_1 + \beta_2 \times 1 + \beta_{12} \times 1 + \beta \mathbf{X}) - F(\beta_1 + \beta_2 \times 0 + \beta_{12} \times 0 + \beta \mathbf{X})\} \\
&\quad - \{F(\beta_2 \times 1 + \beta \mathbf{X}) - F(\beta_2 \times 0 + \beta \mathbf{X})\} \\
&= F(\beta_1 + \beta_2 + \beta_{12} + \beta \mathbf{X}) - F(\beta_1 + \beta \mathbf{X}) - F(\beta_2 + \beta \mathbf{X}) + F(\beta \mathbf{X}) \tag{1}
\end{aligned}$$

What does the interaction effect, $\{\Delta^2 \Pr(y)\}/(\Delta x_1 \Delta x_2)$, mean in substantive terms? To answer this question, we need to understand what each of the four elements on the right-hand side of (1) represents. To make the interpretation more concrete, let us assume we want to assess the effect of gender and race on health. Among other determinants, one's health is a function of both gender (0 = male, 1 = female) and race (0 = majority, 1 = minority), and the two condition each other's effect. The dependent variable is a dummy indicator of health (0 = poor and 1 = good). For models with a binary dependent variable, $y \in \{0, 1\}$, where the probability space is bounded between 0 and 1, a popular nonlinear function is the cumulative standard logistic distribution, $\Lambda(\cdot)$, which leads to the logit model. Specifically, this distribution function is $\Lambda(\beta \mathbf{X}) = (e^{\beta \mathbf{X}})/(1 + e^{\beta \mathbf{X}}) = 1/(1 + e^{-\beta \mathbf{X}}) = \Pr(y = 1)$, where $\Pr(y = 1)$ is the probability of the event occurring (Clark and Golder 2023, 297).

For our example, we can write the probability of being in good health as

$$\Pr(h = 1) = \Lambda(\beta_f f + \beta_r r + \beta_{fr} fr + \beta \mathbf{X}) = \frac{1}{1 + e^{-(\beta_f f + \beta_r r + \beta_{fr} fr + \beta \mathbf{X})}}$$

where h is the health indicator, f stands for female, and r stands for race. When we replace y with h , x_1 with f , x_2 with r , and $F(\cdot)$ with $\Lambda(\cdot)$, the four right-hand-side elements of (1) become

$$\frac{\Delta^2 \Lambda(h)}{\Delta f \Delta r} = \Lambda(\beta_f + \beta_r + \beta_{fr} + \beta \mathbf{X}) - \Lambda(\beta_f + \beta \mathbf{X}) - \Lambda(\beta_r + \beta \mathbf{X}) + \Lambda(\beta \mathbf{X})$$

$$= \frac{1}{1 + e^{-(\beta_f + \beta_r + \beta_{fr} + \beta \mathbf{X})}} \quad \{\Pr(h = 1 \mid f = 1, r = 1)\} \quad (2)$$

$$- \frac{1}{1 + e^{-(\beta_f + \beta \mathbf{X})}} \quad \{\Pr(h = 1 \mid f = 1, r = 0)\} \quad (3)$$

$$- \frac{1}{1 + e^{-(\beta_r + \beta \mathbf{X})}} \quad \{\Pr(h = 1 \mid f = 0, r = 1)\} \quad (4)$$

$$+ \frac{1}{1 + e^{-(\beta \mathbf{X})}} \quad \{\Pr(h = 1 \mid f = 0, r = 0)\} \quad (5)$$

To help with the interpretation, I note in curly braces what each element means in terms of predicted probabilities. For example, (2) captures the probability of being in good health ($h = 1$) for a woman ($f = 1$) who is a member of a racial minority ($r = 1$). Similarly, (3) captures the probability of being in good health for a woman who is a member of the majority group ($r = 0$), and so on. For convenience, I will reference the respective probabilities by the corresponding equation number (that is, $\Pr_{\text{Eq. (2)}}$, $\Pr_{\text{Eq. (3)}}$, $\Pr_{\text{Eq. (4)}}$, and $\Pr_{\text{Eq. (5)}}$).

Without altering the result, we can rearrange the four probabilities as the difference between two distinct differences: $(\Pr_{\text{Eq. (2)}} - \Pr_{\text{Eq. (4)}}) - (\Pr_{\text{Eq. (3)}} - \Pr_{\text{Eq. (5)}})$. The first parenthetical statement, $(\Pr_{\text{Eq. (2)}} - \Pr_{\text{Eq. (4)}})$, captures the difference in the probability of being healthy between a minority female and a minority male. Put differently, this is the effect of gender on health for racial minorities. The second parenthetical statement, $(\Pr_{\text{Eq. (3)}} - \Pr_{\text{Eq. (5)}})$, captures the difference in the probability of being healthy between a woman and a man from the majority group. Thus, this is the effect of gender on health for the racial majority. Finally, the difference in the effect of gender *between* the minority and majority groups is the interaction effect. Substantively, it captures the relative strength of the two gender effects. For instance, a positive value on the interaction effect would indicate that the effect of gender on well-being is more pronounced for minorities than for the majority group.⁹

The interpretation of the interaction effect would be similar if gender were interacted with a continuous variable such as age. In this case, the interaction effect would capture the relative strength of the two gender effects attributable to an n -unit increase in age. More specifically, we would compare the effect of gender between the current group of respondents and a counterfactual group where respondents were one year older (assuming the standard one-unit increase in x).

9. To focus on the effect of racial status on health, we could write the discrete double difference as $(\Pr_{\text{Eq. (2)}} - \Pr_{\text{Eq. (3)}}) - (\Pr_{\text{Eq. (4)}} - \Pr_{\text{Eq. (5)}})$. Here we compare the effect of being a member of a racial minority for women (the first parenthetical statement) and the same effect for men (the second parenthetical statement). A positive value on the interaction effect (that is, the difference between the two effects) would indicate that the effect of race on well-being is more pronounced for females than for males. Because the cross or second derivatives are symmetric, $(\{\partial \Pr(y)\} / \{\partial x_1\}) / \{\partial x_2\} = (\{\partial \Pr(y)\} / \{\partial x_2\}) / \{\partial x_1\}$, the interaction effect is the same regardless of whether we group the probabilities to describe the effect of gender or that of race.

Last, we could also have an interaction between two continuous variables, say, age and income (measured in thousands of dollars). In this example, the interaction effect would capture the effect on health of a one-unit increase in income (that is, \$1,000) attributable to a one-unit change in respondents' age. More specifically, we would compare the effect of increasing the respondents' income by \$1,000 with the effect of the same income bump for a counterfactual group where respondents were one year older.¹⁰

7 Computing interaction effects with `ginteff`

I illustrate the capabilities of the `ginteff` command using data from the Second National Health and Nutrition Examination Survey, available from the StataCorp website. The dependent variable (DV) is `health`, which codes the respondents' well-being on a five-point scale (that is, poor, fair, average, good, and excellent). The independent variables are `age`, `female`, and `race`; `age` is numeric, whereas `female` and `race` are factors. `female` is coded 0 for males and 1 for females. `race` is a three-category variable in which 1 = white, 2 = black, and 3 = other. In the upcoming examples, `female` and `race` are the interacted variables, and `age` is used as a control.¹¹ To keep things concise, I frame the discussion around the effect of gender on health. But the estimated effects also apply to race because cross-partial derivatives are symmetric. Last, these exercises are for illustration purposes only. In particular, I do not test a specific hypothesis or include all relevant predictors.

7.1 A binomial logit example

For this exercise, I use a dichotomous indicator of health coded 1 if the respondent's health is above average and 0 otherwise. Specifically, the dummy DV is obtained by collapsing the poor, fair, and average levels into one category and the good and excellent levels into another category.

After getting the data, I first fit an additive logit model with no interactions as a reference point. The coefficient on `female` is negative and statistically significant, which means that, keeping everything else constant, females are less likely to be in good health than men. Because logistic regression is a nonlinear model, the coefficient estimate is not the same as the average marginal effect. Setting `race` and `age` at their observed values, I compute the effect of gender on health using `margins`. The `generate(obs_noint)` option calculates the predicted probability, separately for males and females, for each case in the data (the `_noint` suffix is shorthand for "no interaction"). The results are saved in two variables called `obs_noint1` and `obs_noint2`. The difference between these variables is the individual-level marginal effect of gender. The `level(90)` option simply instructs `margins` to report the 90% confidence interval (CI) instead of the default, the 95% CI.

10. I discuss the interpretation of three-way interaction effects in online appendix A.1.

11. I present and discuss the corresponding three-way interactions in online appendix A.2.

```

. webuse nhanes2f, clear
. keep health race female age
. clonevar health_2l = health
(2 missing values generated)
. recode health_2l (1/3=0) (4/5=1) // two-level health
(10,335 changes made to health_2l)
. label variable health_2l "Dummy Health"
. logit health_2l i.female i.race age, nolog level(90)
Logistic regression
Log likelihood = -6512.4514
Number of obs = 10,335
LR chi2(4) = 1291.33
Prob > chi2 = 0.0000
Pseudo R2 = 0.0902

```

health_2l	Coefficient	Std. err.	z	P> z	[90% conf. interval]	
female Female	-.1419507	.0420207	-3.38	0.001	-.2110687	-.0728328
race						
Black	-.9702392	.0727429	-13.34	0.000	-1.089891	-.8505877
Other	-.4587717	.1525501	-3.01	0.003	-.7096942	-.2078492
age	-.0406942	.0012659	-32.15	0.000	-.0427765	-.0386119
_cons	2.046528	.0688644	29.72	0.000	1.933256	2.1598

```

. margins, dydx(female) generate(obs_noint) level(90)
Average marginal effects
Model VCE: OIM
Expression: Pr(health_2l), predict()
dy/dx wrt: 1.female
Number of obs = 10,335

```

	Delta-method		z	P> z	[90% conf. interval]	
	dy/dx	std. err.				
female Female	-.0312231	.0092384	-3.38	0.001	-.0464188	-.0160273

Note: dy/dx for factor levels is the discrete change from the base level.

I plot the `margins` results in figure 1, where the average marginal effect of gender is indicated by the solid square mark and the horizontal line is the 90% CI. On average, the probability of being healthy is 0.031 $[-0.046, -0.016]$ percentage points lower for females. This effect is statistically significant because the 90% CI does not contain zero. Thus, compared with men, women are significantly less likely to be in good health. To judge whether the average value represents a good measure of central tendency, I also plot the individual marginal effects for all cases in the data. These are indicated by the string of dots below the average effect. Because overlapping values may obscure the distribution of individual effects, the median is superimposed. While the average and median values are very similar, there is still variation among individual responses. This variation is not due to interactions directly modeled (there are none) but rather to the fact that, in nonlinear analyses, even unconditional effects depend on the values of all model covariates.¹²

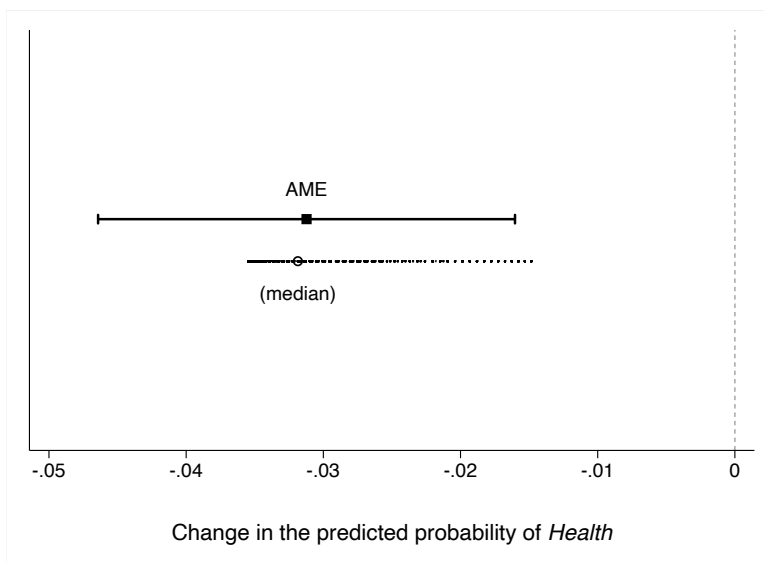


Figure 1. The marginal effect of gender on health (dichotomous indicator)

Next I fit a new model where `female` and `race` are interacted. To compute the average interaction effect, I then issue the `ginteff` command. The `obseff(obs_12w)` option instructs `ginteff` to also compute the interaction effect for each observation (the `_12w` suffix is shorthand for “logit model with a two-way interaction”). These results are saved in two variables, `obs_12w1` and `obs_12w2`, one for each contrast of `race` (that is, change from the base level).

12. The code to produce the figures not generated by the `ginteffplot` command (that is, figure 1, 4a, 5a, and 5b) is provided in the replication files.

```
. logit health_2l i.female##i.race age, nolog level(90)
Logistic regression                               Number of obs = 10,335
                                                    LR chi2(6)      = 1296.53
                                                    Prob > chi2     = 0.0000
                                                    Pseudo R2      = 0.0906
Log likelihood = -6509.849
```

health_2l	Coefficient	Std. err.	z	P> z	[90% conf. interval]	
female						
Female	-.1165629	.0445974	-2.61	0.009	-.189919	-.0432067
race						
Black	-.8073243	.1039329	-7.77	0.000	-.9782786	-.6363699
Other	-.5510827	.213072	-2.59	0.010	-.9015549	-.2006105
female#race						
Female#Black	-.3131646	.1450113	-2.16	0.031	-.5516869	-.0746422
Female#Other	.1922582	.3048587	0.63	0.528	-.3091897	.6937061
age	-.0407121	.0012661	-32.15	0.000	-.0427947	-.0386295
_cons	2.034095	.0692179	29.39	0.000	1.920241	2.147948

```
. ginteff, dydxs(female race) obseff(obs_l2w) level(90)
```

Interaction Effects

```
Statistic      :      Average interaction effect
Standard error :      Delta-method
Δ(i.x1)        :      dy/dx w.r.t. x1; x1 : b0.i(1).female
Δ(i.x2)        :      dy/dx w.r.t. x2; x2 : b1.i(2.3).race
Number of obs  =      10,335
Expression     :      Pr(health_2l), predict()
```

	Statistic	Std. Err.	[90% Conf.	Interval]
Δ(1.x1)#Δ(2.x2)	-.0561421	.02817198	-.10248089	-.00980331
Δ(1.x1)#Δ(3.x2)	.04228932	.06580824	-.06595559	.15053423

Note: dy/dx for factor levels is the discrete change from the base level.

Before moving to interpreting and presenting the results, let us examine the `ginteff` output to understand what information each line conveys. The first line of the output (that is, the left-aligned text at the top) spells out that interaction effects are the substantive quantities of interest computed by the `ginteff` command. The second line clarifies that the main statistic reported in the results table is the average interaction effect. The third line specifies how the standard error of the average interaction effect is calculated (here the delta method).

The next lines introduce the interacted variables, x_1 and x_2 . (For three-way interactions, there will also be an x_3 variable.) For each individual variable, the associated line lists its full name and details how its effect, $\Delta(x_*)$, is computed. Let us consider x_2 , which corresponds to variable `race`. The first thing to note is that the label of the effect, $\Delta(i.x2)$, includes `i.`, conveying that the respective variable is a factor. After the colon, the $\Delta(\cdot)$ notation is explained. In this case, it refers to the partial derivative of $\Pr(y)$ with respect to x_2 (abbreviated to dy/dx); `w.r.t.` is shorthand for “with respect

to.” Because the variable at hand is a factor, the partial derivative is the discrete change from the base level. After the semicolon, the `x2` label is elaborated upon. The additional details comprise the base level (here `b1`, meaning `race = 1`) and the list of factor levels for which the discrete change is to be calculated (enumerated in the `i()` parentheses). In this example, this is all other levels of `race`, `{2, 3}`, which is the default. But one can specify a single level or, for factor variables with many levels, a subset of levels (see example 2 from *The ginteff User’s Manual* for a concrete example). Finally, the full name of the variable is displayed at the end.¹³

The next line of the output notes the number of observations used in the calculation. The last part of the output header spells out the expression of the response for which the effect is calculated. For our example, this is the probability of a positive outcome, $\Pr(\text{health_21} = 1)$, which is the default after logistic regression. But we could have requested the linear prediction instead via the `predict(xb)` option or some other specified prediction. Knowing the type of the response is important because it informs the interpretation of the interaction effect. Because the response is the probability of a positive outcome in our case, the interaction effect can be interpreted as the change in the predicted probability of being in good health attributable to a change in both gender and race. Thus, the interpretation of the interaction effect depends on the outcome metric of the response, which is determined in part by the estimation model. For example, after linear regression, the interaction effect captures the change in the average value of y , but the change in the predicted number of events after a Poisson regression.¹⁴

When the analyst calculates the interaction effect for multiple `at()` scenarios, the legend showing the fixed values of covariates for each scenario will be displayed just above the table of results (see example 2 from *The ginteff User’s Manual*).

The results table shows the value of four estimates: the average interaction effect, its standard error, and the lower and upper limits of the associated CI. These are the table columns. Each row is associated with a distinct interaction effect, and the respective label clarifies the specific scenario. In our example, there are two outputs, one for each contrast of `race`. The first row output, labeled $\Delta(1.x1)\#\Delta(2.x2)$, is the interaction effect associated with a discrete change in x_1 (`female`) from the base level to the factor level 1 and a discrete change in x_2 (`race`) from the base level to the factor level 2. In practical terms, this means moving from males to females and from whites to blacks. The variables represented by x_1 and x_2 , and their respective base levels, are indicated in the output header. The label of the second output, $\Delta(1.x1)\#\Delta(3.x2)$, indicates that the respective interaction effect is attributable to the same discrete change in `female`,

13. The output looks somewhat different for numeric variables, whose effect can be calculated via the partial derivative or the first difference. For the partial derivative approach, dy/dx , the value at which the variable in question is set before taking the derivative (for example, observed values, mean) would be displayed (see example 3 from *The ginteff User’s Manual*). For the first-difference approach, denoted by $(y|x_*+n_*)-(y|x_*)$, both the value of the n -unit increase and the baseline (for example, observed values, mean) would be displayed (see example 4 from *The ginteff User’s Manual*).

14. When `ginteff` computes the interaction effect for multiple outcomes, the response expression output is broken down by individual outcome (see the next section, *An ordered logit example*).

but a change in `race` from the base level to the factor level 3 (racial minorities other than black).

In figure 2a, I graph the estimated average and individual interaction effects separately for both outcomes. Plotting the `ginteff` results is easily achieved using the companion `ginteffplot` command. For instance, figure 2a1 was generated simply by typing

```
. ginteffplot
```

The `ginteffplot` command automatically retrieves the point estimate of the average interaction effect (the solid square mark) and its upper and lower confidence limits (the horizontal capped spike). The significance level of the CI is set by the previous `ginteff` command. The name for the dependent variable used in the x -axis title (that is, `Dummy health`) is the label of `health_21`, which I defined when creating this variable. The graph indicates that, on average, the negative effect of gender is statistically larger for blacks compared with whites (that is, the base category of `race`). Thus, women are generally less likely than men to have above-average health, and furthermore, black women fare worse than white women. The substantive difference is a decrease in the probability of being in good health of 0.056 $[-0.102, -0.010]$ percentage points.

By contrast, women from racial groups other than black seem to fare better than white women. Specifically, the positive estimate in figure 2a2 indicates that the negative effect of gender on health is smaller for minority respondents. However, this effect is not statistically significant. Compared with figure 2a1, this graph has several extra features. First, the x title is more informative because it spells out what the outcome metric is, namely, predicted probability. Second, the plot displays a vertical line at the zero value to more easily judge whether the interaction effect is statistically significant. Third, it reports the full range of individual effects for all cases in the data with the median value superimposed. Because the individual effects are clustered and the median and mean values are very similar, the average represents a good measure of central tendency in this case. The command line to produce the enhanced figure 2a2 is

```
. ginteffplot, obseff(median) output(2)
> xttitle("Change in the predicted probability of {it:Health}", size(4)) zeroline
```

Option `output(2)` indicates that we want to plot the second row estimate from the results table.

By default, the `ginteffplot` command considers only the values for an individual output. This is fine when we estimate a single interaction effect or are mainly interested in assessing whether a given effect is statistically significant. However, if we want to evaluate one interaction effect in relation to another, we need to adjust the graphs so that they use the same scale. More specifically, all plots should have a common x axis. This can be achieved using the option `xcommon()`—where we can specify one other `ginteff` output, `#`; a subset of outputs, `numlist`; or all outputs, `*`—to be put on a common x -axis scale. I illustrate this option in figure 2b, which shows the same interaction effects from figure 2a but with plots that share the same x axis. The two graphs were produced by typing

```
. ginteffplot, obseff(median) xcommon(2)
> xtitle("Change in the predicted probability of {it:Health}", size(4)) zeroline
. ginteffplot, obseff(median) output(2) xcommon(1)
> xtitle("Change in the predicted probability of {it:Health}", size(4)) zeroline
```

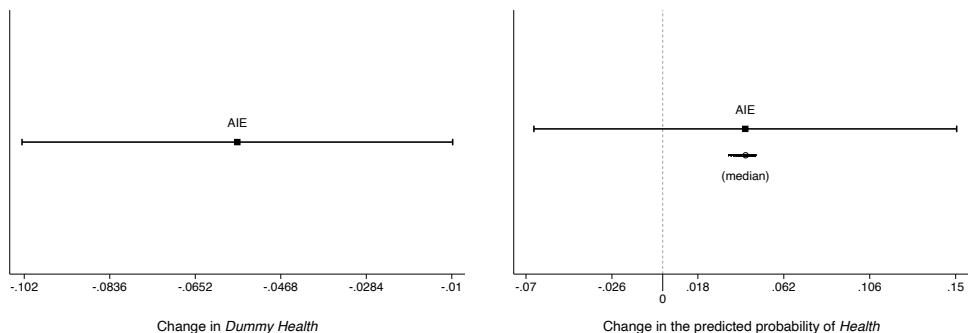
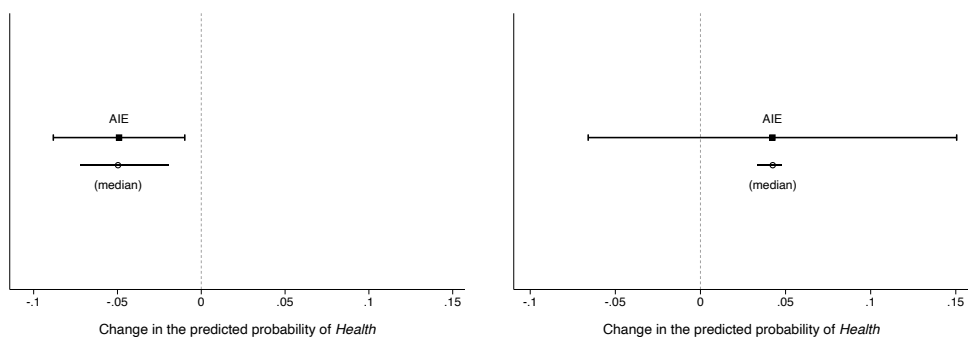
(a) Graphs with individual x axes(b) Graphs with a common x axis

Figure 2. The interaction effect between gender and race on health (dichotomous indicator)

Having the graphs on the same scale facilitates comparisons across outcomes and scenarios, but this should not be taken as a definitive significance of differences test. Specifically, when the CIs of two point estimates overlap, the estimates may or may not be different from one another (Goldstein and Healy 1995; Radean Forthcoming; Schenker and Gentleman 2001). Importantly, this is the case even if the point estimates have different signs and only one of them is statistically significant (Gelman and Stern 2006). This is the situation in our example because the two interaction effects are -0.056 $[-0.102, -0.010]$ and 0.042 $[-0.066, 0.151]$. When the CIs overlap, the solution is to conduct a standard significance of differences test. For that, we need to first save the results, along with the estimated variance–covariance matrix in `e()`, by specifying option `post`. We can then pass these estimates to other commands (for example, `nlcom`) and conduct additional tests.

To illustrate this procedure, I reissue the previous `ginteff` command, using `quietly` to suppress the output. The main difference is that now the results are saved in `e()` because option `post` is specified. We can confirm this by listing the coefficient vector `e(b)`. It is easy to see that these are the `ginteff` estimates, not the coefficients from the logit model. Last, I use `nlcom` to test whether the difference between the two interaction effects is statistically significant. Because the 90% CI of the difference in estimates contains 0, the respective effects are not statistically distinct at the 0.1 significance level. This is despite all conjectural evidence pointing to the contrary (that is, the compared estimates have opposite effect signs and different statistical significance status).

```
. quietly ginteff, dydxs(female race) post
. matrix list e(b)
e(b) [1,2]
      _inteff1  _inteff2
y1  -.0561421  .04228932
. nlcom _b[_inteff2] - _b[_inteff1], level(90)
      _nl_1:  _b[_inteff2] - _b[_inteff1]
```

	Coefficient	Std. err.	z	P> z	[90% conf. interval]	
_nl_1	.0984314	.0701942	1.40	0.161	-.0170277	.2138906

7.1.1 An ordered logit example

One of the advantages of `ginteff` over existing community-contributed commands is that it can accommodate models other than the binomial logit or probit. As an illustration, I next present an application concerning an ordered logit model. The data are the same as before, but now the DV is an ordinal variable, `health_3l`, with three levels, where 1 = poor, 2 = average, and 3 = excellent. Practically, I collapse the poor and fair levels into one category, keep the average category unchanged, and collapse the good and excellent levels into another category.

```
. clonevar health_3l = health
(2 missing values generated)
. recode health_3l (2=1) (3=2) (4/5=3) // three-level health
(9,606 changes made to health_3l)
```

```
. ologit health_3l i.female##i.race age, nolog level(90)
Ordered logistic regression                Number of obs = 10,335
                                           LR chi2(6)    = 1645.32
                                           Prob > chi2   = 0.0000
                                           Pseudo R2    = 0.0760
Log likelihood = -10007.515
```

health_3l	Coefficient	Std. err.	z	P> z	[90% conf. interval]	
female						
Female	-.0848304	.0413361	-2.05	0.040	-.1528223	-.0168386
race						
Black	-.7802284	.0904631	-8.62	0.000	-.929027	-.6314298
Other	-.3585295	.1870633	-1.92	0.055	-.6662212	-.0508378
female#race						
Female#Black	-.2698946	.1224673	-2.20	0.028	-.4713354	-.0684538
Female#Other	.1785221	.2693293	0.66	0.507	-.2644853	.6215294
age	-.0437226	.0011899	-36.75	0.000	-.0456798	-.0417654
/cut1	-3.570566	.0714685			-3.688121	-3.45301
/cut2	-2.137158	.0653756			-2.244691	-2.029625

```
. ginteff, dydxs(female race) obseff(obs_ol2w) level(90)
```

Interaction Effects

```
Statistic      : Average interaction effect
Standard error  : Delta-method
Δ(i.x1)        : dy/dx w.r.t. x1; x1 : b0.i(1).female
Δ(i.x2)        : dy/dx w.r.t. x2; x2 : b1.i(2.3).race
Number of obs  = 10,335
1._pr          : Pr(health_3l==1), predict(pr outcome(1))
2._pr          : Pr(health_3l==2), predict(pr outcome(2))
3._pr          : Pr(health_3l==3), predict(pr outcome(3))
```

	Statistic	Std. Err.	[90% Conf.
1._pr#Δ(1.x1)#Δ(2.x2)	.06109026	.02486272	.02019472
1._pr#Δ(1.x1)#Δ(3.x2)	-.02962508	.04701208	-.10695308
2._pr#Δ(1.x1)#Δ(2.x2)	-.01203469	.00379538	-.01827754
2._pr#Δ(1.x1)#Δ(3.x2)	-.00924215	.01140145	-.02799587
3._pr#Δ(1.x1)#Δ(2.x2)	-.04905557	.02384525	-.08827751
3._pr#Δ(1.x1)#Δ(3.x2)	.03886723	.05830807	-.05704102

	Interval]
1._pr#Δ(1.x1)#Δ(2.x2)	.1019858
1._pr#Δ(1.x1)#Δ(3.x2)	.04770292
2._pr#Δ(1.x1)#Δ(2.x2)	-.00579185
2._pr#Δ(1.x1)#Δ(3.x2)	.00951157
3._pr#Δ(1.x1)#Δ(2.x2)	-.00983362
3._pr#Δ(1.x1)#Δ(3.x2)	.13477547

Note: dy/dx for factor levels is the discrete change from the base level.

Let us consider the same two-way interaction between **gender** and **race**. Upon fitting the new model, I issue the **ginteff** command, which is identical to the one used after the simple logit (save the *stub* name in option **obseff(obs_o12w)**, with the new suffix being shorthand for “ordered logit model with a two-way interaction”). The first thing to note is that there are three times as many output results. Because now we have three distinct health outcomes, the interaction effect is computed separately for each of them. Second, **ginteff** automatically incorporates any auxiliary parameters into the calculations of interaction effects (here **/cut1** and **/cut2**), removing the need for manual manipulation.

Figure 3 graphs the **ginteff** results by outcome and race. For easy comparison, all plots use a common *x*-axis range. Figure 3a1 outlines the difference between the conditional effect of gender for blacks and whites, 0.061 [0.020, 0.102]. The positive estimate means that black women are more likely to be in poor health than white women. When it comes to the average and excellent health outcomes, black women are less likely to be in good health (figures 3a2 and 3a3, respectively). All three interaction effects are statistically significant. The gender effect, however, has the opposite impact when we compare whites with other racial minorities. Specifically, women from minority groups other than black are less likely to be in poor or average health compared with white women and more likely to have excellent health. However, none of these effects are statistically significant (see figure 3b).

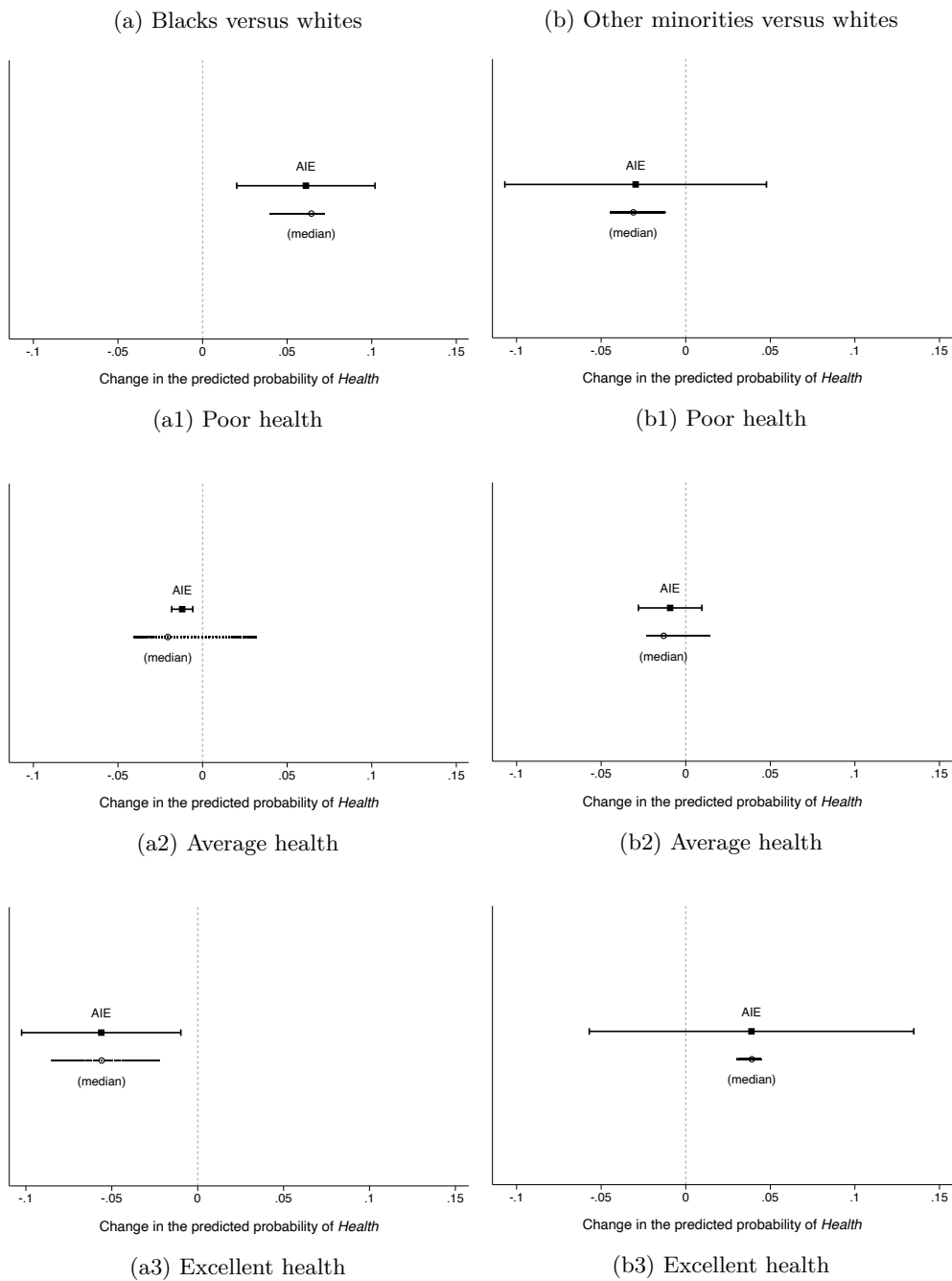


Figure 3. The interaction effect between gender and race on health (three-level indicator)

8 The interaction effect in practical applications

When assessing conditional hypotheses, researchers are typically interested in whether the treatment effect is constant across the levels of the moderator. The quantity of interest in such analyses is the interaction effect because it can provide the answer to the research question for linear and nonlinear models alike. Because of theoretical confusion or a lack of knowhow to compute the interaction effect, many analysts try to get at the interaction without computing the interaction effect. Resorting to heuristics or work-arounds, however, invites mistakes.

A common misconception is that we can draw valid inferences about the significance status of the interaction effect from the statistical significance level of the interaction term coefficient. In nonlinear models, however, a statistically significant interaction term is neither necessary nor sufficient for significant interactive effects. Conflating the interaction effect with the coefficient on the interaction term is another frequent misunderstanding. This is partly because, in a linear regression, the interaction effect is equal to the coefficient on the product term; that is, $\{\partial \Pr(y)\}/(\partial x_1 \partial x_2) = \beta_{x_1 x_2}$. But this is not the case for nonlinear models.

Because of the erroneous association, questions about the coefficient on the interaction term are often misguided questions about the interaction effect. The many Statalist entries on this topic attest to how acute the problem is (for example, Statalist thread [2016, 2017a,b] to reference a few). Specifically, “[p]eople often ask what the ME [marginal effect] of an interaction term is [... even though] there is not one” (Williams 2012, 329) and are willing to go to great lengths to obtain it.¹⁵ For example, to force an estimate for “the average marginal effect of the interaction,” one user proposed forgoing the Stata operator for interactions and manually generating the product between the interacted variables. Thus, instead of running the proper command

```
probit y a##b
```

the user suggested the following work-around:

```
generate c = a*b
probit y a b c
margins, dydx(c)
```

By defrauding `margins` into thinking that `c` is a real variable, one may obtain an estimate for its marginal effect, but, as Clyde Schechter succinctly put it, “[t]hat number is meaningless nonsense” (Statalist thread 2017a). After helpful comments and clarifications from Statalist contributors, it turns out the sought “effect of the interaction term” was actually the interaction effect.

With respect to the practical challenges of computing the interaction effect, some struggle to account for the simultaneous change in the second interacted variable. Let us look at a concrete example. Radean (2019) examines the interaction effect between

15. How frequent is “often”? Sometimes, the frequency is measured in minutes, which led a Statalist contributor to rhetorically ask, “What is the probability that two people would ask a question based on the same misconception within minutes of each other?” (Statalist thread, 2017a).

office benefits and ideological preferences on the probability of party switching in Brazil. The dependent variable is coded 1 if a legislator who is a member of party *A* affiliates midterm with party *B* and 0 otherwise. **President's Coalition** is coded 1 if the legislator's party is a member of the president's legislative coalition and 0 otherwise. This is taken as a proxy for access to state resources. **Ideological Congruence** is the distance between the legislator's ideological position and that of the party.

Because **President's Coalition** is a dummy variable, its effect is calculated as the discrete change from the base level. But is the effect constant across the range of ideological congruence? To answer this question, Radean (2019) reports the effect of **President's Coalition** at minimum and maximum values of congruence with 84% CIs (see figure 1.1a, 147). Unlike with the standard 95% CI, one can visually judge significance of differences at the 0.05 level from the overlap of the customized 84% intervals (Goldstein and Healy 1995; Schenker and Gentleman 2001). The use of this particular interval level, however, is appropriate only if the compared estimates are normally distributed, are independent (that is, not correlated), and have identical standard errors (Radean Forthcoming). These are very restrictive and unrealistic assumptions, and the study does not indicate whether all three conditions are simultaneously met. Using an imprecise interval level may, in turn, lead to incorrect inferences. Ultimately, from the information provided, we cannot determine whether the effect of office benefits changes as ideological congruence also changes.

The solution to all the problems discussed above is to compute the interaction effect—a task made easy by the **ginteff** command. Examining the interaction effect can lend additional support for the research hypothesis, but it may also provide evidence that contradicts it. Either way, the interaction effect should not be ignored.

8.1 Replication of a previous study

In this section, I illustrate one type of substantive findings that could be missed if researchers overlook the interaction effect. To do so, I replicate an analysis from Heller and Mershon (2005) on the effect of the electoral system and party discipline on party switching.¹⁶ The electoral system is made operational in terms of party- versus candidate-centered electoral rules. While the electorates vote for an individual candidate in candidate-centered electoral systems, they cast a party vote in party-centered systems (implicitly voting for all candidates on that party's list). The **Electoral Law** variable is coded 1 if the legislator was elected under candidate-centered electoral rules and 0 if the legislator was elected on a party list.

Furthermore, the degree of party-label clarity (that is, information about the party's policy stance) is used as a proxy for party discipline (that is, the control that party leaders exercise over the rank-and-file members). When party labels are clear, there is less uncertainty about the policy preference of party leadership. When the labels are blurry, legislators may sometimes find themselves at odds with their party's position. Thus,

16. Berry, DeMeritt, and Esarey (2010) are the first to have used this example to illustrate that a statistically significant product term is not a necessary condition for meaningful interactions.

leaders of parties with blurry labels have to enforce party discipline more frequently, which increases a legislator's incentives to switch.

In terms of theoretical expectations, representatives elected under candidate-centered rules should be less likely to switch. Because voters can single them out on the ballot at the next election, they have higher incentives to keep faith with the electorate. By contrast, legislators elected on (closed) party lists are to some extent insulated from voter retribution. The underlying assumption here is that the electorate prefers loyal representatives, who do not jump ship when a better offer comes along. The negative effect of candidate-centered rules on the probability of party switching should be more pronounced in the context of clear party labels because the costs of strict party discipline are less onerous in this context (Heller and Mershon 2005, 538–539).

The output below replicates the logistic regression from Heller and Mershon (see model 3, table 4, 550). With a p -value of 0.274, the coefficient on the interaction term is far from the conventional levels of statistical significance. Based on this information, the authors infer that there is little empirical support for the conditional hypothesis and do not investigate further. Specifically, they do not compute either the marginal effect of the electoral system or the interaction effect.¹⁷

```
. use heller_mershon_jop2005_replication, clear
. logit anycross electlaw##blurrylabel age newcomer groupsize govstatusn
> southisland if elected96, nolog level(90)
```

Logistic regression

Log likelihood = -122.67409	Number of obs = 630
	LR chi2(8) = 83.98
	Prob > chi2 = 0.0000
	Pseudo R2 = 0.2550

anycross	Coefficient	Std. err.	z	P> z	[90% conf. interval]	
1.electlaw	.6580585	1.257445	0.52	0.601	-1.410255	2.726372
1.blurrylabel	3.254135	1.076709	3.02	0.003	1.483105	5.025164
electlaw# blurrylabel 1 1	-1.441158	1.318086	-1.09	0.274	-3.609217	.7269013
age	.0007566	.0182891	0.04	0.967	-.0293264	.0308395
newcomer	.8352592	.3596385	2.32	0.020	.2437065	1.426812
groupsize	-.0209664	.0047239	-4.44	0.000	-.0287366	-.0131962
govstatusno_d	-1.612178	.4880202	-3.30	0.001	-2.414899	-.8094558
southisland	1.191375	.3575821	3.33	0.001	.6032052	1.779546
_cons	-3.852148	1.370978	-2.81	0.005	-6.107207	-1.597089

But, in a logit model, a statistically significant interaction term is not necessary for significant interactive effects (Ai and Norton 2003; Berry, DeMeritt, and Esarey 2010). To elucidate the matter, I first compute the marginal effect of the electoral system for both types of party labels using the `margins` command. These results are plotted in

17. By contrast, for analyses where the coefficient on the product term is statistically significant, the authors discuss the moderating effect of party labels on the effect of the electoral system (for example, see the analysis from model 1, table 3, 548, and the associated discussion on page 547).

figure 4a. The effect is positive in the clear label scenario, suggesting that legislators elected under candidate-centered rules are more likely to switch than their party-list counterparts. As the authors note, this result does not support the theory, but the effect is also not statistically significant. The marginal effect of the electoral system is negative and significant in the blurry label scenario, which is in line with the theoretical expectation. While now we have a better idea of the conditional effect of the electoral system on party switching, we still cannot tell whether the effect is statistically different between the two scenarios. This is because the estimates' CIs overlap. For a definitive answer, I compute the interaction effect using the `ginteff` command.

```
. margins r.electlaw@blurrylabel, contrast(nowald) level(90)
Contrasts of predictive margins                                Number of obs = 630
Model VCE: OIM
Expression: Pr(anycross), predict()
```

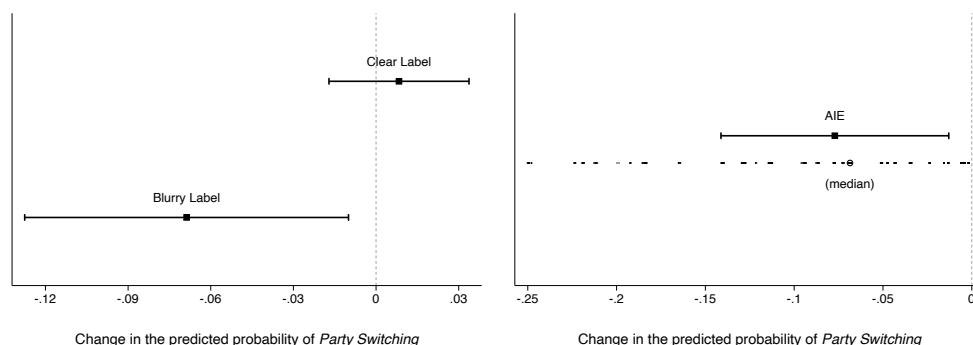
	Delta-method			
	Contrast	std. err.	[90% conf. interval]	
electlaw@blurrylabel				
(1 vs 0) 0	.0083313	.0154511	-.0170835	.033746
(1 vs 0) 1	-.0687639	.0357297	-.127534	-.0099938

```
. ginteff, dydxs(electlaw blurrylabel) obseff(obs_p) level(90)
Interaction Effects
    Statistic      :          Average interaction effect
    Standard error  :          Delta-method
    Δ(i.x1)         :          dy/dx w.r.t. x1; x1 : b0.i(1).blurrylabel
    Δ(i.x2)         :          dy/dx w.r.t. x2; x2 : b0.i(1).electlaw
    Number of obs   =          630
    Expression      :          Pr(anycross), predict()
```

	Statistic	Std. Err.	[90% Conf.	Interval]
Δ(1.x1)#Δ(1.x2)	-.07709514	.03898491	-.14121962	-.01297066

Note: dy/dx for factor levels is the discrete change from the base level.

The statistically significant interaction effect (graphed in figure 4b) indicates that the effect of the electoral system on party switching is in fact distinct between clear and blurry party-label scenarios. Thus, the authors were too quick to dismiss the idea that party-label clarity conditions the effect of electoral rules. This is a piece of information that cannot be gleaned from either the logit coefficients or the conditional effects. More generally, there are cases where only by computing the interaction effect can we ascertain whether the treatment effect varies significantly with the levels of the moderator. In sum, by overlooking the interaction effect, we ignore crucial evidence for testing conditional hypotheses, which can lead to us either understating or, more problematically, overstating the extent of the empirical support for our theories.



(a) The marginal effect of the electoral system across types of party labels

(b) The interaction effect between the electoral system and party labels

Figure 4. The effect of the electoral system on party switching

8.2 Alternative approaches for continuous-by-continuous interactions

Besides computing the interaction effect, there are other options to explore how two continuous variables interact. One alternative is to plot the marginal effect of x_1 across the range of the moderating variable x_2 , $[\Pr(y|x_1 + n; x_2 = \min) - \Pr(y|x_1; x_2 = \min)], \dots, [\Pr(y|x_1 + n; x_2 = \max) - \Pr(y|x_1; x_2 = \max)]$ (see Brambor, Clark, and Golder [2006]). Marginal-effect graphs are useful for outlining the trajectory of the conditional effect but are not designed to compare the effect of x_1 at alternative values of x_2 . Specifically, if there is overlap between the CIs of the effect of x_1 at the minimum and maximum values of x_2 , we cannot tell whether the effect changes significantly with x_2 .

To illustrate the problem, let us consider a logistic regression with a dummy health indicator and a continuous-by-continuous interaction between `height` and `weight`.¹⁸ Figure 5a graphs the effect on health of increasing `height` by one standard deviation from its observed values across the range of `weight`. As expected, an increase in height always has a positive effect on health but more so for heavier people. This is because it decreases the body mass index by a greater amount. However, we cannot tell whether the effect of the counterfactual increase in `height` is statistically different between the minimum and maximum values of `weight`, because the associated 95% CIs overlap.

18. The fitted model is `logit health_21 c.weight##c.height age i.female i.race`, and the full regression output is shown in online appendix B.

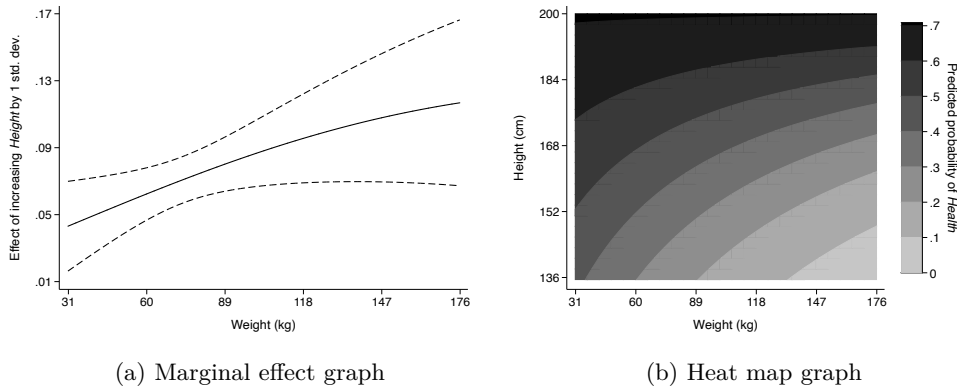


Figure 5. Examining continuous-by-continuous interactions

Heat maps are another popular approach to examine continuous-by-continuous interactions (see Huber [2017]). This type of plot graphs the predicted probability of y across the range of both x_1 and x_2 . The benefit of such graphs is that they cover a wide range of feasible values. Because no individual probabilities are identified, though, heat maps do not typically reveal the estimated uncertainty. Thus, we cannot judge whether a given change in predicted probability is statistically significant. Illustrating this problem, figure 5b graphs the predicted probability of being healthy across the range of both `height` and `weight`. It is easy to note that one has the lowest probability of being in good health when one scores low on height and high on weight (a scenario associated with a high body mass index). Conversely, higher probabilities are associated with low weight and high height scores. That said, from this graph we cannot tell whether $\Pr(\text{health} = 1 \mid \text{height} = \text{max}; \text{weight} = \text{min})$ is statistically different from $\Pr(\text{health} = 1 \mid \text{height} = \text{min}; \text{weight} = \text{max})$.

Unlike other empirical approaches, the interaction effect allows us to directly assess whether the change in the effect of x_1 due to x_2 also changing is statistically significant. Thus, it can be used to assess interactive theories. In fact, establishing whether the treatment effect is distinct at different values of the moderator is the crux of conditional hypothesis testing.

9 Conclusion

Interaction analyses are useful tools to examine complex socioeconomic outcomes where the effect of one variable depends on the presence or levels of another variable. Interaction effects capture the simultaneous change in two (or more) covariates, and their computation is challenging for models with a nonlinear link function (for example, binomial logit or probit) or models involving auxiliary parameters (for example, the correlation parameter in bivariate probit, the cutpoints in ordered logit, etc.). To complicate matters, in nonlinear analyses, the coefficient on the interaction term does not tell us the direction, magnitude, or significance of the interaction effect. For analyses where

the interaction effect cannot be inferred from the model estimates, I introduce a new command that automatically computes two- and three-way interaction effects.

`ginteff` accommodates a large suite of estimation models and allows researchers to use either the partial derivative or the first difference to model the effect of the interacted variables. While `ginteff` is a specialized command, its many options make it fairly flexible. For example, a general concern with nonlinear models is that there is no one single effect, be it conditional or unconditional. In such instances, one can use the `at()` option to specify any number of relevant scenarios or compute the interaction effect for each observation via `obseff()`.

Last, it is important to acknowledge that `ginteff` is a postestimation command in that it retrieves information from a regression model for further analysis but does not produce new estimates as such. This means that the `ginteff` results are only as good as the data fed into the command. In particular, `ginteff` assumes that the multiplicative model is correctly specified. For example, all constitutive terms are included (see Brambor, Clark, and Golder [2006]), there are no unmodeled interaction or quadratic terms, the underlying model assumptions hold, etc.

Concerning unmodeled terms, Beiser-McGrath and Beiser-McGrath (2020) show that omitted product terms can bias the included terms.¹⁹ As a possible solution, the study considers a suite of parametric and nonparametric estimators (that is, the adaptive lasso, kernel regularized least squares, and Bayesian additive regression trees). The advantage of these estimators is that they can select the covariates that belong in the model from a very large set of potential controls without leading to overfitting. One drawback is that they are more conservative; that is, the CI of relevant terms more frequently includes zero (729). Based on Monte Carlo simulations, the authors conclude that, on average, the adaptive lasso is the best approach. If using an alternative estimator, the analyst has to compute the interaction effect by hand while accounting for any constraints associated with that estimator.²⁰ If the results from the alternative estimator and those from the standard model are substantively similar (that is, there are no omitted relevant terms), researchers may use `ginteff` to compute the interaction effect using the estimates from the standard model.

Many researchers take for granted that model assumptions hold in their particular application without assessing the validity of these assumptions. But, when this is not the case, the estimates may be fragile and model dependent. This in turn can lead to incorrect inferences. Hainmueller, Mummolo, and Xu (2019) show that this is so even for the more innocuous case of linear regression. Specifically, the authors consider

19. The omitted product term bias can occur when the analyst includes a nonrelevant interaction term, say, between x and z_2 , but omits a relevant interaction, for example, between x and z_1 , which shares a constitutive term with the included interaction term. Specifically, instead of estimating $y = \beta_x x + \beta_{z_1} z_1 + \beta_{x z_1} x z_1 + \beta_{z_2} z_2$, the analyst fits the incorrect model $y = \beta_x x + \beta_{z_1} z_1 + \beta_{z_2} z_2 + \beta_{x z_2} x z_2$ (Beiser-McGrath and Beiser-McGrath 2020, 710–711).

20. For instance, `margins` cannot be used to compute marginal effects after lasso models. This is because “`margins` requires a full coefficient vector and variance matrix for those coefficients. The lasso inference commands can only tell us about a subset of that coefficient vector and associated variance matrix” (StataCorp 2021a, 110). Because `ginteff` relies on `margins` for the computation of the interaction effect, it also cannot be used after lasso models.

two common assumptions: 1) the linear interaction effect changes at a constant rate with the moderator, and 2) there is sufficient common support in the data to compute valid conditional effects. Based on a literature survey, they find that these assumptions often fail in practice, so they propose some diagnostics. A binning estimator (where a continuous moderator is broken into several bins) can provide a sense of the effect heterogeneity. It may also alert the analyst if the data are sparse. Another diagnostic tool is the kernel smoothing estimator. This estimation strategy relaxes the linearity assumption and estimates a flexible functional form of the treatment effect across the moderator's range. If the diagnostic tests reveal that model assumptions hold, the research can compute the interaction effect using `ginteff`. If not, the analyst needs to first address the underlying problem.

Even if we have the appropriate research design and our model is correctly specified, we still need to exercise caution when computing substantive quantities of interest. We often make assumptions not only at the estimation stage but also in the postestimation phase. The latter type of assumption may be underappreciated but is equally important to obtain practically meaningful estimates. As an example, consider the oft-used fixed-effects logit model, which can be easily fit in Stata via the `xtlogit` command with the `fe` option. The standard panel-data model is

$$\Pr(y_{it} = 1|x_{it}, \alpha_i) = \frac{e^{\beta x_{it} + \alpha_i}}{1 + e^{\beta x_{it} + \alpha_i}}, \quad i = 1, \dots, N, \quad t = 1, \dots, T_i$$

where $(e^\beta)/(1 + e^\beta)$ is the standard logistic cumulative distribution function, α_i is the individual effect, i indexes individuals, and T_i is the number of observations on each individual. Fixed-effects models are attractive because they can account for time-invariant, unobserved individual characteristics. This in turn minimizes the risk that the coefficients on the observed predictors (the β s) are affected by the omitted variable bias. The downside is that we cannot make valid inferences about quantities of interest that require estimates of the fixed effects. The problem is that it is not possible to estimate α_i consistently when T_i is fixed (for a formal discussion, see Greene [2004, 106]). Providing the intuition for why adding more data cannot solve the problem, Wooldridge (2020, 467) explicates that “as we add each additional cross-sectional observation, we add a new α_i . No information accumulates on each α_i when T is fixed.”

This means we cannot compute predicted probabilities or partial effects unless we choose an arbitrary value for α . There is no optimal a priori value “[b]ecause the distribution of α_i is unrestricted—in particular, $E(\alpha_i)$ is not necessarily zero” (Wooldridge 2010, 558).²¹ But this is exactly what is typically assumed. Case in point, the `margins` command assumes $\alpha_i = 0$ for every i when used after `xtlogit` with the `fe` option.²² Some analysts forcefully advise against calculating partial effects at $\alpha_i = 0$ because estimates computed using an arbitrary value are not particularly meaningful (Baetschmann et al. 2020; Santos Silva and Kemp 2016). Because `ginteff` draws on calculations from `margins`, all assumptions are carried over. While evaluations of how problematic a par-

21. To keep with the previous notation, I replace c_i with α_i in the quote from Wooldridge (2010).

22. This assumption is documented in the `xtlogit postestimation` manual entry (StataCorp 2021b, 315).

ticular assumption is are bound to be subjective, at the very least researchers should be aware of the assumptions made in the estimation and postestimation phases.

10 Acknowledgments

The author thanks Daina Chiba, the editor of the *Stata Journal*, and an anonymous reviewer for helpful comments and suggestions.

11 Programs and supplemental materials

To install a snapshot of the corresponding software files as they existed at the time of publication of this article, type

```
. net sj 23-2
. net install st0711      (to install program files, if available)
. net get st0711         (to install ancillary files, if available)
```

The User's manuals and online appendix mentioned in the text can be found by installing these files.

For additional information as well as any future software updates, visit <https://mariusradean.org>.

12 References

- Ai, C., and E. C. Norton. 2003. Interaction terms in logit and probit models. *Economics Letters* 80: 123–129. [https://doi.org/10.1016/S0165-1765\(03\)00032-6](https://doi.org/10.1016/S0165-1765(03)00032-6).
- Baetschmann, G., A. Ballantyne, K. E. Staub, and R. Winkelmann. 2020. feologit: A new command for fitting fixed-effects ordered logit models. *Stata Journal* 20: 253–275. <https://doi.org/10.1177/1536867X20930984>.
- Beiser-McGrath, J., and L. F. Beiser-McGrath. 2020. Problems with products? Control strategies for models with interaction and quadratic effects. *Political Science Research and Methods* 8: 707–730. <https://doi.org/10.1017/psrm.2020.17>.
- Berry, W. D., J. H. R. DeMeritt, and J. Esarey. 2010. Testing for interaction in binary logit and probit models: Is a product term essential? *American Journal of Political Science* 54: 248–266. <https://doi.org/10.1111/j.1540-5907.2009.00429.x>.
- Brambor, T., W. Clark, and M. Golder. 2006. Understanding interaction models: Improving empirical analyses. *Political Analysis* 14: 63–82. <https://doi.org/10.1093/pan/mpi014>.
- Cameron, A. C., and P. K. Trivedi. 2022. *Microeconometrics Using Stata*. 2nd ed. College Station, TX: Stata Press.

- Clark, W. R., and M. Golder. 2023. An Introduction to Interaction Models. Unpublished manuscript, Pennsylvania State University.
- Cornelißen, T., and K. Sonderhof. 2009. Partial effects in probit and logit models with a triple dummy-variable interaction term. *Stata Journal* 9: 571–583. <https://doi.org/10.1177/1536867X0900900404>.
- Gelman, A., and H. Stern. 2006. The difference between “significant” and “not significant” is not itself statistically significant. *American Statistician* 60: 328–331. <https://doi.org/10.1198/000313006X152649>.
- Goldstein, H., and M. J. R. Healy. 1995. The graphical presentation of a collection of means. *Journal of the Royal Statistical Society, Series A* 158: 175–177. <https://doi.org/10.2307/2983411>.
- Greene, W. 2004. The behaviour of the maximum likelihood estimator of limited dependent variable models in the presence of fixed effects. *Econometrics Journal* 7: 98–119. <https://doi.org/10.1111/j.1368-423X.2004.00123.x>.
- Hainmueller, J., J. Mummolo, and Y. Xu. 2019. How much should we trust estimates from multiplicative interaction models? Simple tools to improve empirical practice. *Political Analysis* 27: 163–192. <https://doi.org/10.1017/pan.2018.46>.
- Heller, W. B., and C. Mershon. 2005. Party switching in the Italian Chamber of Deputies, 1996–2001. *Journal of Politics* 67: 536–559. <https://doi.org/10.1111/j.1468-2508.2005.00328.x>.
- Huber, C. 2017. In the spotlight: Visualizing continuous-by-continuous interactions with margins and twoway contour. *Stata News*, vol. 32, no. 1. <https://www.stata.com/stata-news/news32-1/spotlight/>.
- Kam, C. D., and R. Franzese, Jr. 2007. *Modeling and Interpreting Interactive Hypotheses in Regression Analysis*. Ann Arbor, MI: University of Michigan Press.
- Norton, E. C., H. Wang, and C. Ai. 2004. Computing interaction effects and standard errors in logit and probit models. *Stata Journal* 4: 154–167. <https://doi.org/10.1177/1536867X0400400206>.
- Pinzon, E. 2016. Effects of nonlinear models with interactions of discrete and continuous variables: Estimating, graphing, and interpreting. The Stata Blog: Not Elsewhere Classified. <https://blog.stata.com/2016/07/12/effects-for-nonlinear-models-with-interactions-of-discrete-and-continuous-variables-estimating-graphing-and-interpreting/>.
- Radean, M. 2019. Sometimes you cannot have it all: Party switching and affiliation motivations as substitutes. *Party Politics* 25: 140–152. <https://doi.org/10.1177/1354068816688363>.

- . Forthcoming. The significance of differences interval: Assessing the statistical and substantive difference between two quantities of interest. *Journal of Politics*. <https://doi.org/10.1086/723999>.
- Santos Silva, J. M. C., and G. C. R. Kemp. 2016. Partial effects in fixed effects models. U.K. Stata Users Group meeting proceedings. https://www.stata.com/meeting/uk16/slides/santos_uk16.pdf.
- Schenker, N., and J. F. Gentleman. 2001. On judging the significance of differences by examining the overlap between confidence intervals. *American Statistician* 55: 182–186. <https://doi.org/10.1198/000313001317097960>.
- StataCorp. 2021a. *Stata 17 Lasso Reference Manual*. College Station, TX.
- . 2021b. *Stata 17 Longitudinal-Data/Panel-Data Reference Manual*. College Station, TX.
- Statalist thread (The Stata Forum). 2016. Statalist forums: Continuous interaction variables and margins. <https://www.statalist.org/forums/forum/general-stata-discussion/general/1324886-continuous-interaction-variables-and-margins>.
- . 2017a. Statalist forums: Marginal effect of interaction variable in probit regression using Stata. <https://www.statalist.org/forums/forum/general-stata-discussion/general/1379890-marginal-effect-of-interaction-variable-in-probit-regression-using-stata>.
- . 2017b. Statalist forums: The margin effect of interaction term in nonlinear model (e.g. xtgee, family(binomial) link(probit)). <https://www.statalist.org/forums/forum/general-stata-discussion/general/1379888-the-margin-effect-of-interaction-term-in-nonlinear-model-e-g-xtgee-family-binomial-link-probit>.
- Wiggins, V. 2004. Statalist archive: Re: st: Obtaining marginal effects and their standard errors after estimations with interactions. <https://www.stata.com/statalist/archive/2013-01/msg00293.html>.
- Williams, R. 2012. Using the margins command to estimate and interpret adjusted predictions and marginal effects. *Stata Journal* 12: 308–331. <https://doi.org/10.1177/1536867X1201200209>.
- . 2021. Marginal effects for continuous variables. University of Notre Dame. <https://www3.nd.edu/~rwilliam/stats3/Margins02.pdf>.
- Wooldridge, J. M. 2010. *Econometric Analysis of Cross Section and Panel Data*. 2nd ed. Cambridge, MA: MIT Press.
- . 2020. *Introductory Econometrics: A Modern Approach*. 7th ed. Boston: Cengage Learning.

About the author

Marius Radean is a lecturer in the Department of Government at the University of Essex.