

Seminar 4

Using OLS in Stata:
 Functional Form, Data Scaling and Selection of Regressors

1 Running a Regression with Different Functional Forms Using OLS

1. Import the Stata data file "reg01" from the e-course platform.
2. Run the linear regression given by

$$\log(\text{income}) = \beta_0 + \beta_1 \text{educ} + \beta_2 \text{jobexp} + \beta_3 \text{race} + u$$

3. Interpret the OLS coefficients (intercept and slope coefficients). When doing it, make sure to discuss the sign of the coefficient, its magnitude and ceteris-paribus interpretations, as well as statistical significance of the coefficients. How did the results change compared to a case where `income` was in level form? For reference, below is the regression output where `income` was in level form:

Source	SS	df	MS	Number of obs	=	20
Model	1538.92019	3	512.973396	F(3, 16)	=	29.16
Residual	281.505287	16	17.5940804	Prob > F	=	0.0000
				R-squared	=	0.8454
				Adj R-squared	=	0.8164
Total	1820.42548	19	95.8118671	Root MSE	=	4.1945

income	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
educ	1.981124	.3231024	6.13	0.000	1.296178 2.66607
jobexp	.6419622	.1811106	3.54	0.003	.2580248 1.0259
race	.5707931	2.871949	0.20	0.845	-5.517466 6.659052
_cons	-7.863763	5.369166	-1.46	0.162	-19.24589 3.518362

4. Run the linear regression given by
- $$\text{income} = \beta_0 + \beta_1 \text{educ} + \beta_2 \text{jobexp} + \beta_3 \text{jobexp}^2 + \beta_4 \text{race} + u$$
5. Interpret the OLS coefficients (intercept and slope coefficients). When doing it, make sure to discuss the sign of the coefficient, its magnitude and ceteris-paribus interpretations, as well as statistical significance of the coefficients.

6. Run the linear regression given by

$$income = \beta_0 + \beta_1 educ + \beta_2 jobexp + \beta_3 jobexp * race + \beta_4 race + u$$

7. Interpret the OLS coefficients (intercept and slope coefficients). When doing it, make sure to discuss the sign of the coefficient, its magnitude and ceteris-paribus interpretations, as well as statistical significance of the coefficients.

2 Goodness-of-Fit and Selection of Regressors

1. We want to choose between the following models

$$income = \beta_0 + \beta_1 educ + \beta_2 jobexp + \beta_3 jobexp^2 + \beta_4 race + u$$

and

$$income = \beta_0 + \beta_1 educ + \beta_2 \log(jobexp) + \beta_3 race + u$$

Run the respective regressions in Stata and make the case for one of the models using adjusted R-squared.