# Interpreting logit models

Luca J. Uberti
Department of Economics and Management
University of Luxembourg
Esch-sur-Alzette, Luxembourg
luca.uberti@uni.lu

**Abstract.**   The parameters of logit models are typically difficult to interpret, and the applied literature is replete with interpretive and computational mistakes. In this article, I review a menu of options to interpret the results of logistic regressions correctly and effectively using Stata. I consider marginal effects, partial effects, (contrasts of) predictive margins, elasticities, and odds and risk ratios. I also show that interaction terms are typically easier to interpret in practice than implied by the recent literature on this topic.

**Keywords:** st0663, logit, binary outcome models, nonlinear models, interpretation, interaction terms

## 1   Introduction

Applied researchers estimate logit models to infer the causal effect of one or more treatments (for instance, infertility) on the probability of a binary outcome (joining the labor force) for a population of interest (married women). In contrast to the linear probability model (LPM), the logit model always produces predicted probabilities within the meaningful range $[0, 1]$. "From a practical perspective," however, "the most difficult aspect of logit [. . .] models is presenting and interpreting the results" (Wooldridge 2020, chap. 17). Researchers often mislabel their findings or fail to present them in an intuitive and compelling fashion. Computational mistakes are not uncommon. This pedagogical review article summarizes and illustrates a menu of options to interpret the coefficients of logit models correctly and effectively.[1] It covers approaches—for instance, elasticities and risk ratios—that are rarely discussed systematically in econometrics textbooks. In addition, it shows that interaction terms in logit models are much easier to interpret in practice than implied by the recent literature on this topic (Ai and Norton 2003; Norton, Wang, and Ai 2004). Readers may refer to the more primary sources cited here for additional details.

Conditional on a vector $\mathbf{X}$ of covariates, the expected value of a binary dependent variable $(Y)$ equals the probability of a positive outcome: $E(Y_i|\mathbf{X}_i) = 1 \times P(Y_i = 1|\mathbf{X}_i) + 0 \times [1 - P(Y_i = 1|\mathbf{X}_i)] = P(Y_i = 1|\mathbf{X}_i)$. In the LPM, this conditional probability is given a linear specification,

$$P(Y_i = 1|\mathbf{X}_i) = \mathbf{X}_i\boldsymbol{\beta}$$

---

1. Unless otherwise stated, all the arguments and examples presented here carry over, mutatis mutandis, to the case of other binary outcome models such as probit.

and the coefficients ($\boldsymbol{\beta}$) are estimated by ordinary least squares. In reality, however, probabilities cannot exceed the $[0, 1]$ interval by definition, and the data-generating process is nonlinear. Thus, a commonly used functional form is the logistic function, which is sigmoid shaped and has limit 0 and 1 as $\mathbf{X}_i\boldsymbol{\beta}$ tends to $-\infty$ and $+\infty$, respectively:

$$P(Y_i = 1|\mathbf{X}_i) = \frac{1}{1 + e^{-\mathbf{X}_i\boldsymbol{\beta}}} \tag{1}$$

In Stata, the maximum likelihood estimates of the $\boldsymbol{\beta}$'s in (1) can be obtained with the `logit` command.[2]

To illustrate, I estimate a labor supply model using survey data on 753 married women (Mroz 1987). `inlf` ("in labor force"), the outcome, is an indicator variable that takes the value 1 if a woman reports working for a wage and 0 otherwise. `nokid` takes the value 1 if a woman has no children younger than six years old at the time of the survey and 0 otherwise.[3] `nwifeinc` ("nonwife income") measures the husband's annual earnings, in thousands of euros.

```
. webuse mroz
. generate nokid = (kidslt6==0)
. logit inlf nwifeinc i.nokid, robust nolog
```

```
Logistic regression                              Number of obs =      753
                                                 Wald chi2(2)  =    37.78
                                                 Prob > chi2   =   0.0000
Log pseudolikelihood = -493.81541                Pseudo R2     =   0.0409
```

| inlf | Coefficient | Robust std. err. | z | P>\|z\| | [95% conf. interval] | |
|---|---|---|---|---|---|---|
| nwifeinc | -.0210956 | .0071266 | -2.96 | 0.003 | -.0350635 | -.0071278 |
| 1.nokid | 1.060382 | .1924336 | 5.51 | 0.000 | .683219 | 1.437545 |
| _cons | -.1496803 | .2195877 | -0.68 | 0.495 | -.5800644 | .2807037 |

# 2 Marginal effects (continuous covariates)

The impact on $P(Y_i = 1|x_i, \mathbf{X}_i)$ of a "marginal" (infinitesimally small) change in $x_i$, a continuous variable, can be obtained by differentiating (1) with respect to $x_i$,

$$\frac{\partial P}{\partial x_i} = \beta \left[ \frac{e^{-x_i\beta - \mathbf{X}_i\boldsymbol{\beta}}}{(1 + e^{-x_i\beta - \mathbf{X}_i\boldsymbol{\beta}})^2} \right] \tag{2}$$

where $\beta$ denotes the coefficient on $x_i$. This quantity measures the "marginal effects" of $x_i$ (see also Williams [2012] for an introductory account). First, note that in contrast to the linear probability model, where $\partial P/\partial x_i = \beta$, the value of $\partial P/\partial x_i$ given by (2)

---

2. Estimation by nonlinear least squares is much less common.

3. For identification, we assume that the husband's earnings depend solely on his own education and labor market experience and that a woman's childlessness results from a randomly assigned treatment—for instance, medical complications or genetic conditions.

is not equal to $\beta$, because it also depends on the scale factor in brackets. Thus, the coefficients reported in the `logit` output are not marginal effects and do not have a probability interpretation. Second, note that the scale factor is a function of $x_i$, the variable of interest, and the other covariates. Thus, $\partial P/\partial x_i$ is not constant as in the LPM but varies for different values of $x_i$ and $\mathbf{X}_i$. To obtain a summary statistic, we could evaluate $\partial P/\partial x_i$ at the mean values of $x_i$ and $\mathbf{X}_i$. This is called marginal effects at the means. Until recently, this was the most common approach in applications:[4]

```
. quietly margins, dydx(nwifeinc) atmeans
```

Yet, when some of the covariates (`nokid`) are binary, the sample means do not describe any particular woman in the sample. Thus, it makes better sense to "average over" the covariates—that is, compute $\partial P/\partial x_i$ at the values of $x_i$ and $\mathbf{X}_i$ actually observed in each woman in the sample (`asobserved` in Stata jargon) and then take an average across observations. The quantity thus obtained measures the average marginal effects. This is now Stata's default option:

```
. margins, dydx(nwifeinc)
Average marginal effects                                Number of obs = 753
Model VCE: Robust

Expression: Pr(inlf), predict()
dy/dx wrt:  nwifeinc
```

|          |         | Delta-method |       |       |                      |           |
|---------:|--------:|-------------:|------:|------:|---------------------:|----------:|
|          | dy/dx   | std. err.    | z     | P>\|z\| | [95% conf. interval] |           |
| nwifeinc | -.0048879 | .0016117   | -3.03 | 0.002 | -.0080467          | -.0017291 |

On average, a 1,000-euro increase in a married man's annual income leads to a 0.49 percentage-point drop in the probability that his wife might join the labor force.

# 3   Predicted probabilities (continuous covariates)

Very often, a more informative way to convey the results is to plot $\widehat{P}(Y_i = 1|\mathbf{X}_i)$, the predicted probability of being in the labor force, conditional on selected values of covariates. Here we allow `nwifeinc` to range between 0 and 100,000 euros (the sample range), distinguishing between women with and without children. Users could graph $\widehat{P}$ using a combination of `margins` and `marginsplot` (see StataCorp [2021, 1501–1504] for details). Yet a more flexible, user-friendly approach is provided by Royston's (2013) `mcp` (or `marginscontplot`) command, which may be downloaded from the *Stata Journal* website (type `net search marginscontplot` and select `gr0056`). This command plots the response variable (here the probability that `inlf` = 1) at given values of the predictors, averaging over the remaining covariates, if there are any. The fixed values of the predictors can be specified with the `at1`(*numlist*) and `at2`(*numlist*) options, which in this example refer to `nwifeinc` and `nokid`, respectively:

---

4. Part of the reason for this is that earlier Stata commands, such as `mfx` and `dprobit`, allowed only this option.

```
. mcp nwifeinc nokid, at1(0(10)100) at2(0 1) ci plotopts(ycommon)
```
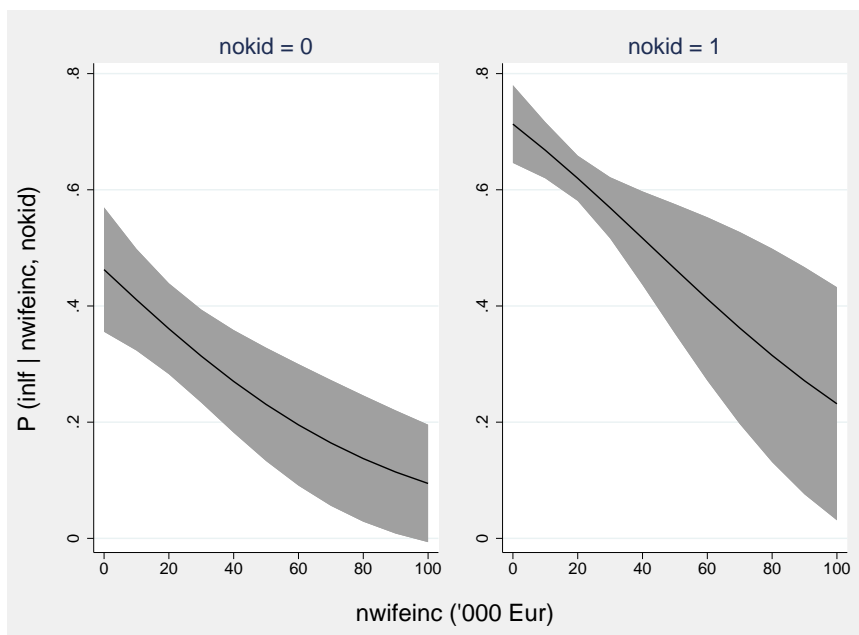


Figure 1. Predicted probabilities of being in the labor force (`inlf` = 1) with 95% confidence interval

Figure 1[5] allows us to read off the vertical axes the probability of labor-force participation at given values of `nokid` and `nwifeinc` and to clearly visualize the negative effect of `nwifeinc` (the downward slope of the curves).

As an alternative (or in addition) to plotting a diagram, users can also obtain a table of predicted probabilities. This can be accomplished with `margins`—or, more conveniently, with Long and Freese's (2014) `mtable` command, which is part of the `spost13` suite of commands downloadable from J. Scott Long's Indiana University webpage (type `net search spost13` and select `spost13_ado`). This command uses `margins` behind the scenes to construct tables of predictions. As always, users can specify `at(`*atspec*`)` options to indicate which predicted probabilities should be displayed and in what order. For instance, some of the probabilities shown on the vertical axis of figure 1 may be obtained by typing

---

5. I created the graph with the command shown and then enhanced the *y*-axis and *x*-axis titles of the resulting graph using the Graph Editor.

```
. mtable, at(nokid = (0) nwifeinc = (0(20)100))
> at(nokid = (1) nwifeinc = (0(20)100)) brief
Expression: Pr(inlf), predict()
```

|    | nwifeinc | nokid | Pr(y) |
|----|----------|-------|-------|
| 1  | 0        | 0     | 0.463 |
| 2  | 20       | 0     | 0.361 |
| 3  | 40       | 0     | 0.270 |
| 4  | 60       | 0     | 0.195 |
| 5  | 80       | 0     | 0.137 |
| 6  | 100      | 0     | 0.095 |
| 7  | 0        | 1     | 0.713 |
| 8  | 20       | 1     | 0.620 |
| 9  | 40       | 1     | 0.517 |
| 10 | 60       | 1     | 0.412 |
| 11 | 80       | 1     | 0.315 |
| 12 | 100      | 1     | 0.232 |

For instance, the probability that a childless (`nokid` = 1) woman with a low-income husband (`nwifeinc` = 20) is economically active is 62%.

# 4   Logit versus LPM

Note that, consistent with (2), the slope of the curves shown in figure 1—the marginal effects of `nwifeinc`—is a function of `nwifeinc`. Yet, in this application, it is clear that the marginal effects are fairly constant across most of the distribution of earnings, implying a near-linear relationship between $P$ and `nwifeinc`.[6] Thus, in this particular case, a simple LPM specification might provide a satisfactory approximation of the true relationship:

```
. regress inlf nwifeinc i.nokid, robust
  (output omitted)
```

The estimated coefficient on `nwifeinc` ($-0.00479$) is fairly close to the average marginal effects implied by the `logit` model ($-0.00489$). The predicted probabilities are also very similar across the two models, although only the logit probabilities are strictly within the meaningful 0–1 range:

```
. quietly logit inlf nwifeinc i.nokid, robust nolog
. predict P_logit, pr
. quietly regress inlf nwifeinc i.nokid, robust
. predict P_LPM, xb
. summarize P_logit P_LPM
```

| Variable |  Obs |     Mean | Std. dev. |       Min |      Max |
|----------|------|----------|-----------|-----------|----------|
| P_logit  |  753 | .5683931 |   .116784 | .1020306  | .7132691 |
| P_LPM    |  753 | .5683931 | .1166285  | -.0003462 | .7147611 |

---

6. In this dataset, 95% of sampled women have husbands with annual income below 41,100 euros.

```
. correlate P_logit P_LPM
(obs=753)
                 |   P_logit      P_LPM
-----------------+------------------------
        P_logit  |    1.0000
          P_LPM  |    0.9987     1.0000
```

Of course, when the data-generating process is highly nonlinear, the discrepancy between LPM and logit may be much greater, and the predicted probabilities obtained from LPM highly misleading, if meaningful at all (see also Allison [2017]).

# 5   Partial effects (continuous covariates)

In nonlinear models, marginal effects provide a good approximation of the impact of a change in $x_i$ only if the change is "small". Suppose we would like to know how much less likely a woman is to join the labor force if her husband's annual income goes up by one standard deviation ($\delta = 11635$ euros). In applications, a very common approach is to multiply 0.49 by 11.63 and conclude that $P$ declines by 5.70 percentage points. Yet, since 11,635 is far from a "marginal" increment, this computation is technically incorrect. For a discrete change in $x_i$, researchers should estimate the (average) difference in predicted probabilities:[7]

$$P(Y_i = 1|x_i + \delta, \mathbf{X}_i) - P(Y_i = 1|x_i, \mathbf{X}_i) = \frac{1}{1 + e^{-[(x_i+\delta)\beta + \mathbf{X}_i\boldsymbol{\beta}]}} - \frac{1}{1 + e^{-(x_i\beta + \mathbf{X}_i\boldsymbol{\beta})}}$$

This calculation cannot be easily accomplished using `margins`. Instead, researchers can use the `mchange` command developed by Long and Freese (2014).[8] Exploiting the `margins` command behind the scenes, `mchange` returns the effects of marginal or discrete changes in a specified treatment variable (here `nwifeinc`), while providing a user-friendly interface. For marginal effects, users can request the `amount(marginal)` option. For the partial effect of a standard-deviation change ($= \delta$) in the specified regressor, the `amount(sd)` option should be invoked:

```
. quietly logit inlf nwifeinc i.nokid, robust nolog
. mchange nwifeinc, amount(sd) stats(change se zvalue pvalue ll ul) decimals(4)
> brief
logit: Changes in Pr(y) | Number of obs = 753
Expression: Pr(inlf), predict(pr)
                 |   Change    Std Err   z-value    p-value        LL         UL
-----------------+-----------------------------------------------------------------
nwifeinc         |
            +SD  |  -0.0576    0.0192    -3.0074     0.0026    -0.0951    -0.0201
```

For other discrete changes, users can customize the magnitude of $\delta$ by specifying the `delta(#)` option, which overrides `amount(sd)`.

---

7. Of course, as a (less advisable) alternative, researchers can also compute the difference in predicted probability at the means.

8. This command is also part of the `spost13` suite of commands. See section 3.

In this example, the correctly estimated partial effect (5.76) is close to the figure obtained by multiplying the average marginal effects by the standard deviation (5.70). Yet, when the estimated conditional probability is highly nonlinear in the covariates, this back-of-the-envelope calculation may produce seriously misleading estimates. Users are therefore advised to always compute the correct partial effects as the default option.

# 6    Margins and contrasts of margins (binary covariates)

When the treatment $x_i$ is binary (for example, `nokid`), or more generally a factor variable, researchers can evaluate the magnitude of the estimated effect in at least two ways. First, they may compute the predicted probability of a positive outcome ("predictive" or "conditional margins" in Stata jargon), treating all the women in the sample as if they had $[\widehat{P}(Y_i = 1|x_i = 1, \mathbf{X}_i)]$ or had no kids $[\widehat{P}(Y_i = 1|x_i = 0, \mathbf{X}_i)]$. Second, researchers can calculate the difference between these two predicted probabilities, which is known in Stata jargon as the "contrast of margins".[9] Again, users can either average over the covariates $(\mathbf{X}_i)$, as we do here, or alternatively hold them fixed at specific values by requesting the `at(`*atspec*`)` or `atmeans` option:

```
. margins i.nokid
Predictive margins                                      Number of obs = 753
Model VCE: Robust
Expression: Pr(inlf), predict()
```

|  |  | Delta-method |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Margin | std. err. | z | P>\|z\| | [95% conf. interval] | |
| nokid |  |  |  |  |  |  |
| 0 | .3624543 | .0394764 | 9.18 | 0.000 | .2850821 | .4398266 |
| 1 | .6181144 | .0196075 | 31.52 | 0.000 | .5796844 | .6565445 |

```
. margins r.nokid
Contrasts of predictive margins                         Number of obs = 753
Model VCE: Robust
Expression: Pr(inlf), predict()
```

|  | df | chi2 | P>chi2 |
| --- | --- | --- | --- |
| nokid | 1 | 33.64 | 0.0000 |

|  | Contrast | Delta-method std. err. | [95% conf. interval] | |
| --- | --- | --- | --- | --- |
| nokid |  |  |  |  |
| (1 vs 0) | .2556601 | .044078 | .1692688 | .3420514 |

---

9. For a categorical or ordinal variable with more than two levels, researchers can of course compute the contrast between each of the factor levels and the base (reference) level ($x_i = 0$).

While the `i.` prefix requests an indicator for each level (category) of the variable `nokid`, the `r.` prefix requests a contrast between each level (in this case, just `nokid = 1`) and the reference category (`nokid = 0`). On average, 62 out of 100 women without kids join the labor force, while only 36 women with kids do so (predictive margins). Accordingly, childlessness increases a women's likelihood of joining the labor force by almost 26 percentage points (contrast of margins). This contrast is significantly different from zero.

Note that the contrast of margins is typically defined as the marginal effects of the factor variable. Thus, an identical estimate may be obtained by typing `margins, dydx(nokid)`. Here the Stata output helpfully reminds users that `dy/dx for factor levels is the discrete change from the base level`, rather than the instantaneous (marginal) rate of change.[10] A very common mistake, which is also to be found in the Stata manual (for instance, see StataCorp [2021, 1468]), is to conclude based on the estimated contrast of margins that childless women are almost 26% more likely to join the labor force. This is wrong because it implies an elasticity rather than a marginal effect.

# 7 Elasticities (continuous and binary covariates)

In some cases, it may be more appropriate or informative to report the relative (that is, percentage) change in probability, rather than the absolute change, induced by a treatment. If $x_i$ is continuous (for example, `nwifeinc`), the percentage change is given by the (semi)elasticity:

$$\frac{\partial P}{\partial x_i}\frac{1}{P} = \frac{\partial \ln P}{\partial x_i} = \beta\left(\frac{e^{-x_i\beta-\mathbf{X}_i\boldsymbol{\beta}}}{1+e^{-x_i\beta-\mathbf{X}_i\boldsymbol{\beta}}}\right)$$

```
. margins, eydx(nwifeinc)

Average marginal effects                          Number of obs = 753
Model VCE: Robust

Expression: Pr(inlf), predict()
ey/dx wrt:  nwifeinc
```

|          | ey/dx    | Delta-method std. err. | z     | P>\|z\| | [95% conf. interval] |           |
|----------|----------|------------------------|-------|--------|----------------------|-----------|
| nwifeinc | −.009105 | .0031042               | −2.93 | 0.003  | −.0151891            | −.0030209 |

On average, a 1,000-euro increase in the husband's income leads to a 0.91% decrease in the probability that the wife might join the labor force.

---

10. Users should be careful to specify `nokid` as a factor variable in the preceding estimation command. Without the `i.` prefix in the `logit` syntax, `nokid` would be treated by `margins` as a continuous variable. Thus, its coefficient would be computed based on (2), rather than as a contrast of predictive margins.

When the treatment variable (for example, `nokid`) is binary, the strict definition of relative change leads to the following expression for the response:

$$\frac{P(Y_i|x_i = 1, \mathbf{X}_i) - P(Y_i|x_i = 0, \mathbf{X}_i)}{P(Y_i|x_i = 0, \mathbf{X}_i)} = \frac{1 + e^{-\mathbf{X}_i\boldsymbol{\beta}}}{1 + e^{-\beta - \mathbf{X}_i\boldsymbol{\beta}}} - 1 \tag{3}$$

For a small $\Delta P$, log differences approximate percentage changes. Thus, an alternative formula for the response is

$$\ln P(Y_i|x_i = 1, \mathbf{X}_i) - \ln P(Y_i|x_i = 0, \mathbf{X}_i) = \ln(1 + e^{-\mathbf{X}_i\boldsymbol{\beta}}) - \ln(1 + e^{-\beta - \mathbf{X}_i\boldsymbol{\beta}}) \tag{4}$$

Perhaps unfortunately, Stata uses (4), not (3), to run `margins, eydx(i.nokid)` after `logit`.[11] There is no shorthand command to compute (3), and users wishing to estimate the "exact" definition of relative change should code this expression manually using `margins`'s `expression(`*pnl_exp*`)` option (see also Uberti [2017]). I illustrate this procedure more explicitly in sections 8 and 9 below.

Estimates based on these definitions of "relative change" are presented in table 1. Based on (3), childless women are 72% more likely to join the labor force. For each woman, the change in probability is measured relative to the base level—that is, the probability of a positive outcome for that woman assuming that she had kids ($x_i = 0$). Then, the percentage changes for all the different women in the sample are averaged. Based on (4), childless women are 54% more likely to participate in the labor force. In this case, $\Delta P$ ($= 0.256$) is not exactly "small", and so (4) does not provide a good approximation of percentage change.

Table 1. Alternative measures of "relative change"

| Response formula | Estimate | (s.e.) |
|---|---|---|
| (3) $N^{-1}\Sigma\{\Delta P/P(Y_i|x_i = 0, \mathbf{X}_i)\}$ | 0.721 | (0.199) |
| (4) $N^{-1}\Sigma(\Delta \ln P)$ | 0.541 | (0.115) |
| (3a) $N^{-1}\Sigma(\Delta P)/\overline{P}(Y_i|x_i, \mathbf{X}_i)$ | 0.450 | (0.078) |
| (3b) $N^{-1}\Sigma\{\Delta P/P(Y_i|x_i, \mathbf{X}_i)\}$ | 0.477 | (0.093) |

NOTES: $\Delta P = P(Y_i|x_i = 1, \mathbf{X}_i) - P(Y_i|x_i = 0, \mathbf{X}_i)$. $N$ refers to the number of observations, and the summation operator is over all observations $i$.

There are more ways to define the notion of relative change. Researchers, for instance, could measure the average change in probability (that is, the contrast of margins) relative to the average observed probability $\overline{P}$ of a positive outcome (the mean of the dependent variable). This approach, denoted as (3a) in table 1, is common in applications because it is easy to compute. Lastly, researchers may compare the change in

---

11. Personal communication with Stata Technical Support (January 2021).

probability for each woman with the expected probability of a positive outcome for that woman (who may or may not have kids) and then take an average (3b).

The estimates of (3a) and (3b) are quite similar to each other, and to the log difference (4). Yet they are all considerably smaller in magnitude than the estimate of average percentage change strictly defined—that is, (3). Researchers should decide case by case which number is most meaningful, if any.

# 8   Odds and risk ratios (binary covariates)

So far, the effects of binary treatments were presented on an additive scale. Sometimes, it may be more informative to present the results on a multiplicative scale (Buis 2010). For the logit model, the most common options are odds and risk ratios. The odds are defined as the ratio of the probability of an event occurring to the probability of it not occurring: $P(Y_i = 1|x_i, \mathbf{X}_i)/P(Y_i = 0|x_i, \mathbf{X}_i)$, where $x_i$ is a binary variable. Thus, the odds of a woman joining the labor force are the expected number of women in the labor force for every woman that is not in the labor force. It follows that, for a change in $x_i$ from 0 to 1, the odds ratio is given by

$$\text{Odds ratio} = \frac{P(Y_i = 1|x_i = 1, \mathbf{X}_i)/P(Y_i = 0|x_i = 1, \mathbf{X}_i)}{P(Y_i = 1|x_i = 0, \mathbf{X}_i)/P(Y_i = 0|x_i = 0, \mathbf{X}_i)} = e^{\beta} \tag{5}$$

where $\beta$ is the coefficient on $x_i$. This expression is attractive because it may be computed by simply exponentiating $\beta$.[12] To see this, simply substitute (1) into (5), and simplify the fractions. Note that, from (5), it also follows that the point estimates returned by `logit` (the $\beta$'s) are the odds ratios in logs, or log odds, which do not admit of a straightforward interpretation.

For `nokid`, the odds ratio can be obtained in postestimation simply by typing

```
. nlcom exp(_b[1.nokid])
        _nl_1: exp(_b[1.nokid])
```

| inlf | Coefficient | Std. err. | z | P>|z| | [95% conf. interval] |
|---|---|---|---|---|---|
| _nl_1 | 2.887473 | .5556469 | 5.20 | 0.000 | 1.798426   3.976521 |

The odds of being in the labor force—the number of women in the labor market for every woman in home production—are almost three times higher for childless women than for women with children. Alternatively, users could call the `logit` command with the `or` option (or type `logistic` instead of `logit`). These commands return all the estimated coefficients in exponentiated form. The problem, of course, is that some of the variables may be continuous, and exponentiating the coefficient of a continuous variable is not very useful for interpretation.

---

12. This result does not carry over to `probit`.

Often, reporting the ratio of two probabilities—the risk ratio—may be preferable to reporting the ratio of two odds:

$$\text{Risk ratio} = \frac{P(Y_i = 1|x_i = 1, \mathbf{X}_i)}{P(Y_i = 1|x_i = 0, \mathbf{X}_i)} = \frac{1 + e^{-\mathbf{X}_i\boldsymbol{\beta}}}{1 + e^{-\beta - \mathbf{X}_i\boldsymbol{\beta}}}$$

Note that this quantity equals one plus the relative change given by (3). Just like relative changes, risk ratios are laborious to calculate in Stata:

```
. margins, expression((1+exp(-nwifeinc*_b[nwifeinc]-0*_b[1.nokid]-_b[_cons])) /
> (1+exp(-nwifeinc*_b[nwifeinc]-1*_b[1.nokid]-_b[_cons]))))
warning: option expression() does not contain option predict() or xb().

Predictive margins                                       Number of obs = 753
Model VCE: Robust

Expression: (1+exp(-nwifeinc*_b[nwifeinc]-0*_b[1.nokid]-_b[_cons])) /
            (1+exp(-nwifeinc*_b[nwifeinc]-1*_b[1.nokid]-_b[_cons]))
```

|       |   Margin | Delta-method std. err. |    z | P>\|z\| | [95% conf. interval] |          |
|-------|----------|------------------------|------|---------|----------------------|----------|
| _cons | 1.720799 |              .1988242  | 8.65 |  0.000  |             1.331111 | 2.110487 |

On the upside, they often provide a clear and compelling formulation of the results: childless women are 1.72 times (72%) more likely to be in the labor force than women with children. When the predicted probabilities are small (which is not the case in our example), odds ratios approximate risk ratios (Buis 2010). In general, however, they are different quantities and should not be confused.

# 9   Interaction terms

Binary choice models may also include interaction (or higher-order) terms. For instance, the effect of the husband's income on a woman's labor-supply decision may be thought to depend on the woman's level of education. While women with few years of education (`educ`) may leave the labor market when the husband begins to earn well, highly educated women with career ambitions may be more likely to stay on irrespective of their husband's income. This argument motivates the inclusion of a multiplicative interaction term between the treatment of interest (`nwifeinc`) and the effect modifier (`educ`), leading to

$$P(Y_i = 1|\mathbf{X}_i) = L(u_i) = L(\beta_1 \texttt{nwifeinc}_i + \beta_2 \texttt{educ}_i + \beta_{12} \texttt{nwifeinc}_i$$
$$\times \texttt{educ}_i + \gamma \texttt{nokid}_i) \tag{6}$$

where $L$ denotes the logistic function $(= 1/1 + e^{-u})$ and $u$ is a linear index of the covariates:

```
. logit inlf c.nwifeinc c.educ c.nwifeinc#c.educ i.nokid, robust nolog

Logistic regression                             Number of obs =      753
                                                Wald chi2(4)  =    75.22
                                                Prob > chi2   =   0.0000
Log pseudolikelihood = -468.34427               Pseudo R2     =   0.0904
```

| inlf | Coefficient | Robust std. err. | z | P>\|z\| | [95% conf. interval] | |
|---|---|---|---|---|---|---|
| nwifeinc | −.0477618 | .0503386 | −0.95 | 0.343 | −.1464236 | .0509001 |
| educ | .249319 | .0798804 | 3.12 | 0.002 | .0927563 | .4058818 |
| c.nwifeinc# c.educ | .0007658 | .003817 | 0.20 | 0.841 | −.0067154 | .008247 |
| 1.nokid | 1.2594 | .2016902 | 6.24 | 0.000 | .864094 | 1.654705 |
| _cons | −3.015328 | 1.021068 | −2.95 | 0.003 | −5.016584 | −1.014072 |

```
. testparm nwifeinc c.nwifeinc#c.educ

 ( 1)  [inlf]nwifeinc = 0
 ( 2)  [inlf]c.nwifeinc#c.educ = 0

         chi2(  2) =    20.11
       Prob > chi2 =    0.0000
```

A Wald-type test suggests that `nwifeinc` and `nwifeinc × educ` are jointly significant. Relying on intuitions carried over from linear models, we may find it tempting to look at the positive sign on the estimated coefficient on `nwifeinc × educ` and conclude that an additional year of education lessens the negative impact of the husband's earnings on the wife's labor-supply decision, in line with expectations. Yet it is now well known that, in nonlinear models, this conclusion is flawed (Ai and Norton 2003).

The reason is simple, though far from self-evident. By definition, the interaction effect describes how much a marginal change in `educ` modifies the impact of the treatment, here `nwifeinc`: $\partial\{\partial P/\partial(\texttt{nwifeinc})\}/\partial(\texttt{educ}) = \partial^2 P/\partial(\texttt{nwifeinc})\partial(\texttt{educ})$.[13] In the LPM, this cross-partial derivative is constant and equal to $\partial P/\partial(\texttt{nwifeinc} \times \texttt{educ})$, the coefficient on the multiplicative term. In nonlinear models such as `logit`, however, this equality does not hold in general. Furthermore, similar to the other expressions discussed so far, $\partial^2 P/\partial(\texttt{nwifeinc})\partial(\texttt{educ})$ also depends on the values of covariates. Indeed, even its sign need not be the same as the sign of the estimated coefficient on the multiplicative term. Ai and Norton (2003) derive formulas for $\partial^2 P/\partial(\texttt{nwifeinc})\partial(\texttt{educ})$ for logit (and probit) models.

In Stata, the correct interaction effect $(\partial^2 P/\partial(\texttt{nwifeinc})\partial(\texttt{educ}))$ after `logit` can be computed with the `inteff` command developed by Norton, Wang, and Ai (2004) and downloadable from the *Stata Journal*'s website (type `net search inteff` and select

---

13. If the researcher does not want to treat `nwifeinc` and (especially) `educ` as continuous, he or she should use discrete cross-differences rather than cross-derivatives.

st0063).[14] This command (which does not support factor-variable notation) produces a numerical estimate of the average interaction effect (`_logit_ie` in the output below), treating all the individuals `asobserved`:

```
. generate nwifeinc_educ = nwifeinc*educ
. quietly logit inlf nwifeinc educ nwifeinc_educ nokid, robust nolog
. inteff inlf nwifeinc educ nwifeinc_educ nokid
Logit with two continuous variables interacted
```

| Variable | Obs | Mean | Std. dev. | Min | Max |
|---|---|---|---|---|---|
| _logit_ie | 753 | .0004244 | .00062 | −.0009878 | .0010581 |
| _logit_se | 753 | .0008147 | .0002286 | .0000674 | .0016656 |
| _logit_z | 753 | .547078 | 1.268062 | −7.134073 | 5.891206 |

Note this estimate is an average: as already mentioned, $\partial^2 P/\partial(\texttt{nwifeinc})\partial(\texttt{educ})$ is different for different values of covariates. Here the mean (0.0004244) conceals a large amount of heterogeneity, with the magnitude of the interaction effect ranging between $-0.00099$ and $0.00106$ depending on the values of covariates. Somewhat unfortunately, the statistics reported as `_logit_se` and `_logit_z` (third column) refer to the average standard error and $z$ statistic, rather than the standard error and $z$ statistic of the average interaction effect reported just above.

To be sure, the `inteff` command was developed before factor-variable notation, and `margins` were introduced in Stata 11 (2009). Most of `inteff`'s output can now be reproduced more flexibly, albeit more laboriously, using `margins` (see below).

In practice, however, things are even easier. In many (probably, most) settings, the researcher is not primarily interested in how much the effect of a treatment changes in response to a marginal change in the modifier (that is, the interaction effect) but in what the treatment effect actually is for various degrees of exposure to the modifier. Thus, the derivative of (6) with respect to `nwifeinc` is often more useful for interpretation than $\partial^2 P/\partial(\texttt{nwifeinc})\partial(\texttt{educ})$:[15]

$$\frac{\partial P}{\partial(\texttt{nwifeinc})} = (\beta_1 + \beta_{12}\texttt{educ}_i)L'(u) = (\beta_1 + \beta_{12}\texttt{educ}_i)\left(\frac{e^{-u_i}}{(1+e^{-u_i})^2}\right) \qquad (7)$$

As is standard in applications involving linear models, $\partial P/\partial(\texttt{nwifeinc})$ can easily be plotted as a function of `educ`, the effect modifier, averaging over the remaining covariates. Of course, researchers may also hold some of or all the covariates fixed at representative values, or at the means, if they wish. This is what we do here. Figure 2[16] evaluates (7) for a) women with kids (`nokid` = 0) and wealthy husbands (income of 60,000 euros); and b) childless women with husbands on very low incomes (10,000 euros).

---

14. Cornelißen and Sonderhof's (2009) extension of `inteff` (called `inteff3`) is capable of handling triple interactions but only when all the regressors are dummy variables.
15. Note that if $\beta_{12}$ equals 0, (7) reduces to (2).
16. I created the graph with the command shown and then enhanced the $y$-axis and $x$-axis titles of the resulting graph using the Graph Editor.

```
. quietly logit inlf c.nwifeinc c.educ c.nwifeinc#c.educ i.nokid, robust nolog
. quietly margins, dydx(nwifeinc) at(educ=(5(1)17) nokid=(0) nwifeinc=(60))
. quietly marginsplot, recast(line) recastci(rarea) name(kidhigh) scheme(sj)
> title("nokid=0, nwifeinc=60")
. quietly margins, dydx(nwifeinc) at(educ=(5(1)17) nokid=(1) nwifeinc=(10))
. quietly marginsplot, recast(line) recastci(rarea) name(nokidlow) scheme(sj)
> title("nokid=1, nwifeinc=10")
. graph combine kidhigh nokidlow, ycommon xcommon
```
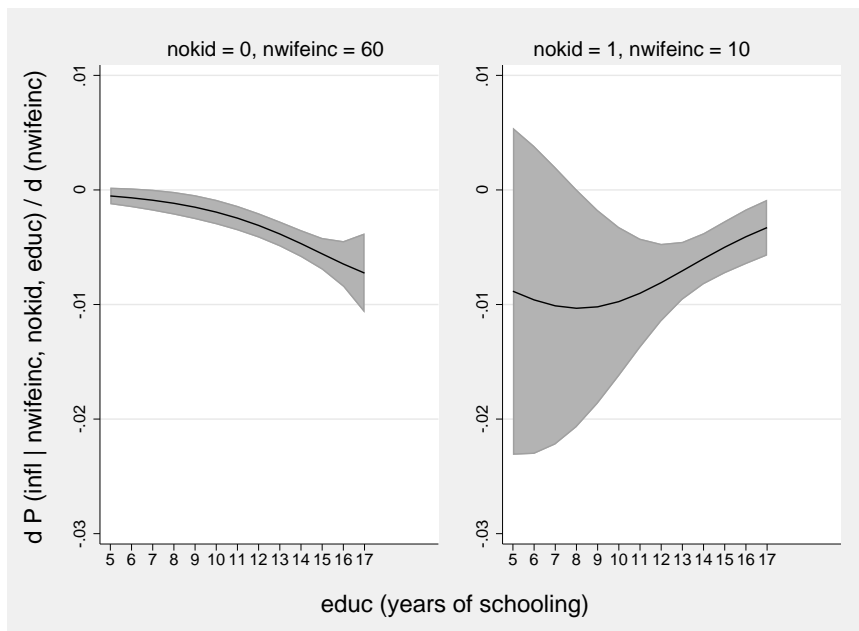


Figure 2. Average marginal effects of `nwifeinc` with 95% confidence interval

It is clear from figure 2 that whether $\partial P/\partial(\texttt{nwifeinc})$ is upward or downward sloping—that is, whether $\partial^2 P/\partial(\texttt{nwifeinc})\partial(\texttt{educ})$, the interaction effect, is positive or negative—depends on the values of covariates. For childless women with husbands on low incomes (right panel, figure 2), education mitigates the substitution effect triggered by the husband's income $[\partial^2 P/\partial(\texttt{nwifeinc})\partial(\texttt{educ}) > 0]$, at least past a given point ($\texttt{educ} \cong 8$). This finding is in line with theoretical expectations. For women with kids and wealthy husbands (left panel, figure 2), things work differently. An additional year of schooling amplifies the substitution effect of the husband's earnings—that is, it makes the estimated effect more negative $[\partial^2 P/\partial(\texttt{nwifeinc})\partial(\texttt{educ}) < 0]$. This finding runs counter to theoretical expectations.

In many (probably, most) applications, a diagram along the lines of figure 2 will provide more than enough insights to interpret the regression results, with no need to compute the interaction effect itself using `inteff`. Some researchers, however, may wish to take it a step further and report the magnitude and statistical significance of

$\partial^2 P/\partial(\texttt{nwifeinc})\partial(\texttt{educ})$, the interaction effect, at the representative values of covariates displayed in figure 2. This calculation cannot be accomplished by `inteff`. Rather, users should differentiate (7) with respect to `educ` (see Ai and Norton [2003, 124]) and code the resulting formula for the response manually:

$$\frac{\partial^2 P}{\partial(\texttt{nwifeinc})\partial(\texttt{educ})} = \beta_{12}L'(u) + (\beta_1 + \beta_{12}\texttt{educ}_i)(\beta_2 + \beta_{12}\texttt{nwifeinc}_i)L''(u)$$

$$= \beta_{12}L(u)\{1 - L(u)\} + (\beta_1 + \beta_{12}\texttt{educ}_i)(\beta_2 + \beta_{12}\texttt{nwifeinc}_i)$$
$$L(u)\{1 - L(u)\}\{1 - 2L(u)\} \tag{8}$$

For instance, for the left panel in figure 2, the interaction effect (the slope of the curve) can be obtained by first defining a macro for $L(u)$ and then coding (8) into `margin`'s `expression(`*pnl_exp*`)` option:

```
. global L "(1 / (1+exp(-_b[c.nwifeinc]*c.nwifeinc - _b[c.educ]*c.educ -
> _b[c.nwifeinc#c.educ]*c.nwifeinc#c.educ - _b[1.nokid]*nokid - _b[_cons])))"
. margins, expression( _b[c.nwifeinc#c.educ] * $L * (1 - $L ) + (_b[c.nwifeinc]
> + _b[c.nwifeinc#c.educ]*c.educ)*(_b[c.educ] + _b[c.nwifeinc#c.educ]*c.nwifeinc)
> * $L * (1 - $L ) * (1 - 2*$L ) )  at(nokid=0 nwifeinc=60 educ=(5(3)17))
warning: option expression() does not contain option predict() or xb().
Adjusted predictions                              Number of obs = 753
Model VCE: Robust

Expression: _b[c.nwifeinc#c.educ] * (1 / (1+exp(-_b[c.nwifeinc]*c.nwifeinc -
            _b[c.educ]*c.educ - _b[c.nwifeinc#c.educ]*c.nwifeinc#c.educ -
            _b[1.nokid]*nokid - _b[_cons]))) * (1 - (1 /
            (1+exp(-_b[c.nwifeinc]*c.nwifeinc - _b[c.educ]*c.educ -
            _b[c.nwifeinc#c.educ]*c.nwifeinc#c.educ - _b[1.nokid]*nokid -
            _b[_cons]))) ) + (_b[c.nwifeinc] +
            _b[c.nwifeinc#c.educ]*c.educ)*(_b[c.educ] +
            _b[c.nwifeinc#c.educ]*c.nwifeinc) * (1 /
            (1+exp(-_b[c.nwifeinc]*c.nwifeinc - _b[c.educ]*c.educ -
            _b[c.nwifeinc#c.educ]*c.nwifeinc#c.educ - _b[1.nokid]*nokid -
            _b[_cons]))) * (1 - (1 / (1+exp(-_b[c.nwifeinc]*c.nwifeinc -
            _b[c.educ]*c.educ - _b[c.nwifeinc#c.educ]*c.nwifeinc#c.educ -
            _b[1.nokid]*nokid - _b[_cons]))) ) * (1 - 2*(1 /
            (1+exp(-_b[c.nwifeinc]*c.nwifeinc - _b[c.educ]*c.educ -
            _b[c.nwifeinc#c.educ]*c.nwifeinc#c.educ - _b[1.nokid]*nokid -
            _b[_cons]))) )
1._at: nwifeinc = 60
       educ     =  5
       nokid    =  0
2._at: nwifeinc = 60
       educ     =  8
       nokid    =  0
3._at: nwifeinc = 60
       educ     = 11
       nokid    =  0
4._at: nwifeinc = 60
       educ     = 14
       nokid    =  0
5._at: nwifeinc = 60
       educ     = 17
       nokid    =  0
```

|  | Margin | Delta-method std. err. | z | P>\|z\| | [95% conf. interval] |  |
|---|---|---|---|---|---|---|
| _at |  |  |  |  |  |  |
| 1 | -.0001418 | .0000526 | -2.70 | 0.007 | -.0002449 | -.0000388 |
| 2 | -.0003024 | .0000567 | -5.34 | 0.000 | -.0004135 | -.0001913 |
| 3 | -.0005812 | .000123 | -4.73 | 0.000 | -.0008222 | -.0003402 |
| 4 | -.0008751 | .0002281 | -3.84 | 0.000 | -.0013222 | -.0004281 |
| 5 | -.0007097 | .001173 | -0.61 | 0.545 | -.0030086 | .0015893 |

By dropping the `at(nokid=0 nwifeinc=60 educ=(5(3)17))` option from the command above, users can also replicate the average interaction effect outputted by `inteff` (= 0.0004244). Usefully, they can also obtain an estimate of the standard error of this average effect (= 0.0007605), which is not returned by `inteff`.

In sum, `margins` (and in particular its `expression()` option) expands the range of possibilities made available by `inteff` (only at the cost of making the coding a bit more laborious). Yet I suggest that things are much easier in practice. Describing how the marginal effects of the variable of interest (`nwifeinc`) vary at different values of the effect modifier (`educ`), as in figure 2, is quicker and usually more informative than reporting interaction effects [that is, the values returned by (8)].

## 10 Conclusion

When fitting binary choice models, applied researchers should carefully consider the best way to interpret the coefficients, repressing a knee-jerk temptation to compute and report marginal effects indiscriminately after running a `logit` or `probit` regression. Plotting the predicted probabilities is often a visually effective solution. When some of the regressors are binary, predictive margins (that is, the predicted probabilities) and their contrasts (difference) are the way to go. Elasticities can be a tricky option because percentage changes in probability can be difficult to grasp intuitively and are sensitive to the choice of response formula. By contrast, risk ratios can often provide an effective metric to convey the results, although they are tedious to compute in Stata. In the presence of interaction terms, things are actually easier in practice than implied in the recent literature on this topic: researchers can simply use `margins` to plot the marginal effects of the treatment of interest as a function of the effect modifier. Usually, the resulting diagram conveys the nature of heterogeneous effects in a clear and intuitive fashion.

Logit (and probit) provides a more accurate description of binary choice processes than the LPM (Allison 2017). With a little effort, `logit` can be combined with `margins` or other related postestimation commands, or both, to obtain interpretable numerical estimates. Of course, when (and only when) a linear approximation can be made with little loss of information, as in the example discussed in this article, researchers can fall back on the LPM and estimate by ordinary least squares, simplifying the interpretation

of the coefficients. The tools reviewed in this article, however, suggest that interpreting logit models in Stata is not prohibitively hard and that logit (rather than the LPM) should be the default approach.

# 11   References

Ai, C., and E. C. Norton. 2003. Interaction terms in logit and probit models. *Economics Letters* 80: 123–129.

Allison, P. 2017. In defense of logit—Part 2. Statistical Horizons. https://statisticalhorizons.com/in-defense-of-logit-part-2.

Buis, M. 2010. Stata tip 87: Interpretation of interactions in nonlinear models. *Stata Journal* 10: 305–308. https://doi.org/10.1177/1536867X1001000211.

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.

Long, J. S., and J. Freese. 2014. *Regression Models for Categorical Dependent Variables Using Stata*. 3rd ed. College Station, TX: Stata Press.

Mroz, T. A. 1987. The sensitivity of an empirical model of married women's hours of work to economic and statistical assumptions. *Econometrica* 55: 765–799. https://doi.org/10.2307/1911029.

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.

Royston, P. 2013. marginscontplot: Plotting the marginal effects of continuous predictors. *Stata Journal* 13: 510–527. https://doi.org/10.1177/1536867X1301300305.

StataCorp. 2021. *Stata 17 Base Reference Manual*. College Station, TX.

Uberti, L. J. 2017. Stata tip 128: Marginal effects in log-transformed models: A trade application. *Stata Journal* 17: 774–778. https://doi.org/10.1177/1536867X1701700315.

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.

Wooldridge, J. M. 2020. *Introductory Econometrics: A Modern Approach*. 7th ed. Boston: Cengage Learning.

**About the author**

Luca J. Uberti is a researcher in the Department of Economics and Management, University of Luxembourg. His research is in institutional and comparative economics, growth and development, and applied statistical methods.