Prelude to Comparing Coefficients Between Nested Models

Richard Williams, University of Notre Dame, https://www3.nd.edu/~rwilliam/
Last revised February 28, 2017

Following are results of two bivariate and one multivariate logistic regressions using a data set I constructed:

```
. use https://www3.nd.edu/~rwilliam statafiles/standardized.dta, clear
. logit ybinary x1, nolog
. logit ybinary x2, nolog
. logit ybinary x1 x2, nolog
```

Usually, when we add variables to a model (at least in OLS regression), the effects of variables added earlier go down. However, in this case, we see that the coefficients for x1 and x2 increase (seemingly) dramatically when both variables are in the model, i.e. in the separate bivariate regressions the effects of x1 and x2 are .7388678 and .4886751, but in the multivariate regressions the effects are 1.78923 and 1.173144, more than twice as large as before. This leads to two questions:

1. If we saw something similar in an OLS regression, what would we suspect was going on? In other words, in an OLS regression, what can cause coefficients to get bigger rather than smaller as more variables are added?
2. In a logistic regression, why might such an interpretation be totally wrong?