2023 Doctoral Winter School
13 – 17 February
14.00 – 18.00 (CET, Ljubljana)

Advanced Panel-Data Econometrics (ECTS: 4)

SPRUK ROK, University of Ljubljana, Slovenia

Aims of the course:

In recent years, panel datasets dominate the empirical research paradigm with observations at increasingly lower levels and research spanning many topics ranging from behavioural economics to political economy. The aim of the course is to provide a theoretical, methodological and applied overview of econometric models for panel data where observations are available on at least two dimensions.

The first part of the course relates to static and dynamic panel data with emphasis on the endogeneity of covariates through the correlation with individual heterogeneity and correlation with idiosyncratic shocks. Instrumental variables estimators will be discussed in-depth through the generalized method of moments. The second part of the course consists of problem evaluation techniques such as difference-in-differences, synthetic controls and event studies to evaluate the impact of policies on economic outcomes and behaviour.

At the end of the course, students will be able to critically evaluate the empirical literature based on panel data and policy evaluation techniques, estimate their own issue of interest and write a scholarly paper on the topic of interest.

Course syllabus:

1. **Introduction to panel-data econometrics:**
    a. Ordinary least squares
    b. Instrumental variables
    c. Heteroskedasticity
    d. Unit roots and stationarity
2. **Static panels:**
    a. Clustering the data structure
    b. Endogeneity, simultaneity, measurement error
    c. Dealing with unobserved heterogeneity
    d. Variance decompositions
    e. Correlated random effects and slopes
    f. Instrumental variables in panel data: testing and identification
    g. Unbalanced and weakly balanced panels in the presence of sample selection bias
3. **Dynamic panels:**
    a. Nickell's bias
    b. Overfitting versus weak instrumental variables

c. Arellano-Bond estimator and its extensions
d. First-differences and instrumental variables
e. Generalized method of moments estimator and instrumental variables
f. Principal components analysis and latent instruments
g. Multidimensional panels: estimating models with high-dimensional / multi-way fixed effects

4. **Heterogeneous and non-stationary panels:**
   a. Integration and cointegration
   b. First- and second-generation unit root tests
   c. Testing and correcting for cointegration through Pedroni and Westerlund tests
5. **Difference-in-differences policy evaluation**
   a. Standard difference-in-differences model with q-parallel trends
   b. Difference-in-differences estimator with staggered adoption
   c. Difference-in-differences estimator with multiple-period heterogeneous treatment
   d. Fuzzy difference-in-differences and two-way fixed-effects extension
   e. Event-style difference-in-differences
6. **Synthetic control analysis:**
   a. Preliminaries
   b. Model construction and estimation in the presence of parallel trend assumption violation
   c. Testing and specification issues
   d. Estimating and fitting the synthetic control models
   e. Sparsity and confidence intervals
   f. Temporal and spatial placebo analyses
   g. Prediction-based confidence intervals with synthetic controls
   h. Non-parametric synthetic control analysis with kernel-based regularization of bandwidth

Machine learning extensions of synthetic controls through ridge regression, elastic nets and LASSO

Tentative schedule:

In each session, students are expected to actively participate and contribute to the discussion in the class. A solid command of concepts and their empirical applications are expected at the end of each session.

Session 1: Introduction to panel-data econometrics
Readings:
Wooldridge (2010), chapters 1-10
Baltagi (2021), chapters 1-3
Arellano (2003), chapter 1
Alvarez, J., & Arellano, M. (2003). The time series and cross-section asymptotics of dynamic panel data estimators. Econometrica, 71(4), 1121-1159.

Session 2: Static panels:

Arellano (2003), chapters 2 through 4

Hadri, K. (2000). Testing for stationarity in heterogeneous panel data. The Econometrics Journal, 3(2), 148-161.

Baltagi (2021), chapters 4-3

Session 3: Dynamic, heterogeneous and non-stationary panels:

Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. The Review of Economic Studies, 58(2), 277-297.

Bun, M. J., & Carree, M. A. (2005). Bias-corrected estimation in dynamic panel data models. Journal of Business & Economic Statistics, 23(2), 200-210.

Chudik, A., & Pesaran, M. H. (2015). Common correlated effects estimation of heterogeneous dynamic panel data models with weakly exogenous regressors. Journal of Econometrics, 188(2), 393-420.

Hadri, K. (2000). Testing for stationarity in heterogeneous panel data. The Econometrics Journal, 3(2), 148-161.

Hayakawa, K. (2007). Small sample bias properties of the system GMM estimator in dynamic panel data models. Economics Letters, 95(1), 32-38.

Judson, R. A., & Owen, A. L. (1999). Estimating dynamic panel data models: a guide for macroeconomists. Economics Letters, 65(1), 9-15.

Moral-Benito, E., Allison, P., & Williams, R. (2019). Dynamic panel data modelling using maximum likelihood: an alternative to Arellano-Bond. Applied Economics, 51(20), 2221-2232.

Pedroni, P. (2004). Panel cointegration: asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis. Econometric Theory, 20(3), 597-625.

Phillips, P. C., & Sul, D. (2007). Bias in dynamic panel estimation with fixed effects, incidental trends and cross section dependence. Journal of Econometrics, 137(1), 162-188.

Wawro, G. (2002). Estimating dynamic panel data models in political science. Political Analysis, 10(1), 25-48.

Westerlund, J., & Edgerton, D. L. (2008). A simple test for cointegration in dependent panels with structural breaks. Oxford Bulletin of Economics and Statistics, 70(5), 665-704.

Session 4: Difference-in-differences

Abadie, A. (2005). Semiparametric difference-in-differences estimators. The Review of Economic Studies, 72(1), 1-19.

Athey, S., & Imbens, G. W. (2006). Identification and inference in nonlinear difference-in-differences models. Econometrica, 74(2), 431-497.

Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. Journal of Econometrics, 225(2), 200-230.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. The Review of Economics and Statistics, 90(3), 414-427.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2011). Robust inference with multiway clustering. Journal of Business & Economic Statistics, 29(2), 238-249.

Cameron, A. C., & Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. Journal of Human Resources, 50(2), 317-372.
Conley, T. G., & Taber, C. R. (2011). Inference with "difference in differences" with a small number of policy changes. The Review of Economics and Statistics, 93(1), 113-125.
Correia, S., Guimarães, P., & Zylkin, T. (2020). Fast Poisson estimation with high-dimensional fixed effects. The Stata Journal, 20(1), 95-115.
De Chaisemartin, C., & d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. American Economic Review, 110(9), 2964-96.
De Chaisemartin, C., & d'Haultfoeuille, X. (2018). Fuzzy differences-in-differences. The Review of Economic Studies, 85(2), 999-1028.
Donald, S. G., & Lang, K. (2007). Inference with difference-in-differences and other panel data. The Review of Economics and Statistics, 89(2), 221-233.
Guimaraes, P., & Portugal, P. (2010). A simple feasible procedure to fit models with high-dimensional fixed effects. The Stata Journal, 10(4), 628-649.
Sant'Anna, P. H., & Zhao, J. (2020). Doubly robust difference-in-differences estimators. Journal of Econometrics, 219(1), 101-122.

Session 5: Synthetic control method and extensions

Abadie, A., & L'Hour, J. (2021). A penalized synthetic control estimator for disaggregated data. Journal of the American Statistical Association, 116(536), 1817-1834.
Abadie, A. (2021). Using synthetic controls: Feasibility, data requirements, and methodological aspects. Journal of Economic Literature, 59(2), 391-425.
Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W., & Wager, S. (2021). Synthetic difference-in-differences. American Economic Review, 111(12), 4088-4118.
Athey, S., Bayati, M., Doudchenko, N., Imbens, G., & Khosravi, K. (2021). Matrix completion methods for causal panel data models. Journal of the American Statistical Association, 116(536), 1716-1730.
Ben-Michael, E., Feller, A., & Rothstein, J. (2021). The augmented synthetic control method. Journal of the American Statistical Association, 116(536), 1789-1803.
Botosaru, I., & Ferman, B. (2019). On the role of covariates in the synthetic control method. The Econometrics Journal, 22(2), 117-130.
Chernozhukov, V., Wüthrich, K., & Zhu, Y. (2021). An exact and robust conformal inference method for counterfactual and synthetic controls. Journal of the American Statistical Association, 116(536), 1849-1864.
Ferman, B., & Pinto, C. (2021). Synthetic controls with imperfect pretreatment fit. Quantitative Economics, 12(4), 1197-1221.
Gobillon, L., & Magnac, T. (2016). Regional policy evaluation: Interactive fixed effects and synthetic controls. Review of Economics and Statistics, 98(3), 535-551.
Xu, Y. (2017). Generalized synthetic control method: Causal inference with interactive fixed effects models. Political Analysis, 25(1), 57-76.

Teaching methods/Online tools and software:

Combined lectures and tutorials.

Course materials/List of readings:

Core literature:

- Wooldridge, J. (2010). *Econometric analysis of cross-section and panel data*. MIT Press.
- Baltagi, B. (2021). *Econometric analysis of panel data*. Springer.
- Arellano, M. (2003). *Panel data econometrics*. Oxford University Press.

E-reader distributed to the students at the initial session:

Abadie, A. (2005). Semiparametric difference-in-differences estimators. The Review of Economic Studies, 72(1), 1-19.
Abadie, A., & L'Hour, J. (2021). A penalized synthetic control estimator for disaggregated data. Journal of the American Statistical Association, 116(536), 1817-1834.
Abadie, A. (2021). Using synthetic controls: Feasibility, data requirements, and methodological aspects. Journal of Economic Literature, 59(2), 391-425.
Alvarez, J., & Arellano, M. (2003). The time series and cross-section asymptotics of dynamic panel data estimators. Econometrica, 71(4), 1121-1159.
Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W., & Wager, S. (2021). Synthetic difference-in-differences. American Economic Review, 111(12), 4088-4118.
Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. The Review of Economic Studies, 58(2), 277-297.
Athey, S., & Imbens, G. W. (2006). Identification and inference in nonlinear difference-in-differences models. Econometrica, 74(2), 431-497.
Athey, S., Bayati, M., Doudchenko, N., Imbens, G., & Khosravi, K. (2021). Matrix completion methods for causal panel data models. Journal of the American Statistical Association, 116(536), 1716-1730.
Ben-Michael, E., Feller, A., & Rothstein, J. (2021). The augmented synthetic control method. Journal of the American Statistical Association, 116(536), 1789-1803.
Botosaru, I., & Ferman, B. (2019). On the role of covariates in the synthetic control method. The Econometrics Journal, 22(2), 117-130.
Bun, M. J., & Carree, M. A. (2005). Bias-corrected estimation in dynamic panel data models. Journal of Business & Economic Statistics, 23(2), 200-210.
Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. Journal of Econometrics, 225(2), 200-230.
Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. The Review of Economics and Statistics, 90(3), 414-427.
Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2011). Robust inference with multiway clustering. Journal of Business & Economic Statistics, 29(2), 238-249.
Cameron, A. C., & Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. Journal of Human Resources, 50(2), 317-372.
Chernozhukov, V., Wüthrich, K., & Zhu, Y. (2021). An exact and robust conformal inference method for counterfactual and synthetic controls. Journal of the American Statistical Association, 116(536), 1849-1864.
Chudik, A., & Pesaran, M. H. (2015). Common correlated effects estimation of heterogeneous dynamic panel data models with weakly exogenous regressors. Journal of Econometrics, 188(2), 393-420.

Conley, T. G., & Taber, C. R. (2011). Inference with "difference in differences" with a small number of policy changes. The Review of Economics and Statistics, 93(1), 113-125.
Correia, S., Guimarães, P., & Zylkin, T. (2020). Fast Poisson estimation with high-dimensional fixed effects. The Stata Journal, 20(1), 95-115.
De Chaisemartin, C., & d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. American Economic Review, 110(9), 2964-96.
De Chaisemartin, C., & d'Haultfoeuille, X. (2018). Fuzzy differences-in-differences. The Review of Economic Studies, 85(2), 999-1028.
Donald, S. G., & Lang, K. (2007). Inference with difference-in-differences and other panel data. The Review of Economics and Statistics, 89(2), 221-233.
Ferman, B., & Pinto, C. (2021). Synthetic controls with imperfect pretreatment fit. Quantitative Economics, 12(4), 1197-1221.
Gobillon, L., & Magnac, T. (2016). Regional policy evaluation: Interactive fixed effects and synthetic controls. Review of Economics and Statistics, 98(3), 535-551.
Guimaraes, P., & Portugal, P. (2010). A simple feasible procedure to fit models with high-dimensional fixed effects. The Stata Journal, 10(4), 628-649.
Hadri, K. (2000). Testing for stationarity in heterogeneous panel data. The Econometrics Journal, 3(2), 148-161.
Hahn, J., & Shi, R. (2017). Synthetic control and inference. Econometrics, 5(4), 52.
Hayakawa, K. (2007). Small sample bias properties of the system GMM estimator in dynamic panel data models. Economics Letters, 95(1), 32-38.
Judson, R. A., & Owen, A. L. (1999). Estimating dynamic panel data models: a guide for macroeconomists. Economics Letters, 65(1), 9-15.
Moral-Benito, E., Allison, P., & Williams, R. (2019). Dynamic panel data modelling using maximum likelihood: an alternative to Arellano-Bond. Applied Economics, 51(20), 2221-2232.
Pedroni, P. (2004). Panel cointegration: asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis. Econometric Theory, 20(3), 597-625.
Phillips, P. C., & Sul, D. (2007). Bias in dynamic panel estimation with fixed effects, incidental trends and cross section dependence. Journal of Econometrics, 137(1), 162-188.
Sant'Anna, P. H., & Zhao, J. (2020). Doubly robust difference-in-differences estimators. Journal of Econometrics, 219(1), 101-122.
Wawro, G. (2002). Estimating dynamic panel data models in political science. Political Analysis, 10(1), 25-48.
Westerlund, J., & Edgerton, D. L. (2008). A simple test for cointegration in dependent panels with structural breaks. Oxford Bulletin of Economics and Statistics, 70(5), 665-704.
Xu, Y. (2017). Generalized synthetic control method: Causal inference with interactive fixed effects models. Political Analysis, 25(1), 57-76.

Course credit/Examination Methods:

Students needing a credit for the course will have to successfully pass the exam which will take place in the last day of the course.

Biographical note:

Rok Spruk is an assistant professor of economics at the School of Economics and Business, University of Ljubljana. He is a member of American Economic Association and Econometric Society. His research has been published in various high-impact journals including Review of International Organizations, Journal of Comparative Economics, Journal of Economic Surveys, Cliometrica, Empirical Economics and Journal of Regional Science among several others. His fields of research include methods of causal inference, empirical political economy, economic growth and development and economic history. His research on the institutional transformation and origins of world income distribution has been listed by Edward Elgar's encyclopedia on critical writings as one of the most influential works in the field of institutions and economic growth. He is the recipient of the Young Scholars prize for exceptional achievements by University of Ljubljana.