Partial effects in probit and logit models with a triple dummy variable interaction term

Thomas Cornelißen[†] **and Katja Sonderhof**[‡]

Leibniz Universität Hannover, Discussion Paper No. 386

Revised version: June 2009

ISSN: 0949–9962

Abstract

In non-linear regression models, such as probit or logit models, coefficients cannot be interpreted as partial effects. The partial effects are usually non-linear combinations of all regressors and regression coefficients of the model. We derive the partial effects in such models with a triple dummy variable interaction term. The formulae derived here are implemented in the Stata program `inteff3`. The program applies the delta method in order to also compute the standard errors of the partial effects. We illustrate the use of the program by an empirical application, analyzing how the gender gap in labor market participation is affected by the presence of children and a university degree. We find that the presence of children increases the gender gap in labor market participation, but that this increase is smaller for more highly educated individuals.

Keywords: probit model, dummy variables, interaction terms, partial effects, Stata, labor market participation

JEL-Classification: C25, C87, J21

[†]Centre for Research and Analysis of Migration, UCL Department of Economics, t.cornelissen@ucl.ac.uk
[‡]Institute of labor Economics, Leibniz Universität Hannover, sonderhof@aoek.uni-hannover.de

1 Introduction

Regression analysis usually aims at estimating the partial effect of a regressor on the outcome variable, holding effects of other regressors constant. The partial effect of a continuous regressor is given by the partial derivative of the expected value of the outcome variable with respect to that regressor. For discrete regressors the effect is usually computed by the difference in predicted values for a given change in the regressor. In the linear regression model, the partial effect of a regressor is given by the regression coefficient. In non-linear regression models, such as probit and logit models, the partial effects are more complicated: they are usually non-linear combinations of all regressors and regression coefficients of the model.

When an interaction term of two variables is included in the model, the interaction effect of the two variables is given by the cross-partial derivative (or difference, in the case of discrete regressors) of the expectation of the dependent variable with respect to the two interacted variables. In a linear model this is simply the coefficient on the interaction term. In a non-linear model, the cross-derivative or difference is usually a non-linear combination of all regressors and all coefficients of the model. Ai and Norton (2003) and Norton et al. (2004) derive the formulae of interaction effects of two interacted variables in a logit and probit model.

In this paper we look at the case of probit and logit models, in which three dummy variables are included alongside with their pairwise interactions and their triple interaction. This case occurs when the effect of a binary regressor on a binary dependent variable is allowed to vary over combinations of two sub-groups. For example, one may be interested in the way a university degree and the presence of children affect the gender difference in labor market participation. To this effect, one may run a binary choice model of labor market participation including dummies for female, university degree and presence of children, as well as their pairwise and triple interaction terms[1].

We present the partial effects in an analogous way as Ai and Norton (2003) and Norton et al. (2004). The standard errors of the partial effects can be computed using the delta method (see e.g. Davidson/MacKinnon 2004, p.202). We implemented the computation of

[1] A similar application of a probit or logit model with a triple dummy variable interaction term is the difference-in-difference-in-differences estimator with a binary dependent variable (Gruber 1994, Gruber and Poterba 1994). However, Puhani (2008) argues that the treatment effect in non-linear difference-in-differences models is not given by the interaction effect à la Ai and Norton (2003). In fact, computing the interaction effect à la Ai and Norton (2003) would not ensure that the difference-in-differences treatment effect is bounded between 0 and 1.

the partial effects and their standard errors in a companion Stata program `inteff3`. The program is available by typing `net search inteff3` in Stata, and requires at least Stata version 9. It covers partial effects in probit and logit models, but only treats interactions of dummy variables, not of continuous variables.

The paper proceeds as follows. Section 2 derives the partial effects of the three dummy variables and their interactions in probit and logit models. Section 3 describes the Stata ado-file `inteff3` and presents a short empirical application. Section 4 concludes.

2 The partial interaction effects in probit and logit models with a triple dummy variable interaction term

The model with a triple dummy variable interaction term is

$$\begin{aligned} P(y=1|x_1,x_2,x_3,\tilde{\mathbf{x}}) &= F(\beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 \\ &\quad + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 + \beta_{123} x_1 x_2 x_3 + \tilde{\mathbf{x}}\tilde{\beta}) \\ &= F(\mathbf{x}\beta) \end{aligned} \tag{1}$$

where subscripts for observations are dropped for simplicity, y is the binary dependent variable, x_1, x_2 and x_3 are dummy variables to be interacted, β_j are the associated coefficients, and $\tilde{\mathbf{x}}\tilde{\beta}$ denotes the linear combination of all remaining explanatory variables and coefficients. In the case of a probit model, F is the standard normal cumulative density function. In the case of a logit model, it is the cumulative density function of the logistic distribution.

For continuous variables, partial effects are usually computed as the derivative of the dependent variable with respect to the regressor of interest. As the dummies x_1, x_2 and x_3 and their interactions are discrete variables, their partial effects are more appropriately derived by partial differences rather than partial derivatives. The partial effect of the dummy variable x_1 is then the change in the predicted probability of $y = 1$ when x_1 changes from 0 to 1 and all other variables are held constant at specific values:

$$\begin{aligned} \frac{\Delta F(\mathbf{x}\beta)}{\Delta x_1} &= F(\beta_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_2 + \beta_{13} x_3 + \beta_{23} x_2 x_3 + \beta_{123} x_2 x_3 + \tilde{\mathbf{x}}\tilde{\beta}) \\ &\quad - F(\beta_2 x_2 + \beta_3 x_3 + \beta_{23} x_2 x_3 + \tilde{\mathbf{x}}\tilde{\beta}) \end{aligned} \tag{2}$$

The effects of of the dummies x_2 and x_3 can be derived analogously.
The interaction effect of x_1 and x_2 captures how x_2 affects the effect of x_1 on y. This is a second difference, or cross difference, i.e. it is the change of the (first) difference given in (2), for a change of x_2 from 0 to 1:

$$\begin{aligned}\frac{\Delta^2 F(\mathbf{x}\beta)}{\Delta x_1 \Delta x_2} &= F(\beta_1 + \beta_2 + \beta_3 x_3 + \beta_{12} + \beta_{13} x_3 + \beta_{23} x_3 + \beta_{123} x_3 + \tilde{\mathbf{x}}\tilde{\beta}) \\ &\quad -F(\beta_1 + \beta_3 x_3 + \beta_{13} x_3 + \tilde{\mathbf{x}}\tilde{\beta}) - F(\beta_2 + \beta_3 x_3 + \beta_{23} x_3 + \tilde{\mathbf{x}}\tilde{\beta}) \\ &\quad +F(\beta_3 x_3 + \tilde{\mathbf{x}}\tilde{\beta}). \end{aligned} \tag{3}$$

The interaction effects of x_1 and x_3 and of x_2 and x_3 can be derived in the same way.
The triple interaction effect is a third difference. It is the change of the second difference in (3) when x_3 changes from 0 to 1 and all other variables are held constant at specific values:

$$\begin{aligned}\frac{\Delta^3 F(\mathbf{x}\beta)}{\Delta x_1 \Delta x_2 \Delta x_3} &= F(\beta_1 + \beta_2 + \beta_3 + \beta_{12} + \beta_{13} + \beta_{23} + \beta_{123} + \tilde{\mathbf{x}}\tilde{\beta}) \\ &\quad -F(\beta_1 + \beta_2 + \beta_{12} + \tilde{\mathbf{x}}\tilde{\beta}) - F(\beta_1 + \beta_3 + \beta_{13} + \tilde{\mathbf{x}}\tilde{\beta}) \\ &\quad -F(\beta_2 + \beta_3 + \beta_{23} + \tilde{\mathbf{x}}\tilde{\beta}) + F(\beta_3 + \tilde{\mathbf{x}}\tilde{\beta}) \\ &\quad +F(\beta_2 + \tilde{\mathbf{x}}\tilde{\beta}) + F(\beta_1 + \tilde{\mathbf{x}}\tilde{\beta}) - F(\tilde{\mathbf{x}}\tilde{\beta}). \end{aligned} \tag{4}$$

With given estimates of the coefficients of the non-linear model, $\hat{\beta}$, equations like (2)-(4) can be used to derive estimates of the partial effects. As the partial effects are non-linear functions of the underlying parameters estimates $\hat{\beta}$, their standard errors can be computed using the delta method (see e.g. Davidson/MacKinnon 2004, p.202). Let $\mathbf{g}(\hat{\beta})$ be a column vector of k partial effects g_i, $(i = 1, \ldots, k)$. Then, for the given estimated covariance matrix of the regression coefficients, $\hat{\mathbf{V}}(\hat{\beta})$, the covariance matrix of $\mathbf{g}$, can be estimated according to the delta method by

$$\hat{V}(\mathbf{g}) = \hat{\mathbf{G}}\hat{\mathbf{V}}(\hat{\beta})\hat{\mathbf{G}}', \tag{5}$$

where $\hat{\mathbf{G}} \equiv \mathbf{G}(\hat{\beta})$ is the matrix $\partial \mathbf{g}(\beta)/\partial \beta'$. The ith row of $\mathbf{G}(\hat{\beta})$ is the vector of partial derivatives of the ith function with respect to $\hat{\beta}'$ or, in other words, the typical element in row i and column j of $\mathbf{G}(\hat{\beta})$ is $\partial g_i(\beta)/\partial \beta_j$ (Davidson/MacKinnon 2004, p. 208).
Hence, the method requires the derivatives of the partial effects (of the type shown in (2)-(4)) with respect to the underlying regression coefficients β. As an example, the

derivatives of the effect (2) with respect to β_1, β_{12}, β_{123} and a coefficient β_j (part of $x\tilde{\beta}$) are represented in the appendix.

We have implemented the computation of the partial effects and their standard errors in the Stata program `inteff3`. The program computes partial effects at means, at values specified by the user, or the average partial effects, which are computed by averaging over the partial effects for each observation in the sample.

3 The Stata ado-file `inteff3` and an empirical application

We illustrate the use of `inteff3` by means of a probit regression of labor market participation[2]. Ideally we would present an empirical application using data from the German Socio Economic Panel (GSOEP), a representative household panel data set. As the GSOEP data is subject to data protection rules that do not allow users to disseminate the data to third parties, using it would not allow us to submit the data we used to generate the output in this paper. We therefore present an empirical example with simulated data. The simulation, however, is based on the real GSOEP data.

We start by extracting the following data from the GSOEP waves 2000 to 2006: a dummy for labor market participation (`particip`) as the dependent variable; dummies for female gender (`female`), university degree (`uni`) and the presence of children (`child`) as the main explanatory variables. From this we generate the following interaction terms:

```
gen fem_child=female*child
gen fem_uni=female*uni
gen child_uni=child*uni
gen fem_chi_uni=female*child*uni.
```

As control variables we also extract variables for age and its square (`age`, `age_sq`), a dummy for German nationality (`german`), six year dummies (`year*`) and 15 state dummies (`state*`).

We then include all explanatory variables into a probit regression of labor force participation which we run on the GSOEP data covering roughly 87,000 observations. After that, we reduce the sample size to 2000 and replace all explanatory variables by random variables with the same mean as the variables observed in the data. Based on these sim-

[2]The program is available by typing `net search inteff3` in Stata, and requires at least Stata version 9. It covers partial effects in probit and logit models, but only treats interactions of dummy variables, not of continuous variables.

ulated random variables, we predict the linear combination $x'\hat{\beta}$ of the estimated probit model and add an error term e to it, drawn from a normal distribution with mean zero and standard deviation .8. We create a simulated dependent variable for labor market participation that is 1 if $x'\hat{\beta}+e > 0$ and 0 otherwise. All output produced in the following is based on this simulated data, but we will also mention the results obtained with the real data, to show that the simulated data reproduces those results reasonably well.

The probit model followed by `inteff3` gives the following results:

```
. probit particip female child uni fem_child fem_uni child_uni fem_chi_uni age
>  age_sq german year2 year3 year4 year5 year6 year7 state1 state2 state3 state4
>  state6 state7 state8 state9 state10 state11 state12 state13 state14 state15
>  state16

Iteration 0:   log likelihood = -829.64661
Iteration 1:   log likelihood = -591.93241
Iteration 2:   log likelihood = -573.11604
Iteration 3:   log likelihood =  -572.0443
Iteration 4:   log likelihood = -572.03799
Iteration 5:   log likelihood = -572.03799

Probit regression                                 Number of obs   =       2000
                                                  LR chi2(31)     =     515.22
                                                  Prob > chi2     =     0.0000
Log likelihood = -572.03799                       Pseudo R2       =     0.3105

------------------------------------------------------------------------------
    particip |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      female |  -.0596092    .265928    -0.22   0.823    -.5808184       .4616
       child |   .7013859   .2811934     2.49   0.013     .1502569    1.252515
         uni |   1.035991   .2525665     4.10   0.000     .5409695    1.531012
   fem_child |  -1.554207   .3508084    -4.43   0.000    -2.241779   -.8666351
     fem_uni |  -.3498625   .3218242    -1.09   0.277    -.9806264    .2809015
   child_uni |  -.5425721   .3414728    -1.59   0.112    -1.211847    .1267023
 fem_chi_uni |   .5205862   .4196237     1.24   0.215    -.3018611    1.343033
         age |   .4486498   .0347122    12.92   0.000     .3806152    .5166844
      age_sq |  -.0053626   .0004532   -11.83   0.000    -.0062509   -.0044744
      german |   .3952258   .1720125     2.30   0.022     .0580875     .732364
       year2 |   .0879536   .1183411     0.74   0.457    -.1439907    .3198979
       year3 |  -.1167532   .1228444    -0.95   0.342    -.3575238    .1240173
       year4 |   .0254419   .1222224     0.21   0.835    -.2141096    .2649934
       year5 |  -.0571422     .13155    -0.43   0.664    -.3149754     .200691
       year6 |  -.1810155   .1152651    -1.57   0.116    -.4069309    .0448998
       year7 |  -.0859129   .1224446    -0.70   0.483       -.3259    .1540742
      state1 |  -.0718204   .2473074    -0.29   0.772    -.5565339    .4128931
      state2 |   .3606021   .1856515     1.94   0.052    -.0032681    .7244723
      state3 |    .103183   .3880872     0.27   0.790     -.657454      .86382
      state4 |  -.0520698   .1537476    -0.34   0.735    -.3534096    .2492699
      state6 |   .2754981   .0993348     2.77   0.006     .0808054    .4701908
      state7 |   -.263168   .2176327    -1.21   0.227    -.6897202    .1633843
      state8 |   .0741054     .16948     0.44   0.662    -.2580694    .4062802
      state9 |   .0641097   .1297169     0.49   0.621    -.1901307      .31835
```

```
     state10 |   .2997607   .1095528     2.74   0.006     .0850412    .5144802
     state11 |  -.1825238   .1996686    -0.91   0.361    -.5738671    .2088194
     state12 |  -.1826433   .2803153    -0.65   0.515    -.7320513    .3667646
     state13 |   .3395475   .1951523     1.74   0.082    -.0429439    .7220389
     state14 |   .2726648   .2267348     1.20   0.229    -.1717273     .717057
     state15 |  -.0041417    .249022    -0.02   0.987    -.4922159    .4839324
     state16 |  -.0416705   .1900511    -0.22   0.826    -.4141638    .3308228
       _cons |   -8.49694   1.087946    -7.81   0.000    -10.62928   -6.364605
------------------------------------------------------------------------------

. inteff3

Dummies and Interactions: female, child, uni, fem_child, fem_uni, child_uni, fem_chi_uni.
Control variable: age age_sq german year2 year3 year4 year5 year6 year7 state1 state2
state3 state4 state6 state7 state8 state9 state10 state11 state12 state13 state14
state15 state16, constant term.

Marginal effect at means of probit estimation sample:
------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      female |  -.1462466   .0140528   -10.41   0.000    -.1737896   -.1187035
       child |  -.0442597   .0129668    -3.41   0.001    -.0696741   -.0188452
         uni |   .1301115   .0236999     5.49   0.000     .0836605    .1765624
   fem_child |   -.217242   .0257771    -8.43   0.000    -.2677641   -.1667198
     fem_uni |   .1391947   .0445354     3.13   0.002      .051907    .2264824
   child_uni |  -.0120989   .0471321    -0.26   0.797    -.1044761    .0802782
 fem_chi_uni |   .2260262   .0878555     2.57   0.010     .0538327    .3982198
------------------------------------------------------------------------------
```

The effect of the variable `female` shows that the probablity of women to participate in the labor market is about 15 percentage points lower than that of men. The default of `inteff3` is to compute partial effects at means. Hence, the gender difference of 15 percentage points applies to a hypothetical average individual with mean values for all regressors. Having a child is associated with a 4 percentage points lower participation rate, and having a university degree with a 13 percentage points higher participation rate for average individuals.

For the two-fold interaction terms there are two possible interpretations. The interaction effect -.22 of female and children (`fem_child`) means that (i) the gender difference is 22 percentage points larger for average individuals with children compared to similar individuals without children, or that (ii) the negative effect of having a child on participation is by 22 percentage points stronger for females than it is for males[3].

[3] When using the term *effect*, which conveys the notion of causality, we implicitly assume that there is no reversed causality (e.g. labor market participation having an effect on fertility) and no unobserved heterogeneity that would bias our effects from being causal.

The effect for fem_uni shows that (i) for university graduates the gender difference is 14 percentage points smaller than for non-graduates, or (ii) for women the positive effect of a university degree on participation is 14 percentage points stronger than it is for men.

The insignificant effect of child_uni implies that (i) the effect of children on participation does not seem to depend on the university degree of the parents, or (ii) the effect of the university degree on participation does not seem to depend on the presence of children.

One possible interpretation of the triple interaction term is as follows. The effect of children on the gender difference in participation is by about 23 percentage points weaker for women with a university degree compared to women without such a degree. While the presence of children does increase the gender gap in participation (fem_child), it does less so for more highly educated women (fem_chi_uni). This empirical result makes sense economically, because more highly educated women usually have higher opportunity costs (higher wages, more interesting jobs) from not participating in the labor market[4].

When not using the simulated but the real data set the partial effects are qualitatively similar but different in size. They are -.21 for female, -.11 for child, .16 for uni, -.38 for fem_child, .03 for fem_uni, .006 for child_uni and .07 for fem_chi_uni.

Next we compare the output of inteff3 after probit with that of a linear probability model.

```
. reg particip female child uni fem_child fem_uni child_uni fem_chi_uni age
> age_sq german year2 year3 year4 year5 year6 year7 state1 state2 state3 state4
> state6 state7 state8 state9 state10 state11 state12 state13 state14 state15
> state16

      Source |       SS       df       MS              Number of obs =    2000
-------------+------------------------------           F( 31,  1968) =   22.06
       Model |  64.1308182    31  2.06873607           Prob > F      =  0.0000
    Residual |  184.528682  1968  .093764574           R-squared     =  0.2579
```

[4]Here we have chosen to interpret the triple interaction term by asking how a university degree changes our interpretation (i) of the coefficient fem_child. But there are all together 6 possibilities of interpreting the triple interaction term, because for each possible interpretation of a pairwise interaction term we can ask how it changes with the remaining dummy variable. For example, we could have asked how the presence of children affects the interpretation (ii) of fem_uni. Interpretation (ii) of fem_uni was that the positive effect of a university degree on participation is about 14 percentage points stronger for women than it is for men. The triple interaction term then means that this male-female difference in the effect of a university degree is stronger by 23 percentage points if children are present than if they are not present.

```
-------------+------------------------------           Adj R-squared =  0.2462
       Total |    248.6595  1999  .124391946           Root MSE      =  .30621

------------------------------------------------------------------------------
    particip |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      female |  -.0110096   .0513559    -0.21   0.830    -.1117274    .0897081
       child |   .1211511   .0487532     2.48   0.013     .0255379    .2167644
         uni |   .1457012   .0427593     3.41   0.001     .0618428    .2295595
   fem_child |   -.363491   .0652491    -5.57   0.000    -.4914556   -.2355263
     fem_uni |  -.0312669    .056697    -0.55   0.581    -.1424594    .0799257
   child_uni |  -.1126186   .0537618    -2.09   0.036    -.2180547   -.0071826
 fem_chi_uni |   .1911874   .0724121     2.64   0.008      .049175    .3331998
         age |   .0944234   .0058371    16.18   0.000     .0829759     .105871
      age_sq |  -.0011218   .0000759   -14.78   0.000    -.0012706    -.000973
      german |   .0533853   .0247257     2.16   0.031      .004894    .1018766
       year2 |   .0126909   .0191028     0.66   0.507    -.0247729    .0501547
       year3 |  -.0222891   .0195318    -1.14   0.254    -.0605943     .016016
       year4 |   .0137031    .019895     0.69   0.491    -.0253144    .0527206
       year5 |  -.0065777   .0203906    -0.32   0.747     -.046567    .0334117
       year6 |  -.0329195   .0183662    -1.79   0.073    -.0689388    .0030998
       year7 |  -.0119072   .0196214    -0.61   0.544    -.0503881    .0265736
      state1 |   .0006947   .0404722     0.02   0.986    -.0786782    .0800676
      state2 |   .0571352   .0346434     1.65   0.099    -.0108064    .1250768
      state3 |   .0205673   .0661412     0.31   0.756    -.1091468    .1502815
      state4 |   -.009332   .0243071    -0.38   0.701    -.0570023    .0383383
      state6 |   .0508116   .0169564     3.00   0.003     .0175572     .084066
      state7 |  -.0348525   .0292841    -1.19   0.234    -.0922835    .0225786
      state8 |   .0241543   .0289179     0.84   0.404    -.0325585    .0808672
      state9 |   .0090815   .0214628     0.42   0.672    -.0330108    .0511737
     state10 |   .0530236   .0190447     2.78   0.005     .0156737    .0903734
     state11 |  -.0464125   .0310093    -1.50   0.135     -.107227    .0144021
     state12 |  -.0347191   .0425799    -0.82   0.415    -.1182254    .0487873
     state13 |   .0665878   .0358917     1.86   0.064    -.0038019    .1369774
     state14 |   .0573612   .0410576     1.40   0.163    -.0231596    .1378821
     state15 |  -.0053813   .0421607    -0.13   0.898    -.0880656    .0773029
     state16 |  -.0006873   .0294542    -0.02   0.981    -.0584521    .0570775
       _cons |  -1.222554   .1779831    -6.87   0.000    -1.571609   -.8734984
------------------------------------------------------------------------------
```

In the linear regression, the coefficient on `female` is the partial effect for those individuals for whom all variables interacted with `female` take on a value of zero. Hence -.01 is the partial gender effect for individuals without university degree and without children. The gender effect for individuals with children but without university degree is obtained by summing up coefficients on `female` and `fem_child`. It is -.01-.36 = -.37. The gender effect for individuals with children and with university degree is -.01-.36-.03+.19 = -.21. The effect of -.15 of the previous `inteff3` output lies somewhere in between these values. This is normal, as we expect the effect for an average individual computed by `inteff3` to be some weighted average of -.01, -.37 and -.21.

If we wanted to use `inteff3` to compute the gender effect for individuals without children and without university degree, and with mean values on all other regressors, then we have to set the regressor values in `inteff3` manually:

```
. inteff3, at(0.5225 0 0 39.2785 1636.299 0.085 0.846 0.8545 0.8595
> 0.867 0.8305 0.855 0.97 0.9585 0.989 0.9115 0.7915 0.9405 0.9395 0.8825 0.8445
> 0.9475 0.973 0.9615 0.971 0.9725 0.942 1)

Dummies and Interactions: female, child, uni, fem_child, fem_uni, child_uni, fem_chi_uni.
Control variable: age age_sq german year2 year3 year4 year5 year6 year7 state1 state2
> state3 state4 state6 state7 state8 state9 state10 state11 state12 state13 state14
> state15 state16, constant term.

Marginal effect at following values:

__000009[1,3]
       female   child     uni
Values   .5225       0       0

__000008[1,25]
             age    age_sq    german     year2     year3     year4     year5
Values   39.2785  1636.299      .085      .846     .8545     .8595      .867

           year6     year7    state1    state2    state3    state4    state6
Values     .8305      .855       .97     .9585      .989     .9115     .7915

          state7    state8    state9   state10   state11   state12   state13
Values     .9405     .9395     .8825     .8445     .9475      .973     .9615

         state14   state15   state16     _cons
Values      .971     .9725      .942         1
------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      female |  -.0150696   .0669203    -0.23   0.822    -.1462311    .1160918
       child |  -.0294894   .0455244    -0.65   0.517    -.1187156    .0597367
         uni |    .134705    .033966     3.97   0.000     .0681328    .2012772
   fem_child |   -.410151   .0840221    -4.88   0.000    -.5748312   -.2454708
     fem_uni |  -.0168247   .0682101    -0.25   0.805     -.150514    .1168646
   child_uni |  -.0120989   .0471321    -0.26   0.797    -.1044761    .0802782
 fem_chi_uni |   .2260262   .0878555     2.57   0.010     .0538327    .3982198
------------------------------------------------------------------------------
```

Here we get -.015 for the effect of `female`, which is close to the OLS coefficient in the earlier linear regression.

A naive approach to computing the interaction effects might be using `mfx` after `probit`, or `dprobit`. However, these commands do not deliver the desired interaction effects:

```
. dprobit particip female child uni fem_child fem_uni child_uni fem_chi_uni
```

```
> age age_sq german year2 year3 year4 year5 year6 year7 state1 state2 state3
> state4 state6 state7 state8 state9 state10 state11 state12 state13 state14
> state15 state16

Iteration 0:   log likelihood = -829.64661
Iteration 1:   log likelihood = -591.93241
Iteration 2:   log likelihood = -573.11604
Iteration 3:   log likelihood =  -572.0443
Iteration 4:   log likelihood = -572.03799
Iteration 5:   log likelihood = -572.03799

Probit regression, reporting marginal effects           Number of obs =   2000
                                                        LR chi2(31)   = 515.22
                                                        Prob > chi2   = 0.0000
Log likelihood = -572.03799                             Pseudo R2     = 0.3105

------------------------------------------------------------------------------
 particip |      dF/dx   Std. Err.      z    P>|z|     x-bar  [    95% C.I.   ]
---------+--------------------------------------------------------------------
  female*|  -.0082118   .0365744    -0.22   0.823     .5225  -.079896  .063473
   child*|   .1093669   .0496024     2.49   0.013     .6045   .012148  .206586
     uni*|   .2190083   .0708413     4.10   0.000      .809   .080162  .357855
fem_ch~d*|  -.3261532   .0967815    -4.43   0.000     .3055  -.515842 -.136465
 fem_uni*|  -.0506015   .0487959    -1.09   0.277     .4175   -.14624  .045037
child_~i*|  -.0766267   .0501354    -1.59   0.112     .4875   -.17489  .021637
fem_ch~i*|   .0595931   .0400244     1.24   0.215      .243  -.018853   .13804
     age |    .061917   .0054205    12.92   0.000   39.2785   .051293  .072541
  age_sq |  -.0007401   .0000695   -11.83   0.000    1636.3  -.000876 -.000604
  german*|   .0428215   .0142196     2.30   0.022      .085   .014952  .070691
   year2*|   .0126869   .0178172     0.74   0.457      .846  -.022234  .047608
   year3*|  -.0151676   .0150066    -0.95   0.342     .8545   -.04458  .014245
   year4*|   .0035582   .0173291     0.21   0.835     .8595  -.030406  .037523
   year5*|  -.0076483   .0170656    -0.43   0.664      .867  -.041096    .0258
   year6*|  -.0228957   .0133482    -1.57   0.116     .8305  -.049058  .003266
   year7*|  -.0113402   .0154457    -0.70   0.483      .855  -.041613  .018933
  state1*|  -.0094331   .0308558    -0.29   0.772       .97  -.069909  .051043
  state2*|   .0626863   .0392473     1.94   0.052     .9585  -.014237   .13961
  state3*|   .0153135   .0617368     0.27   0.790      .989  -.105688  .136315
  state4*|  -.0069645   .0199152    -0.34   0.735     .9115  -.045998  .032069
  state6*|   .0427229   .0171951     2.77   0.006     .7915   .009021  .076425
  state7*|  -.0305984   .0209289    -1.21   0.227     .9405  -.071618  .010422
  state8*|    .010721   .0256829     0.44   0.662     .9395  -.039617  .061059
  state9*|   .0091683   .0192088     0.49   0.621     .8825   -.02848  .046817
 state10*|   .0479791   .0200612     2.74   0.006     .8445    .00866  .087298
 state11*|  -.0223311   .0215186    -0.91   0.361     .9475  -.064507  .019845
 state12*|  -.0221846   .0296781    -0.65   0.515      .973  -.080353  .035984
 state13*|   .0583521    .040506     1.74   0.082     .9615  -.021038  .137742
 state14*|   .0451065   .0440765     1.20   0.229      .971  -.041282  .131495
 state15*|    -.00057   .0341713    -0.02   0.987     .9725  -.067545  .066405
 state16*|  -.0055981   .0248373    -0.22   0.826      .942  -.054278  .043082
---------+--------------------------------------------------------------------
  obs. P |      .8545
 pred. P |   .9274498  (at x-bar)
------------------------------------------------------------------------------
(*) dF/dx is for discrete change of dummy variable from 0 to 1
    z and P>|z| correspond to the test of the underlying coefficient being 0
```

For example, here the effect associated with the triple interaction is .06 and it is not statistically significant. Such a result would have suggested the conclusion that the increase of the gender difference in participation due to the presence of children does not depend on education. The command `dprobit` computes

$$\begin{aligned}\frac{\Delta^3 F(\mathbf{x}\beta)}{\Delta(x_1 x_2 x_3)} &= F(\beta_1+\beta_2+\beta_3+\beta_{12}+\beta_{13}+\beta_{23}+\beta_{123}+\tilde{\mathbf{x}}\tilde{\beta}) \\ &\quad -F(\beta_1+\beta_2+\beta_3+\beta_{12}+\beta_{13}+\beta_{23}+\tilde{\mathbf{x}}\tilde{\beta}). \qquad (6)\end{aligned}$$

In the empirical example we were interested in the interaction effect given in (4). The effect in (6) is very different. In general there is no guarantee that (4) and (6) are of equal sign[5].

Above we demonstrated the use of `inteff3` to compute effects at means or at certain regressor values. The program also allows to compute the partial effects for each individual in the sample and to average these effects. According to Greene (2003, p. 668) this is more advisable than just computing the effect at means. This is possible with `inteff3` by specifying:

```
. inteff3, average pex1(pe1) pex1x2x3(pe123)

Dummies and Interactions: female, child, uni, fem_child, fem_uni, child_uni, fem_chi_uni.
Control variable: age age_sq german year2 year3 year4 year5 year6 year7 state1 state2
> state3 state4 state6 state7 state8 state9 state10 state11 state12 state13 state14
> state15 state16, constant term.

Average marginal effect:
------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      female | -.1639777   .0132023   -12.42   0.000    -.1898537   -.1381017
       child |  -.079771   .0130652    -6.11   0.000    -.1053784   -.0541637
         uni |  .1274121   .0194422     6.55   0.000     .0893061    .1655182
   fem_child |  -.212257   .0255793    -8.30   0.000    -.2623915   -.1621226
     fem_uni |  .0843261   .0387943     2.17   0.030     .0082906    .1603616
   child_uni | -.0075491   .0397794    -0.19   0.849    -.0855153    .0704171
 fem_chi_uni |  .1850893   .5207804     0.36   0.722    -.8356215      1.2058
------------------------------------------------------------------------------
```

The estimates now differ to some extent from those computed at means[6].

A more complete description of the sample distribution of the estimated effects, than

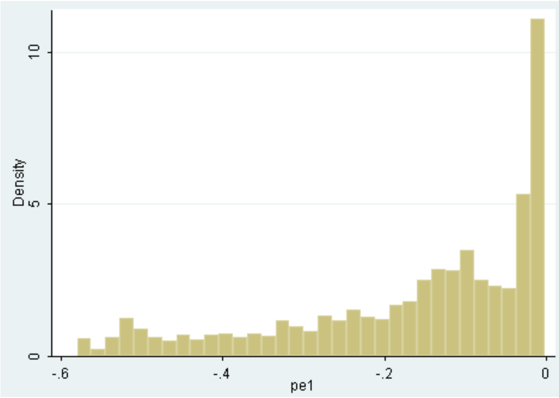
[5] Equation (6) is useful, however, because in a difference-in-difference-in-differences model it represents the treatment effect (see Puhani 2008).

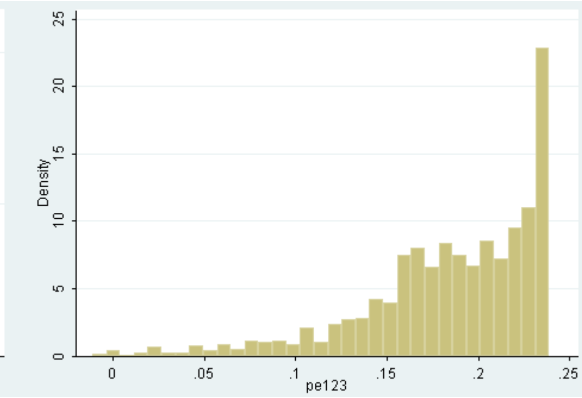
[6] When not using the simulated but the real data set the results are: -.19 for `female`, -.11 for `child`, .14 for `uni`, -.34 for `fem_child`, .02 for `fem_uni`, -.01 for `child_uni` and .06 for `fem_chi_uni`, all except `fem_uni` and `child_uni` being significant at the 1% level.

just reporting the average, would be to report quantiles or to graph the distribution of the effects. The options `pex1()` and `pex1x2x3()` used here save the individual effects of equations (2) and (4) as variables and allow to describe or graph their distribution. The histograms for the effects saved as `pe1` (partial effect of `female`) and `pe123` (partial effect of `fem_chi_uni`) uncover a large amount of heterogeneity:

```
histogram pe1
```

```
histogram pe123
```





4 Conclusion

This paper has derived the partial effects in probit and logit models with three interacted dummy variables. The computation of the partial effects and their standard errors has been implemented in the Stata program `inteff3` which applies the delta method to compute the standard errors of the partial effects. We have demonstrated the use of `inteff3` by means of a probit regression of labor market participation. We have included dummies for female gender, university degree, presence of children, as well as their pairwise and triple interaction terms. This allows to analyse the way a university degree and the presence of children affect the gender difference in labor market participation. We find evidence consistent with the idea that the presence of children increases the gender gap in labor market participation, but that this increase is smaller for more highly educated individuals.

In an analogous way as presented here and as presented in Ai and Norton (2003) and Norton et al. (2004), the effects can be computed for the case of an interaction of three continuous variables or for a mixture of continuous and dummy variables.

References

[1] Ai, C. and E. C. Norton (2003): Interaction terms in logit and probit models, *Economics Letters*, 80, pp. 123-129.

[2] Davidson, R. and J. MacKinnon (2004): *Econometric Theory and Methods*, Oxford University Press, New York.

[3] Greene, W. (2003): *Econometric Analysis*, Prentice Hall, New Jersey.

[4] Gruber, J. (1994): The Incidence of Mandated Maternity Benefits, *The American Economic Review*, Vol. 84, No. 3, pp. 622-641.

[5] Gruber, J. and J. Poterba (1994): Tax Incentives and the Decision to Purchase Health Insurance: Evidence from the Self-Employed, *The Quarterly Journal of Economics*, Vol. 109, No. 3, pp. 701-733.

[6] Norton, E. C., H. Wang and C. Ai (2004): Computing interaction effects and standard errors in logit and probit models, *The Stata Journal*, 4(2), 154-167.

[7] Puhani, P. (2008): The Treatment Effect, the Cross Difference, and the Interaction Term in Nonlinear "Difference-in-Differences" Models, *IZA Discussion Paper No. 3478.*

Appendix

Let g_1 denote the difference $\frac{\Delta F(\mathbf{x}\beta)}{\Delta x_1}$ given in (2). The derivatives of g_1 with respect to β_1, β_{12}, β_{123} and a coefficient β_j (part of $x\tilde{\beta}$) are given by:

$$\begin{aligned}
\frac{\partial g_1}{\partial \beta_1} &= f(\beta_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_2 + \beta_{13} x_3 + \beta_{23} x_2 x_3 + \beta_{123} x_2 x_3 + \tilde{\mathbf{x}}\tilde{\beta}) \\
\frac{\partial g_1}{\partial \beta_{12}} &= f(\beta_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_2 + \beta_{13} x_3 + \beta_{23} x_2 x_3 + \beta_{123} x_2 x_3 + \tilde{\mathbf{x}}\tilde{\beta}) x_2 \\
\frac{\partial g_1}{\partial \beta_{123}} &= f(\beta_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_2 + \beta_{13} x_3 + \beta_{23} x_2 x_3 + \beta_{123} x_2 x_3 + \tilde{\mathbf{x}}\tilde{\beta}) x_2 x_3 \\
\frac{\partial g_1}{\partial \beta_j} &= (f(\beta_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_2 + \beta_{13} x_3 + \beta_{23} x_2 x_3 + \beta_{123} x_2 x_3 + \tilde{\mathbf{x}}\tilde{\beta}) \\
&\quad - f(\beta_2 x_2 + \beta_3 x_3 + \beta_{23} x_2 x_3 + \tilde{\mathbf{x}}\tilde{\beta})) x_j
\end{aligned}$$