

Economics 312
Daily Problem #34

Spring 2014
April 11

Suppose that we have a balanced panel of data with N individuals observed over T periods. We want to allow each individual to have a different intercept, so we specify the model as

$$y_{it} = \beta_{1i} + \beta_2 x_{2it} + \beta_3 x_{3it} + e_{it}.$$

1. Explain intuitively why we get the same estimator if we (a) estimate the model using “differenced” data with no constant term (subtracting the first temporal observation from the later ones for each cross-sectional unit): $y_{it} - y_{i1} = \beta_2 (x_{2it} - x_{2i1}) + \beta_3 (x_{3it} - x_{3i1}) + (e_{it} - e_{i1})$, for $t = 2, 3, \dots, T$, or (b) include unit dummies for all individuals (leaving out the constant term):

$y_{it} = \sum_{j=1}^N \beta_{1j} D_{ji} + \beta_2 x_{2it} + \beta_3 x_{3it} + e_{it}$ for $i = 1, 2, \dots, N$. [Hint: Think about variation across individuals and variation over time, and which kind of variation is used to identify the parameters in each case.]

2. How many observations will be in the part (a) regression? How many coefficients are estimated? How many degrees of freedom are available?

3. How many observations will be in the part (b) regression? How many coefficients are estimated? How many degrees of freedom are available?