It is based on the estimation of a multivariate regression. We will use data from a subsample of the NLSY data on young women aged 14–26 years in 1968. Women were surveyed in each of the 21 years 1968–1988, except for the six years 1974, 1976, 1979, 1981, 1984, and 1986. Our subsample is of 4,711 women in years when employed, not enrolled in school and having completed their education, and with wages in excess of $1/hour but less than $700/hour. To import the data and obtain further information on the dataset use the following command:

webuse nlswork.dta

describe

Please complete the following:

1. Specify the model you will be estimating. You are to analyse ln wage in terms of completed years of schooling (grade), current age and age squared, current years worked (experience) and experience squared, current years of tenure on the current job and tenure squared, whether residing in an area not designated a standard metropolitan statistical area (SMSA), and whether residing in the South.
2. Describe each of the variables in your model. Present any other descriptive statistics that you feel are justified.
3. Estimate your model using the data provided. As the data is panel in nature estimate an OLS model, a least squares dummy variable model, a fixed effects within group model and a random effects model. Identify which model best fits the data.
4. Interpret the results of your final estimation.
5. Discuss any potential violations of the classical assumptions and discuss how you have corrected for these issues.

(i) exhibit a good understanding of how to justify your model specification (ii) exhibit a good understanding of how to interpret the estimated results, (iii) present the assignment in a professional manner and (iv) make sufficient reference to relevant literature in order to explain your model choices and results.