

Assignment 6: Panel Data

PSCI 8357, Spring 2016
March 17, 2016

This assignment must be turned in by the start of class on **Thursday, March 24**. You must follow the [instructions for submitting an assignment](#).

Main Task

This week's assignment is the third and last in a trio using the Neumayer (2005) data. You will once again use the Neumayer (2005) data; see the Assignment 4 handout for details. You will use the same model and composite hypothesis as in the previous assignments, again using the full data instead of a split sample.

1. **Before** running anything, think about your model. Do you expect your estimates to be affected by the clustered/panel structure of the data? Why or why not?
2. Test your composite hypothesis under each of the following estimators. For the sake of space, do *not* report the coefficients and standard errors of each model. Just report the code you used to fit each model, the code you used to test the hypothesis for each estimator, and the results of the corresponding hypothesis test. Bonus points for organizing the hypothesis test results into a table—programmatically, not manually—instead of giving us a sequence of nine different outputs.
 1. OLS with ordinary standard errors.
 2. OLS with heteroskedasticity-consistent standard errors.
 3. OLS with cluster-robust standard errors (clustered on country).
 4. Random effects by country with ordinary standard errors.
 5. Fixed effects by country with ordinary standard errors.
 6. Fixed effects by country with cluster-robust standard errors (clustered on country).
 7. Fixed effects by country and time with ordinary standard errors.
 8. Fixed effects by country and time with cluster-robust standard errors (clustered on country).

9. Fixed effects by country and time with cluster-robust standard errors (clustered on time).
3. If you were writing your results as a research paper and could only choose one of the nine specifications in the previous step to report, which would you choose? Why?
4. Choose any one of the terms in your original model, call it X_j with associated coefficient β_j . You will calculate the *cluster jackknife estimate* of the variance of $\hat{\beta}_j$, the OLS estimate of β_j (Efron and Stein 1981; Wu 1986; Lipsitz and Parzen 1996):
 1. For each cluster $i = 1, \dots, N$, calculate the estimated coefficients $\hat{\beta}^{(-i)}$ by running OLS on all the observations except those in cluster i . Save only the j 'th term of the estimated coefficients, $\hat{\beta}_j^{(-i)}$.
 2. Calculate the jackknife estimate of β_j via the formula

$$\bar{\beta}_j = \frac{1}{N} \sum_{i=1}^N \hat{\beta}_j^{(-i)}.$$

3. Estimate the variance of $\hat{\beta}_j$ via the formula

$$\hat{V}[\hat{\beta}_j] = \frac{N-p}{N} \sum_{i=1}^N (\hat{\beta}_j^{(-i)} - \bar{\beta}_j)^2,$$

where p is the number of terms in the model (including the intercept).

4. Estimate the standard error of β_j as $\sqrt{\hat{V}[\hat{\beta}_j]}$.

Compare the cluster jackknife estimate of the standard error to the original OLS estimate of the standard error and the cluster-robust estimate of the standard error. Which is it closer to?

If you are feeling especially ambitious, you may choose to obtain the cluster jackknife estimate of the variance of $\hat{\beta}$ as a whole (not just a single term). In this case the appropriate formulas would be

$$\bar{\beta} = \frac{1}{N} \sum_{i=1}^N \hat{\beta}^{(-i)}$$

and

$$\hat{V}[\hat{\beta}] = \frac{N-p}{N} \sum_{i=1}^N (\hat{\beta}^{(-i)} - \bar{\beta})(\hat{\beta}^{(-i)} - \bar{\beta})^{\top}.$$

References

- Efron, B., and C. Stein. 1981. "The Jackknife Estimate of Variance." *The Annals of Statistics* 9 (3): 586–96. <http://www.jstor.org/stable/2240822>.
- Lipsitz, Stuart R., and Michael Parzen. 1996. "A Jackknife Estimator of Variance for Cox Regression for Correlated Survival Data." *Biometrics* 52 (1): 291–98. <http://www.jstor.org/stable/2533164>.
- Neumayer, Eric. 2005. "Bogus Refugees? The Determinants of Asylum Migration to Western Europe." *International Studies Quarterly* 49 (3): 389–410. <http://dx.doi.org/10.1111/j.1468-2478.2005.00370.x>.
- Wu, C. F. J. 1986. "Jackknife, Bootstrap and Other Resampling Methods in Regression Analysis." *The Annals of Statistics* 14 (4): 1261–95. <http://www.jstor.org/stable/2241454>.