%MODPROBE

This document describes the features of MODPROBE version 2.0. Not all of the features documented here are available in prior versions of MODPROBE.

```
%MODPROBE (DATA = datafile, Y = dep, X = [covarlist] focvar modvar
                             [,MODVAL = {mval}]
                             [,JN = {j}{0**}]
                             [,ALPHA = {a}(0.05**)]
                             [,EST = {z}(0**)]
                             [,SAVE = {s}{0**}]
                             [,NOMOD = {nm}{0**}]
                             [,PTILES = {pt}{0**}]
                             [,ITERATE = {i}{10000**}]
                             [,CONVERGE = {cc}{.0000001**}]
                             [,CENTER = {d}{0**}]
                             [,HC3 = {hc}{0**}]
                             [,DECIMALS = {dec}{10.4}];
```

Subcommands and variables in brackets are optional
** Default if subcommand is omitted

Overview

`MODPROBE` is an aide for estimating and probing two way interactions in OLS and logistic regression. It estimates model coefficients and standard errors in a model including predictor variables `focvar`, `modvar`, various products involving `modvar` and `focvar`, and any additional variables in `covarlist` to estimate dependent variable `dep`. In addition to estimating the coefficients of the model, it produces tests of the conditional effect of the `focpred` on `dep` at values of `modvar`, also called *conditional effects* or *simple slopes*. For continuous `modvar`, conditional effects of `focvar` are estimated at the sample mean of `modvar` as well as one standard deviation above and below the sample. When `modvar` is dichotomous, it produces conditional effects of `focvar` at the groups defined by codes in `modvar`. The user can also request specific values of `modvar` at which to estimate the conditional effect of `focvar`. If requested, when `modvar` is dichotomous or continuous, `MODPROBE` also estimates the regions of significance for `focvar` as a function of `modvar` using the Johnson-Neyman technique or the Potthoff modification of the Johnson-Neyman technique. The command can also produce a table of estimated values of `dep` from the model for various combinations of `modvar` and `focvar`, helpful for visualizing the interaction between `focvar` and `modvar`. For a discussion of the methods implemented by MODPROBE for continuous or dichotomous `modvar` and/or `focvar` , see Hayes, A. F., & Matthes, J. (2009). *Computational procedures for probing interactions in OLS and logistic regression: SPSS and SAS implementations*. *Behavior Research Methods, 41,* 924-936.

Examples

```
%MODPROBE (DATA=example, Y=know, X=sex age educ attitude, JN=1);
```

- educ is the focal predictor variable and attitude is the proposed moderator variable

- Estimates the regression coefficients of an OLS regression model predicting know from sex, age, educ, attitude, and educ × attitude as predictors.

- Unless attitude is dichotomous, produces the conditional effect of educ when attitude is at the sample mean as well as one standard deviation above and below the sample mean

- Derives the regions of significance for the conditional effect of educ—the values for attitude, if they exist, at which the effect of educ transitions between significant and nonsignificant at the $\alpha = .05$ level of significance.

- Prints the change in R^2 due to the interaction, as well as the *F*-ratio and *p*-value for the change.

```
%MODPROBE (DATA=example, Y=vote, X=sex age tho_comp, est=1);
```

- age is the focal predictor variable and tho_comp is the proposed moderator variable.

- With vote as a dichotomous outcome (1 = yes, 0 = no), estimates the logistic regression coefficients in a model estimating the log odds of vote = 1 (versus vote = 0) from sex, age, tho_comp, and age × tho_comp.

- Generates the estimated effect of age on the log odds of vote = 1 (versus 0) in the groups defined by tho_comp.

- Produces a table of estimated log odds of vote = 1 (versus 0) as well as probabilities for various values of age in each of the groups defined by tho_comp. A SAS data file is also produced in the SAS session containing the data in this table.

Specification of Focal Predictor and Moderator

The predictor variables in the model are specified with a list of variable names following x= in the order covlist focvar and mdvar (covlist is optional). Thus, the moderator variable should be listed last in the list, and the focal predictor specified second to last. All other variables are treated as additional predictors or covariates in the model but do not play roles of focal predictor or moderator.

Multicategorical Covariates

Multicategorical covariates with *k* categories should be presented with a set of *k* − 1 codes constructed manually outside of PROCESS. These codes should be listed as covariates in the MODPROBE command.

Probing the Interaction using the "Pick-a-Point" Approach

By default, MODPROBE will generate the conditional effects or "simple slopes" for `focvar` at values of `modvar` equal to the sample mean as well as a standard deviation above and below the sample mean. If `modvar` is dichotomous, the output will include conditional effects for `focvar` at the two values of `modvar`. Alternatively, the conditional effect can be estimated at the 25%, 50th, and 75th percentiles of the distribution of `modvar`.

The user can request instead the conditional effect of `focvar` at a specific value of `modvar` using the `MODVAL` command, specifying the value of the moderator at which to condition the estimate of the effect of `focvar`.

For example, in addition to the OLS regression model, this command generates the conditional effect of `educ` when `attitude` = 2.

```
%MODPROBE (DATA=example, Y=know, X=sex age educ attitude, MODVAL=2);
```

If the modval option is used in conjunction with the center option, the value specified as the argument in modval should be in the centered metric. Only a single value can be specified as the argument in `MODVAL`. The `MODVAL` command is ignored when used in conjunction with `MCMOD` option.

Replacing the `MODVAL` command with `PTILES = 1`, as below, produces the OLS regression model as well as the conditional effect of `educ` at values of `attitude` that define the 25th, 50th, and 75th percentiles of the distribution in the sample.

```
%MODPROBE (DATA=example, Y=know, X=sex age educ attitude, PTILES=1);
```

Probing the Interaction using the Johnson-Neyman Technique

The Johnson-Neyman technique can be implemented to derive the region(s) of significance for the conditional effect of `focvar`. This method derives the value(s) of `modvar` at which the conditional effect of focvar transitions between statistically significant and nonsignificant. The Johnson-Neyman method, requested with subcommand `JN` and setting `j` to 1, is enabled only if `modvar` is a quantitative variable The macro assumes that if `modvar` has more than 2 discrete values, it is a quantitative variable. The more conservative Potthoff modification to the Johnson-Neyman technique is implemented by setting `j` to 2 in the `JN` subcommand. The Potthoff modification is not available for logistic regression models.

Covariates

The analyst will be using `MODPROBE` to better understand the nature of the interaction between `focval` and `modvar` in the regression. All variables in `covarlist`, if any, are included as additional predictors in the regression model and thus can be interpreted as statistical controls when estimating the interaction and the conditional effects of `focvar` at values of `modvar`.

Mean Centering

Mean centering is sometimes advocated when estimating models with interactions to ease interpretation of the coefficients for variables that define the interaction and to minimize the likelihood of rounding

error creeping into computations. Use of the CENTER subcommand, setting d to 1, will automatically mean center focvar, modvar, or both prior to computation of the product and estimation of the model coefficients. Mean centering has no effect on the interaction coefficient or its test of significance.

Heteroscedasticity-Robust Inference

Use of the HC3 subcommand, setting hc to 1 (i.e., hc3=1), will generate heteroscedasticity-consistent standard errors for all regression coefficients using the HC3 standard error estimator. In addition, inferential tests for the incremental R^2 due to product terms will be based on a modified *F*-test using the covariance matrix of parameter estimates based on the HC3 estimator of covariation. For a discussion, see Hayes and Cai (2007, *Behavior Research Methods, 39,* 709-722).

Binary Outcomes

The MODPROBE command automatically detects whether or not dep is binary and estimates an OLS or logistic model accordingly. If binary, the macro estimates the coefficients in a linear model of the log odds of the event coded with dep. Prior to estimation, dep is recoded such that the higher code in dep is recoded 1 and is treated as the event being modeled. The lower code in dep is recoded to 0.

The coefficients of the model are estimated iteratively using the Newton-Raphson method. In the event of nonconvergence during iteration toward the maximum likelihood solution for model coefficients, the macro allows the user to change the default number of iterations (i) and convergence criteria (cc) using the ITERATE and CONVERGE subcommands, which default to 10000 and .0000001, respectively.

The Potthoff modification to the Johnson-Neyman technique for probing the interaction is not available for binary outcomes.

ALPHA subcommand

The ALPHA subcommand is used to specify the level of significance when estimating regions of significance using the Johnson-Neyman technique. The a argument defaults to 0.05, for regions of significance at the $\alpha = .05$ level. The analyst can set a to .01 or .10 for the $\alpha = .01$ and $\alpha = .10$ significance levels, respectively. No other values of a are acceptable arguments in the ALPHA subcommand

EST and SAVE options

To help visualize and interpret the nature of the interaction, the EST subcommand generates a table of predicted values of dep from the model, using various values of focvar and modvar. This table is generated by setting z to 1 in the EST subcommand. If there are any variables listed in covarlist, these are set to their sample mean when deriving the predicted values. The EST command is disabled when the user probes the interaction using the Johnson-Neyman technique.

In the output, the estimated value of the outcome is listed as "yhat". For OLS regression, this is simply the estimate of dep from the regression model. For logistic regression, yhat is the estimated log odds of the event coded with dep (with the higher code arbitrarily treated as the the event modeled, as described

below). The /EST command for logistic models will also produce a column labeled "prob", which is the estimated log odds converted to a conditional probability of the event using the standard formula prob = e^{yhat} / (1 + e^{yhat}).

When the EST command is used, a version of the data file shown in the output is saved as temporary SAS data file. The data file will be named "ESTVALS". This file can be used to generate a plot of the interaction. For example, the command

```
%MODPROBE (DATA=example, Y=know, X=sex age educ attitude, EST=1, SAVE=1);
```

will generate a SAS data file called ESTVALS containing a column of values of educ across the observed range in the data, a column of values of attitude equal to the mean and plus or minus one standard deviation from the sample mean, and a column of estimated values of know (named "yhat" in the data file) from the model for various combinations of educ and attitude and setting sex and age to their sample mean (note: in the case of dichotomous covariate coded 0 and 1, the mean is just the proportion of the sample coded 1). The ESTVALS data file can then be used to generate a plot of the interaction between educ (focal predictor) and attitude (moderator). For example, PROC PLOT could be used as such:

```
proc plot data = estvals;
  plot yhat*educ = attitude;
  run;
```

WARNING: The /EST command should not be used to generate data for plotting a three way interaction using the procedure described on page 932 of Hayes and Matthes (2009). The estimated values will be incorrect if what the macro interprets as a covariate involves a product of the focal predictor. Use PROCESS for the estimation and probing of a three-way interaction and the generation of data to facilitate plotting.

No Moderation Model

By default MODPROBE estimates a model that includes moderation of the effect of focvar on depvar by modvar. MODPROBE can also estimate a model that fixes the effect of focvar on depvar to be invariant across values of or groups defined by focvar. To do so, set the nm argument in NOMOD to 1 (i.e., "NOMOD= 1").

Output Resolution

By default, MODPROBE allocates 10 characters to numerical output, with four after the decimal place. This can be changed using the decimals option, set the dec argument to the number of characters, a decimal point, and the number of decimal places desired. For example, specifying decimals = 8.3 allocates eight characters to numbers in the output with three following the decimal.

PROCESS and RLM as a substitutes for MODPROBE

PROCESS and RLM are capable of most things MODPROBE can do, but PROCESS and RLM have some features not available in MODPROBE. For example, PROCESS can estimate, plot, and probe three way interactions, and both MODPROBE and RLM can estimate models with a multicategorical focal predictor or moderator.

PROCESS is introduced in Hayes (2013) *An Introduction to mediation, moderation, and conditional process analysis* (http://www.guilford.com/p/hayes3). RLM is introduced in Darlington and Hayes (2016). *Regression and linear models* (http://www.afhayes.com/regression-and-linear-models.html)

Notes

- A case will be deleted from the analysis if missing on any of the variables in the regression model.

- Conditional effects of `focvar` will be estimated regardless of whether the interaction between `focvar` and `modvar` is statistically significant. Use your own judgment when deciding whether or not to interpret conditional effects in the absence of a statistically significant interaction.

- Do not use STRING formatted variables in any of your models. Doing so will produce errors. All variables should be NUMERIC format.

- MODPROBE does not accept sampling weights

- Do not use a multicategorical focal predictor or moderator with MODPROBE. Use PROCESS or RLM to estimate and probe an interaction when your focal predictor or moderator is multicategorical.