

Regression Interpretation Guide

This document includes examples of how to interpret the output for bivariate and multivariate regressions with different types of variables. **NOTE:** None of these examples are currently being used in your hypotheses and you **may not use these combinations of variables in Homework 4.**

Bivariate Regression: Interval IV, Interval DV

Example: ANES 2016 Dataset

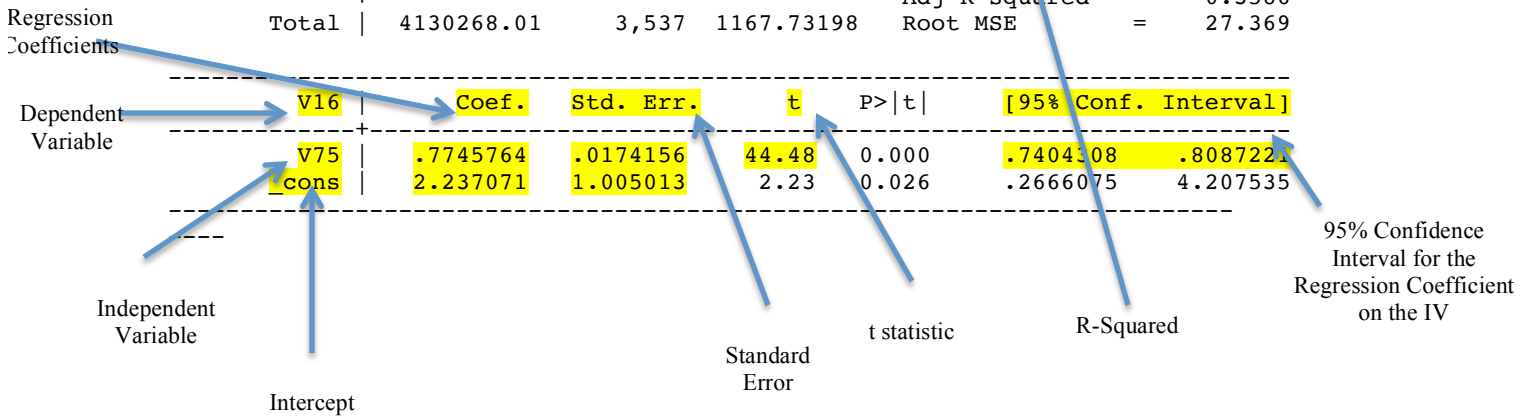
IV = Feelings toward liberals (V75)

DV = Feelings toward the 2016 Democratic presidential candidate (V16)

Stata Code: `regress v16 v75`

| Source | SS | df | MS | Number of obs | = | 3,538 |
|----------|------------|-------|------------|---------------|---|---------|
| Model | 1481676.09 | 1 | 1481676.09 | F(1, 3536) | = | 1978.11 |
| Residual | 2648591.91 | 3,536 | 749.036174 | Prob > F | = | 0.0000 |
| Total | 4130268.01 | 3,537 | 1167.73198 | R-squared | = | 0.3587 |
| | | | | Adj R-squared | = | 0.3586 |
| | | | | Root MSE | = | 27.369 |

| | Coef. | Std. Err. | t | P> t | [95% Conf. Interval] |
|-------|----------|-----------|-------|-------|----------------------|
| V16 | | | | | |
| V75 | .7745764 | .0174156 | 44.48 | 0.000 | .7404108 .8087221 |
| _cons | 2.237071 | 1.005013 | 2.23 | 0.026 | .2666075 4.207535 |



Interpretation:

- Regression Coefficient on the IV: .77
 - There is a positive association between feelings toward liberals and feelings toward the Democratic presidential candidate. Specifically, a one unit increase in favorability toward liberals is associated with a .77 unit increase in favorability toward the Democratic presidential candidate.
- Intercept: 2.24
 - When favorability toward liberals is 0, favorability toward the Democratic presidential candidate is 2.24.
- Standard Errors
 - The standard error of the regression coefficient on favorability toward liberals is .017.
 - The Standard Error of the intercept is 1.005.
- R-Squared: .3587
 - Variation in favorability toward liberals explain 35.87% of the variation in favorability toward the Democratic presidential candidate.

- Statistical Significance:
 - Because 0 is not inside the 95% confidence interval on the regression coefficient for V75 (.74, .80), we can reject the null hypothesis that the slope is 0. We conclude that there is a statistically significant association between feelings toward liberals (IV, V75) and feelings toward the 2016 Democratic presidential candidate (DV, V16)

Multivariate Regression: Interval IV, Interval DV, Ordinal Confound

Example: ANES 2016 Dataset

IV = Feelings toward liberals (V75)

DV = Feelings toward the 2016 Democratic presidential candidate (V16)

Control Variable: Opposition toward the 2010 health care law (V22A)

1=Favor, 2=Neither favor nor oppose, 3=Oppose

Let's say that I think that my attitudes toward the 2010 health care law (Obamacare / the Affordable Care Act) affect my feelings toward liberals *and* my feelings toward the Democratic presidential candidate. Specifically, I think that the more individuals oppose Obamacare, the colder they'll feel toward liberals (my IV) and the colder they'll feel toward the Democratic presidential candidate in 2016.

Stata Code: regress V16 V75 V22A

| Source | SS | df | MS | Number of obs | = | 3,536 |
|----------|------------|-------|------------|---------------|---|---------|
| Model | 2067091.81 | 2 | 1033545.91 | F(2, 3533) | = | 1771.92 |
| Residual | 2060768.12 | 3,533 | 583.291287 | Prob > F | = | 0.0000 |
| | | | | R-squared | = | 0.5008 |
| | | | | Adj R-squared | = | 0.5005 |
| Total | 4127859.93 | 3,535 | 1167.71144 | Root MSE | = | 24.151 |

| V16 | Coef. | Std. Err. | t | P> t | [95% Conf. Interval] |
|-------|-----------|-----------|--------|-------|----------------------|
| V75 | .5091085 | .017507 | 29.08 | 0.000 | .4747836 .5434334 |
| V22A | -16.39318 | .5175456 | -31.67 | 0.000 | -17.4079 -15.37846 |
| _cons | 49.45936 | 1.734354 | 28.52 | 0.000 | 46.05893 52.8598 |

Interpretation

- Regression Coefficient on the IV: .509
 - Holding all other variables constant, there is a positive association between feelings toward liberals and feelings toward the Democratic presidential candidate. Specifically, holding all else constant, a one unit increase in favorability toward liberals is associated with a .509 unit increase in favorability toward the Democratic presidential candidate.
- Regression Coefficient on the Control Variable (V22A Opposition to the health care law): -16.39
 - Holding all other variables constant, there is a negative association between opposition to the ACA and favorability toward the Democratic presidential

candidate. Specifically, a one unit increase in opposition toward the ACA is associated with a 16.39 unit *decrease* in favorability toward the Democratic presidential candidate.

- Intercept: 49.46
 - When favorability toward liberals and opposition to the health care law are both 0, favorability toward the Democratic presidential candidate is 49.46.
- Standard Errors
 - The standard error of the regression coefficient on favorability toward liberals is .0175.
 - The standard error of the regression coefficient on opposition toward Obamacare is 0.518.
 - The Standard Error of the intercept is 1.734.
- R-Squared: .5008
 - Variation in favorability toward liberals and variation in opposition to the ACA explain 50.08% of the variation in favorability toward the Democratic presidential candidate.
- Statistical Significance:
 - IV: Favorability toward Liberals: Because 0 is not inside the 95% confidence interval on the regression coefficient for favorability toward liberals (V75) (.47, .54), we can reject the null hypothesis that the slope is 0. We conclude that there is a statistically significant association between favorability toward liberals and favorability toward the 2016 Democratic presidential candidate, holding opposition to the ACA constant.
 - Control: Opposition to the ACA: Because the 0 is not inside the 95% confidence interval on the regression coefficient for opposition to the ACA (V22A) (-17.4, -15.4), we can reject the null hypothesis that the slope is 0. We can conclude with 95% confidence that there is a statistically significant association between opposition to the ACA (V22A) and favorability toward the 2016 Democratic presidential candidate (V16), holding favorability toward liberals (V75) constant.

Bivariate Regression: Nominal IV, Interval DV

NOTE: If you have a nominal IV with more than two categories, you will need to recode it to two categories of interest. For example, if your IV is which party the respondent thinks is better at handling the nation's economy (V34): 1=Democrats, 2=Republicans, 3=Not much difference, 4=Neither party, you might want to recode this as just comparing those who think Democrats handle the economy better to those who think Republicans handle the economy better, setting the other two categories to missing.

You can do this with the following code: (see the Section 3 Stata Code for more information)

```
recode V34 (4=.)
recode V34 (3=.)
```

Other common examples in which you might need to recode to two categories are: race (e.g. maybe use White/Non-White), marital status (e.g. maybe use married/not married), 2016 presidential vote (e.g. maybe use Clinton / Trump, setting Johnson, Stein, and Other to missing).

You should also recode your nominal variable so that one category takes the value of 1 and the other takes the value of 0. For example, let's set Democrats=1 and Republicans=0

```
recode V34 (1=1)
recode V34 (2=0)
```

Example: ANES 2016 Dataset

IV = 2012 presidential vote (V9)
1=Obama, 2=Romney

DV = Feelings toward the 2016 Democratic presidential candidate (V16)

Stata Code: I'm going to recode V9 so that 1=voted for Obama and 0=voted for Romney

```
recode V9 (1=1)
recode V9 (2=0)
regress V16 V9
```

| Source | SS | df | MS | Number of obs | = | 2,554 |
|----------|------------|-------|------------|------------------|---|---------------|
| Model | 1772793.74 | 1 | 1772793.74 | F(1, 2552) | = | 3158.86 |
| Residual | 1432217.26 | 2,552 | 561.213659 | Prob > F | = | 0.0000 |
| | | | | R-squared | = | 0.5531 |
| | | | | Adj R-squared | = | 0.5530 |
| Total | 3205011 | 2,553 | 1255.39013 | Root MSE | = | 23.69 |

| V16 | Coef. | Std. Err. | t | P> t | [95% Conf. Interval] |
|-------|----------------|-----------------|-------|-------|--------------------------|
| V9 | 53.3309 | .9488858 | 56.20 | 0.000 | 51.47024 55.19157 |
| _cons | 12.625 | .7208622 | 17.51 | 0.000 | 11.21147 14.03853 |

Interpretation:

- Regression Coefficient on the IV: -53.33
 - There is a positive association between voting for Obama in 2012 and feelings toward the 2016 Democratic presidential candidate. Specifically, voting for Obama in 2012 is associated with an *increase* in favorability toward the Democratic presidential candidate of 53.33 units.
- Intercept: 12.63
 - When 2012 vote choice is 0 (voted for Romney), favorability toward the Democratic presidential candidate is 12.63.
- Standard Errors
 - The standard error of the regression coefficient on 2012 vote choice is .949.
 - The Standard Error of the intercept is .721.
- R-Squared: .5531
 - Variation in 2012 vote choice explain 55.31% of the variation in favorability toward the Democratic presidential candidate.
- Statistical Significance:

- Because 0 is not inside the 95% confidence interval for the regression coefficient on V9 (2012 vote choice), we can reject the null hypothesis that the slope is 0. We conclude with 95% confidence that there is a statistically significant association between 2012 vote choice (V9) and favorability toward the 2016 Democratic presidential candidate (V16).

Multivariate Regression: Nominal IV, Interval DV, Interval Control

Example: ANES 2016 Dataset

IV = 2012 presidential vote (V9)

1=Obama, 2=Romney (1=Obama, 0=Romney after the recode above!)

DV = Feelings toward the 2016 Democratic presidential candidate (V16)

Control Variable = Feelings toward Big Business (V75B)

Stata Code: regress V16 V9 V75B

| Source | SS | df | MS | Number of obs | = | 2,528 |
|----------|------------|-------|------------|---------------|---|---------|
| Model | 1747507.6 | 2 | 873753.8 | F(2, 2525) | = | 1551.75 |
| Residual | 1421770.52 | 2,525 | 563.077435 | Prob > F | = | 0.0000 |
| | | | | R-squared | = | 0.5514 |
| | | | | Adj R-squared | = | 0.5510 |
| Total | 3169278.13 | 2,527 | 1254.16625 | Root MSE | = | 23.729 |

| V16 | Coef. | Std. Err. | t | P> t | [95% Conf. Interval] |
|-------|-----------|-----------|-------|-------|----------------------|
| V9 | 53.13119 | .9921955 | 53.55 | 0.000 | 51.18559 55.07678 |
| V75B | -.0078406 | .0220261 | -0.36 | 0.722 | -.0510317 .0353504 |
| _cons | 13.14425 | 1.449764 | 9.07 | 0.000 | 10.3014 15.98709 |

Interpretation

- Regression Coefficient on the IV: 53.13
 - Holding favorability toward big business constant, there is a positive association between voting for Obama in 2012 and favorability toward the 2016 Democratic presidential candidate. Specifically, holding all else constant, voting for Obama in 2012 is associated with a 53.13 unit increase in favorability toward the 2016 Democratic presidential candidate.
- Regression Coefficient on the Control Variable (V75B feelings toward big business): -0.0078
 - Holding 2012 vote choice constant, there is a negative association between favorability toward big business and favorability toward the 2016 Democratic presidential candidate. Specifically, holding vote choice constant, a one unit increase in favorability toward big business is associated with a .0078 unit decrease in favorability toward the 2016 Democratic presidential candidate.
- Intercept: 13.14
 - When favorability toward big business and 2012 vote choice are both 0, favorability toward the 2016 Democratic presidential candidate is 13.14.
- Standard Errors
 - The standard error of the regression coefficient on voting for Obama is .992.

- The standard error of the regression coefficient on favorability toward big business is 0.022.
- The Standard Error of the intercept is 1.45.
- R-Squared: .5514
 - Variation in voting for Obama in 2012 and favorability toward big business explain 55.14% of the variation in favorability toward the 2016 Democratic presidential candidate.
- Statistical Significance:
 - IV: 2012 Vote Choice: Because 0 is not inside the 95% confidence interval for the regression coefficient on 2012 vote choice (V9) (51.2, 55.1), we can reject the null hypothesis that the slope is 0. We can conclude with 95% confidence that there is a statistically significant association between voting for Obama in 2012 (V9) and favorability toward the 2016 Democratic presidential candidate (V16), holding favorability toward big business (V75B) constant.
 - Control: Favorability toward Big Business: Because **0 is inside** the 95% confidence interval of the regression coefficient on favorability toward big business (V75B) (-.051, .035), we **fail to reject** the null hypothesis that the slope is 0. We **do not have sufficient evidence** to conclude that there is a statistically significant association between favorability toward big business (V75B) and favorability toward the 2016 Democratic presidential candidate (V16), holding voting for Obama in 2012 (V9) constant.