OLS Regression Results

Dep. Variable: **Employed** R-squared: 0.967 Model: **OLS** Adj. R-squared: 0.965 Method: F-statistic: Least Squares 415.1 Date: Sat, 08 Feb 2014 **Prob (F-statistic):** 8.36e-12 Log-Likelihood: -14.904 Time: 01:28:29 No. Observations: 16 AIC: 33.81 **Df Residuals:** 14 BIC: 35.35

Df Model: 1

coef std err t P>|t| [95.0% Conf. Int.]

 Intercept
 51.8436
 0.681
 76.087
 0.000
 50.382
 53.305

 GNP
 0.0348
 0.002
 20.374
 0.000
 0.031
 0.038

 Omnibus:
 1.925
 Durbin-Watson:
 1.619

 Prob(Omnibus):
 0.382
 Jarque-Bera (JB):
 1.215

 Skew:
 0.664
 Prob(JB):
 0.545

 Kurtosis:
 2.759
 Cond. No.
 1.66e+03

This summary provides quite a lot of information about the fit. The parts of the table we think are the most important are bolded in the description below.

The left part of the first table provides basic information about the model fit:

Element Description

Dep. Variable Which variable is the response in the model

Model What model you are using in the fit

Method How the parameters of the model were calculated

No. Observations The number of observations (examples)

DF Residuals Degrees of freedom of the residuals. Number of observations – number of

parameters

DF Model Number of parameters in the model (not including the constant term if present)

The right part of the first table shows the goodness of fit

Element Description

R-squared The coefficient of determination. A statistical measure of how well the regression

line approximates the real data points

Adj. R-squared The above value adjusted based on the number of observations and the degrees-

of-freedom of the residuals

F-statistic A measure how significant the fit is. The mean squared error of the model divided

by the mean squared error of the residuals

Prob (F- The probability that you would get the above statistic, given the null hypothesis

statistic) that they are unrelated

Log-likelihood The log of the likelihood function.

AIC The Akaike Information Criterion. Adjusts the log-likelihood based on the number

of observations and the complexity of the model.

BIC The Bayesian Information Criterion. Similar to the AIC, but has a higher penalty

for models with more parameters.

The second table reports for each of the coefficients

	The name of the term in the model
coef	The estimated value of the coefficient
std err	The basic standard error of the estimate of the coefficient. More sophisticated errors are also available.
t	The t-statistic value. This is a measure of how statistically significant the coefficient is.
P > t	P-value that the null-hypothesis that the coefficient = 0 is true. If it is less than the confidence level, often 0.05, it indicates that there is a statistically significant relationship between the term and the response.

[95.0% Conf. The lower and upper values of the 95% confidence interval

Finally, there are several statistical tests to assess the distribution of the residuals

Element	Description
Skewness	A measure of the symmetry of the data about the mean. Normally-distributed errors should be symmetrically distributed about the mean (equal amounts above and below the line).
Kurtosis	A measure of the shape of the distribution. Compares the amount of data close to the mean with those far away from the mean (in the tails).
Omnibus	D'Angostino's test. It provides a combined statistical test for the presence of skewness and kurtosis.
Prob(Omnibus)	The above statistic turned into a probability
Jarque-Bera	A different test of the skewness and kurtosis
Prob (JB)	The above statistic turned into a probability
Durbin-Watson	A test for the presence of autocorrelation (that the errors are not independent.) Often important in time-series analysis
Cond. No	A test for multicollinearity (if in a fit with multiple parameters, the parameters are related with each other).

As a final note, if you don't want to include a constant term in your model, you can exclude it using the minus operator.