

Understanding Earnings

Economics 140B Regression Interpretation Project

You will use the earnings dataset compiled from the Current Population Survey (CPS). There are three components. The first component is to prepare the data for analysis. The second component is to construct and estimate a model and interpret the magnitude and significance of the estimated effects.

Component 1

- 1) Examine the data on income. Carefully describe how you prepared the data for analysis. (For example, if there is a numeric code for missing observations, how did you treat these observations?)
- 2) Examine the data on education, gender, race, and marital status and describe any changes you made to prepare the data for analysis.

Component 2

- 1) Construct and estimate a model that explains income as a function only of education. Be sure to clearly motivate the functional form you choose.
- 2) Include gender and race in your model.
- 3) Include marital status in your model.

Component 3

- 1) Describe why one might be concerned that education is endogenous and what the implications are for the estimates of your model from Component 2.
- 2) Discuss the validity of any potential instruments in the data set.
- 3) If you feel you have a valid instrument, construct and estimate your model from component 2.3 (the model with gender, race and marital status included) using 2 stage least squares.