EM II/QSM II: Development
Problem Set 1 Solutions

Barcelona School of Economics, Spring 2026

Due date: Monday, 19 January, 23:59

Group members: AB, CD, EF, GH.

January 15, 2026

Submission Instructions

Please submit a document with your answers, including Stata or R output and programs, to our TA Janik Deutscher via the Google Classroom by Monday, 19 January, 23:59 at the very latest. You can work in groups of up to four.

1 Random Effects Model

Consider the model with a single regressor x_{it}:

$$y_{it} = \beta_0 + \beta_1 x_{it} + \alpha_i + u_{it}, \tag{1}$$

where α_i represents an unobserved effect fixed over time and u_{it} is a homoskedastic error term which is independent over time t and individuals i. There are N randomly sampled individuals, each observed for $T = 4$ time periods. Assume that $\mathbb{E}(u_{it}|X_i, \alpha_i) = 0$ for all i and that $\mathbb{E}(u_{it}u_{is}|X_i, \alpha_i) = 0$ for any t and $s : t \neq s$ where X_i represents the $T \times 2$ data matrix for individual i.

1. State under which assumptions you would estimate a random effects model in this context. Derive the random effects estimator and show that it is a consistent estimator of $\beta = [\beta_0, \beta_1]'$.

   > **Solution:** Here comes the solution.

2. Derive the (asymptotic) variance-covariance matrix of the random effects estimator.

   > **Solution:** Here comes the solution.

3. Explain how you would implement this estimator using the data in your sample.

> **Solution:** Here comes the solution.

2 Fixed Effects Variance Estimation

1. Show in detail why, in the context of the fixed effects model, we need to use the formula

$$\hat{\sigma}_u^2 = \frac{1}{N(T-1)} \sum_{i=1}^{N} \sum_{t=1}^{T} \tilde{u}_{it}^2 \tag{2}$$

to obtain a consistent estimate of $\hat{\sigma}_u^2$ (we are ignoring the degrees of freedom adjustment for the K regressors here which is asymptotically irrelevant). In particular, show that we need to divide by $N(T-1)$ rather than NT. (Note: for simplicity, it is fine here to work with $\tilde{u}_{it}$ rather than $\hat{\tilde{u}}_{it}$).

> **Solution:** Here comes the solution.

3 Fixed Effects versus First Difference Estimator

1. Consider the following estimation equation:

$$y_{it} = \alpha + x_{it}\beta + f_i + u_{it} \tag{3}$$

for $i = 1, \ldots, N$ and $t = 1, \ldots, T$.

where α is a constant, x_{it} is a single time-varying regressor and the idiosyncratic errors are serially uncorrelated and homoscedastic. Show that if $T = 2$, the fixed effects estimator and first difference estimator (which you obtain from transforming the model to $\Delta y_{it} = \Delta x_{it}\beta + \Delta u_{it}$ and then applying OLS to the transformed model) lead to identical estimates of both the coefficient and its variance.

> **Solution:** Here comes the solution.

4 Empirical Analysis: Children and Life Satisfaction

In this question, we will use panel data methods to understand how having children affects life satisfaction. Download the SOEP practise data set `soep_lebensz_en.dta`, which is available on the course website, and inspect its variables (using Stata or R).

1. Construct a binary variable `has_kids` that indicates if a person at time t has any children at all. Then, regress the (standardized) variable measuring current life satisfaction, `satisf_std`, on your constructed indicator. Include the individual's gender, education, categorical health and indicator variables for each year in the regression, and cluster your standard errors at the level of the individual.

2

First, estimate the effect of the children indicator on life satisfaction in a pooled OLS regression. Then, estimate the effect with a fixed effects regression. What does the difference of the estimated coefficients tell you about the unobserved effect f_i and, in particular, its covariance with `has_kids`?

> **Solution:** Here comes the solution.

2. Why has the coefficient on gender disappeared in the fixed effects regression? Run the same fixed effects regression as in the previous question but this time interact the gender indicator with the children indicator. Are women and men affected differentially? How do you interpret the magnitudes of the estimated coefficients?

> **Solution:** Here comes the solution.

3. Test the effect of having children on life satisfaction in a random effects model. Do the coefficients of the children indicator differ between the fixed and random effects model? What can you infer from this? Can you trust the assumptions of the RE model in this context? Why? Why not?

> **Solution:** Here comes the solution.

4. Perform a formal Hausman test to compare the fixed effects and the random effects model. Do you reject the null hypothesis? What does this result tell you? (Hint: check out the command `hausman` in Stata or `phtest` in R).

> **Solution:** Here comes the solution.