# How to Interpret Regression Coefficients

# SS387 Public Finance

How one interprets the coefficients in regression models will be a function of how the dependent (y) and independent (x) variables are measured. In general, there are three main types of variables used in econometrics: continuous variables, the natural log of continuous variables, and dummy variables. In the examples below we will consider models with three independent variables:

x1i a continuous variable

ln(x2i) the natural log of a continuous variable

x3i a dummy variable that equals 1 (if yes) and 0 (if no)

Listed below are three models. In each case, the right hand side variables are the same, but the dependent variables differ. In each of these regressions, the dependent variable will be measured either as a continuous variable, the natural log or a dummy variable. Define the following dependent variables:

y1i a continuous variable

ln(y2i) the natural log of a continuous variable

y3i a dummy variable that equals 1 (if yes) and 0 (if no)

Below each model is text that describes how to interpret particular regression coefficients.

**Model 1: y1i =** β**0 + x1i**β**1 + ln(x2i)**β**2 + x3i**β**3 +** ε**i**

β1 =∂y1i/∂x1i = a one unit change in x1 generates a β1 unit change in y1i

β2 =∂y1i/∂ln(x2i) = a 100% change in x2 generates a β2 change in y1i

β3 = the movement of x3i from 0 to 1 produces a β3 unit change in y1i

**Model 2: ln(y2i) =** β**0 + x1i**β**1 + ln(x2i)**β**2 + x3i**β**3 +** ε**i**

β1 =∂ln(y2i)/∂x1i = a one unit change in x1 generates a 100\*β1 percent change in y2i β2 =∂ln(y1i)/∂ln(x2i) = a 100% change in x2 generates a 100\*β2 percent change in y2i β3 = the movement of x3i from 0 to 1 produced a 100\*β3 percent change in y2i

**Model 3: y3i =** β**0 + x1i**β**1 + ln(x2i)**β**2 + x3i**β**3 +** ε**i**

β1 =∂y3i/∂x1i = a one unit change in x1 generates a 100\*β1 percentage point change in the probability y3i occurs

β2 =∂y3i/∂ln(x2i) = a 100% change in x2 generates a 100\*β2 percentage point change in the probability y3i occurs

β3 = the movement of x3i from 0 to 1 produced a 100\*β3 percentage point change in the probability that y3i occurs

# An extended example:

Below are results from three regressions generated from one data set. The results parallel the three models outlined above. The data set contains responses from a sample of senior citizens (aged 65+) who are all on Medicare. The regressions have three different outcome measures (total expenditures on medical care (totalexp), the natural log of total medical expenditures (totalexp\_ln) and whether the person has high blood pressure (high\_bp). For each of these dependent variables, there are three potential independent variables, a continuous variable (age), the natural log of a continuous variable (ln of family income) and a dummy variable (obese) that equals 1 if a respondent is obese, =0 0 otherwise.

The sample description and the sample means are presented below.

. desc

Contains data from D:\bill\fall2008\econ30331\meps\_senior.dta obs: 2,970

vars: 6 20 Oct 2008 17:24

size: 77,220 (99.3% of memory free)

-------------------------------------------------------------------------------

storage display value

variable name type format label variable label

-------------------------------------------------------------------------------

|  |  |  |  |
| --- | --- | --- | --- |
| age | byte | %8.0g | age in years |
| totalexp | long | %12.0g | total expenditures on medical |

care, 2005

high\_bp byte %8.0g dummy variable, =1 if have high

|  |  |  |
| --- | --- | --- |
|  | blood | pressure, =0 otherwise |
| income\_ln | float | %9.0g | natural | log of family income |
| totalexp\_ln | float | %9.0g | natural | log of total medical |

expenditures

obese float %9.0g dummy variable, =1 if obese, =0 otherwise

-------------------------------------------------------------------------------

. sum

Variable | Obs Mean Std. Dev. Min Max

-------------+--------------------------------------------------------

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| age | | | 2970 | 74.07576 | 6.228823 | 65 | 85 |
| totalexp | | | 2970 | 8358.247 | 14109.34 | 1 | 235392 |
| high\_bp | | | 2970 | .6703704 | .4701578 | 0 | 1 |
| income\_ln | | | 2970 | 9.557707 | .3464276 | 9.220389 | 9.913537 |
| totalexp\_ln | | | 2970 | 8.045003 | 1.904871 | 0 | 12.36901 |

-------------+-------------------------------------------------------- obese | 2970 .2690236 .4435269 0 1

. reg totalexp age income\_ln obese

|  |  |  |  |
| --- | --- | --- | --- |
| Source | SS df MS | Number of obs | = | 2970 |
| -------------+------------------------------ | F( 3, 2966) | = | 8.20 |
| Model | 4.8607e+09 3 1.6202e+09 | Prob > F | = | 0.0000 |
| Residual | 5.8619e+11 2966 197636123 | R-squared | = | 0.0082 |
| -------------+------------------------------ | Adj R-squared | = | 0.0072 |
| Total | 5.9105e+11 2969 199073579 | Root MSE | = | 14058 |

------------------------------------------------------------------------------

totalexp | Coef. Std. Err. t P>|t| [95% Conf. Interval]

-------------+----------------------------------------------------------------

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| age | | | 202.1078 | 43.41592 | 4.66 | 0.000 | 116.9794 | 287.2362 |
| income\_ln | | | -260.2222 | 772.2026 | -0.34 | 0.736 | -1774.329 | 1253.885 |
| obese | | | 1251.303 | 588.4134 | 2.13 | 0.034 | 97.56308 | 2405.043 |
| \_cons | | | -4462.544 | 7241.433 | -0.62 | 0.538 | -18661.29 | 9736.197 |

------------------------------------------------------------------------------

Interpreting the coefficients:

age: a one year increase in age will increase annual medical spending by $202 income\_ln: a 100% increase in income will reduce medical spending by $260

male: Obese seniors spend $1251 more per year on medical care than the non-obese

. \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\* model 2 \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

. reg totalexp\_ln age income\_ln obese

|  |  |  |  |
| --- | --- | --- | --- |
| Source | SS df MS | Number of obs | = | 2970 |
| -------------+------------------------------ | F( 3, 2966) | = | 23.48 |
| Model | 249.870278 3 83.2900927 | Prob > F | = | 0.0000 |
| Residual | 10523.2502 2966 3.54796029 | R-squared | = | 0.0232 |
| -------------+------------------------------ | Adj R-squared | = | 0.0222 |
| Total | 10773.1205 2969 3.62853502 | Root MSE | = | 1.8836 |

------------------------------------------------------------------------------

totalexp\_ln | Coef. Std. Err. t P>|t| [95% Conf. Interval]

-------------+----------------------------------------------------------------

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| age | | | .0419183 | .0058171 | 7.21 | 0.000 | .0305124 | .0533243 |
| income\_ln | | | -.1696737 | .1034636 | -1.64 | 0.101 | -.3725414 | .0331939 |
| obese | | | .420106 | .0788386 | 5.33 | 0.000 | .2655222 | .5746899 |
| \_cons | | | 6.448543 | .9702434 | 6.65 | 0.000 | 4.546125 | 8.350962 |

------------------------------------------------------------------------------

Interpreting the coefficients:

age: a one year increase in age will increase medical spending by 4.2% income\_ln: a 100% increase in income will reduce medical spending by roughly 17%

male: Obese seniors have 42% higher medical care spending than non-obese seniors.

. reg high\_bp age income\_ln obese

|  |  |  |  |
| --- | --- | --- | --- |
| Source | SS df MS | Number of obs | = | 2970 |
| -------------+------------------------------ | F( 3, 2966) | = | 44.67 |
| Model | 28.371025 3 9.45700834 | Prob > F | = | 0.0000 |
| Residual | 627.921568 2966 .21170653 | R-squared | = | 0.0432 |
| -------------+------------------------------ | Adj R-squared | = | 0.0423 |
| Total | 656.292593 2969 .221048364 | Root MSE | = | .46012 |

------------------------------------------------------------------------------

high\_bp | Coef. Std. Err. t P>|t| [95% Conf. Interval]

-------------+----------------------------------------------------------------

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| age | | | .0053784 | .001421 | 3.79 | 0.000 | .0025922 | .0081646 |
| income\_ln | | | .0914678 | .0252735 | 3.62 | 0.000 | .0419125 | .1410232 |
| obese | | | .1987462 | .0192582 | 10.32 | 0.000 | .1609854 | .2365071 |
| \_cons | | | -.6557299 | .2370055 | -2.77 | 0.006 | -1.120442 | -.191018 |

------------------------------------------------------------------------------

Interpreting the coefficients:

age: a one year increase in age will increase the probability of having high blood pressure by 0.5 percentage points

income\_ln: a 100% increase in income will increase the probability of having high blood pressure by 9.1 percentage points

male: Obese seniors have 19.9 percentage point higher probability of being obese than non- obese seniors