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 "de-meaned" data with no constant term: $y_{it} - \bar{y}_i = \beta_2 \left(x_{2it} - \bar{x}_{2i} \right) + \beta_3 \left(x_{3it} - \bar{x}_{3i} \right) + \left(e_{it} - \bar{e}_i \right)$, where the means are understood to be over time for each individual, or (b) include unit dummies for all individuals (leaving out the constant term): $y_{it} = \sum_{j=1}^{N} \beta_{1j} D_{ji} + \beta_{2i} x_{2it} + \beta_{3i} x_{3it} + e_{it} \, .$

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?