Panel Data Analysis

Instructor: Javier Osorio

josorio1@email.arizona.edu
www.javierosorio.net
Office: 326 Social Sciences Building
Office Hours: by appointment.
Phone: 520-621-4607

Version: 8/29/2021

Course Description:

Political Scientists across sub-fields require analyzing and explaining the behavior of several units of analysis over time. These could be individuals, countries, municipalities, institutions, groups, firms, etc. that display different behavioral trends over time and across space. This course offers graduate students the opportunity to develop methodological skills for understanding and performing quantitative analyses of time series cross-sectional (TSCS) data, also known as panel data. The topics covered in this course include the specification, estimation, evaluation, and inference of fixed and random effects models for continuous, categorical, and count data in time series cross-sectional studies. The basic model specification will be extended to more advanced methods such as dynamic models, instrumental variables, generalized method of moments, and spatial econometric models for panel data.

Course Objectives:

The lectures, in-class exercises, and assignments of this course are designed to help graduate students acquire methodological foundations and develop technical skills for conducting econometric analyses for panel data. After successfully completing this course, students will be able to:

1. Understand the methodological foundations of panel data econometrics.
2. Specify and implement quantitative analyses for panel data analysis.
3. Interpret and evaluate panel data model specifications in applied research.

Learning Outcomes:

The lectures, in-class exercises, and assignments are designed to help graduate students acquire methodological foundations and develop technical skills for conducting TSCS analyses. After successfully completing this course, students will be able to:

1. Articulate the methodological foundations of panel data econometrics.
2. Specify and independently implement quantitative analyses for TSCS data.
3. Interpret and evaluate panel data model specifications in applied research.

Required Texts:

The course will rely on selected sections of the following books. They are available online at the UA Library.

- Cheng Hsiao. *Analysis of Panel Data.* Cambridge University Press, New York, second edition, 2003.
- Badi H. Baltagi. *Econometric Analysis of Panel Data.* Wiley, West Sussex, fifth edition, 2013.
- Edward W. Frees. *Longitudinal and Panel Data.* Cambridge University Press, Cambridge, 2004.
- Yves Croissant and Giovanni Millo. *Panel Data Econometrics with R.* Wiley, Hoboken, NJ, fifth edition, 2019.

Required Software:
- In class sessions, and problem sets will rely on R and RStudio, which are available for free at `https://www.rstudio.com/` and `https://www.rstudio.com/`.
- Students are required to bring their own laptop computer to class.
- Students are expected to have basic familiarity with quantitative analysis using R.
- For the first class session, students are required to have an updated version of R and RStudio in their personal computers.

Course Requirements:
1. **Class attendance and participation**

   - Students are expected to attend to all class sessions and arrive on time.
   - Students considering being absent for religious observances, please send me an email at the beginning of the semester.
   - In the case of an emergency, students are expected to notify the professor before class.
   - Active participation in class is a strong requirement.

2. **Problem sets**

   - Students will receive a number of problem sets throughout the semester. These problem sets will mainly ask you to apply panel data analysis procedures on your own data.
   - My recommendation is that you use these problem sets to start analyzing the data that you will use for your final paper.
   - Students are required to work individually on each problem set.
   - Problem sets are due via D2L on the specified date.
   - Students are required to submit their problem sets using R Markdown and make sure the results can be replicated seamlessly.

3. **Replication sessions**

   - Students will pick a published academic article on a topic of their preference. The article should rely on a panel data analysis as the core of its empirical evaluation. Hopefully, this paper will be on a similar topic as the student's final paper.
   - To facilitate the replication process, try to find an article that used R to conduct the data analysis. In that way you will not struggle with translating Stata or SPSS code into R.
   - Students will inform the instructor the selected article by the date indicated in the Course Content section.
   - Each student will present to the class an analytic assessment of the results published in the paper. Such assessment will include:
     - Brief explanation of the main argument.
     - Brief explanation of the research design.
     - Replication of the original results.

– Identify the strengths of the analysis.
– Conduct a thorough methodological assessment to identify weaknesses.
– Implement a methodological improvement of the original results by modifying the model specifications or considering alternative variables or additional data.
– Discuss the improved results.

4. **Research paper**

- Students should write a research paper on the topic of their interest. The key requirement is for the paper to rely on panel data analysis.
- Students who would like to use survey data, must find a panel survey with at least 7 waves. There is no golden rule about the sufficient time frame for panel data, but less than 7 waves make it hard to model time.
- Towards the end of the semester, we will have the opportunity to workshop your research papers. Each student will make a brief presentation to talk about the theoretical, substantive, and empirical aspects of the research project. The rest of the class will engage to provide constructive and actionable feedback.
- The basic components of a solid research paper include the following: a theoretical or empirical puzzle; an explanatory research question; a clear argument; a brief and sharp literature review; advance a theoretical contribution; conduct an empirical assessment of the theoretical expectations (using panel data analysis, in this case); and a discussion of the implications of the results.
- The best way to approach this requirement is for students to use their research paper as the groundwork of a publishable article. I will be happy to provide continuous advice beyond this course for those papers advancing toward publication.
- The final paper must be no more than 9,000 words including references, tables, and footnotes. Please, use regular margins, Times New Roman font, 12 pts. and double space.
- For professionalization purposes, students are encouraged to write their papers in LaTeX.
- The final paper will be due via D2L.
- Students are responsible for submitting the right version of the paper on time.

Grade Distribution and Scale:

The following table presents the grade distribution for this course :

Grade Distribution	
Assignments	Percentage
Problem sets	30%
Replication session	30%
Final paper	40%

The following table presents the grading scale for a corresponding letter grade:

Grading Scale	
Letter grade	Scale
A	90-100%
B	80-89%
C	70-79%
D	60-69%
E	59% or less

Course Policies:

1. **Difficult Times**

   - From all the things that you can be, please choose to be kind. Be kind to yourself, be kind to all the people that you know. Also, please be kind to all the people that you do not know. We all are struggling.

   - If you are experiencing personal difficulties (of any kind) that prevent your from adequately performing in the class, please send me an email. Timely communication is crucial for properly addressing specific challenges. So, please do not wait until there is a drop in your grade or at the end of the semester. I will try to work with you to find the best possible solution so this class is not an additional burden on you.

   - If you are experiencing unexpected barriers to your success, please note the Dean of Students Office is a central support resource for all students and may be helpful. The Dean of Students Office (`deanofstudents.arizona.edu`) can be reached by phone at 520-621-2057 or by email `DOS-deanofstudents@email.arizona.edu`.

   - If you are facing physical or mental health challenges this semester, please note that Campus Health provides quality medical and mental health care. For medical appointments, call (520-621-9202. For After Hours care, call (520) 570-7898. For the Counseling & Psych Services (CAPS) 24/7 hotline, call (520) 621-3334.

2. **COVID Related Issues**

   - Please keep an eye on the University of Arizona COVID-related policies. As the pandemic evolves, the University will continue to adapt its policies, which is also likely to have consequences for this class. You can find out more at `https://covid19.arizona.edu/`.

   - If you are feeling COVID-related symptoms (even if they are mild) or if you have been in contact with someone who tested positive, **please do not come to class and get tested**. We all should be responsible for our own health and other people's health. So, when in doubt, just stay home and get tested before coming to class.

   - If you test positive to COVID, please follow the UA protocol:
     - `https://covid19.arizona.edu/positive-case-protocol-students`.

3. **Electronic etiquette**

   - Please mute your cell phones in class.

   - Refraining from using distracting electronics or apps during class will help us to engage more effectively in a stimulating intellectual conversation without interruptions.

   - Students are expected to behave maturely. Please demonstrate respect to yourself, your peers, and instructor by giving your full attention and participating in class.

   - Students are required to use their University of Arizona email for all university-related communications.

   - When sending an email, please be respectful, clear, and concise. Long substantive questions are more suitable for class discussions or office hours than email exchanges.

4. **Academic Policies**

- Please consult the UA Academic Policies Website:
  `https://academicaffairs.arizona.edu/syllabus-policies`
- This site includes **important information** regarding:
  - Absence and class participation.
  - Threatening behavior.
  - Accessibility and Accommodations (Disability Resource Center).
  - Code of Academic Integrity.
  - Non-Discrimination and Anti-Harassment Policy.
  - Safety on Campus and in the Classroom

5. **Additional Resources for Students**

- SGPP recently put together a comprehensive list of resources for students. You can find it at: `https://sgpp.arizona.edu/student-resources`
- This site includes information regarding:
  - Important dates and deadlines.
  - General resources.
  - COVID-19 resources.
  - Undergraduate student resources.
  - Graduate student resources.
  - Financial resources.
  - Health, wellness, and safety.
  - Student organizations.
  - Cultural and specialized student resource centers.

6. **Confidentiality of Student Records**:

- Confidentiality policy of the University of Arizona:
  `https://registrar.arizona.edu/privacy-ferpa/ferpa-compliance`

7. **Diversity and Inclusion:**

- The School of Government and Public Policy (SGPP) recognizes the richness of diversity and inclusion as critical components of intellectual and civic excellence. We are committed to fostering environments in which our students, staff, faculty, community partners, and visitors can participate fully, regardless of race, ethnicity, gender identity, sexual orientation, age, socio-economic status, citizenship status, size, ability, language, religion, or any other characteristic.

  We recognize that the disciplines of political science and public affairs do not yet reflect the complexity of our increasingly multicultural society, and that the substance of these disciplines is subject to polarized worldviews. Accordingly, we promote open, critical, and respectful dialogue and practices. To ensure that all members of the SGPP community thrive, we condemn expressions of discrimination and hate, and support pedagogies to explain and challenge all manifestations of unconscious bias in the classroom, on campus, and in the community.

The School has established the Committee for Diversity, Equity, and Inclusion to identify and coordinate activities to support our efforts to foster diverse and inclusive intellectual environments.

- In line with the University of Arizona's vision of diversity, this course will provide an environment of recognition, acceptance, and interaction of all the aspects that enrich diversity in our pursue of excellence.

- Participants in this course are strongly encouraged to move beyond recognizing diversity and inclusion as ideal or desirable goals, and actively engage in embracing, practicing, and promoting diversity and inclusion in the classroom and in their daily activities and relationships.

- At the beginning of the course, students will be asked about their pronouns.

8. **Care-Giving Policy:**

- Babies and children are welcome in class. In case of occasional minor illnesses or unforeseen disruptions in childcare, you can bring your baby or child to class.

- Mothers who are breastfeeding an infant, or expressing milk, may do so in class without the permission of the instructor. I ask everyone in class to be respectful and contributing to a family-friendly environment. If you prefer to breastfeed or breast-pump outside of class, you may take time out of class.

- In all cases where babies and children come to class, I ask that you sit close to the door so that if your little one needs special attention, you may step outside until their need has been met. I also ask non-parenting students to reserve seats near the door for parenting classmates.

- College is hard and even more so if you are a parent or care-giver of another person. Sleep deprivation, tiredness, stress, and unforeseen circumstances are real obstacles for anyone. Feel free to share with me your student-parent-care-giver status so I can offer special accommodation if needs arise. Although I have high expectations for all students in class, I am always happy to offer support for those who need it.

9. **Accessibility and Accommodations:**

- At the University of Arizona, we strive to make learning experiences as accessible as possible. If you anticipate or experience barriers based on disability or pregnancy, please contact the Disability Resource Center (520-621-3268, `https://drc.arizona.edu`) to establish reasonable accommodations.

10. **Safety on Campus and in the Classroom**

- For a list of emergency procedures for all types on incidents, please visit the website of the Critical Incident Response Team (CIRT):

  `https://cirt.arizona.edu/case-emergency/overview`

11. **Subject to Change:**

- Information contained in the course syllabus, other than the grade distribution and UA policies, may be subject to change with advance notice as deemed appropriate by the instructor.

Course Schedule:

The course will follow this schedule:

Month	Day	Session	Topic
August	21	1	Introduction
August	28	2	Panel Data Fundamentals
September	4		No Class - Labor Day
September	11	3	Panel Balance and Missing Data
September	18	4	Review of Linear Regression
September	15	5	Fixed Effects Models
October	2	6	Random Effects Models
October	9	7	Tests for Fixed or Random Effects Models
October	16	8	Review of TSCS topics I
October	23	9	Dichotomous data
October	30	10	Count data
November	6	11	Dynamic Models
November	13	12	Spatial Panel Data
November	20		Final Paper Workshop
November	27	13	Review of TSCS topics II
December	4	14	Review of TSCS topics III
December	11		Final paper due

Course Content:

August 21
Session 1 - Introduction

- Introduction
- R Markdown
- High-Performance Computing (HPC)

August 28
Session 2 - Panel Data Fundamentals

- Characteristics of panel data.
- Panel data visualization
  Assigned readings:

  - Cheng Hsiao. *Analysis of Panel Data.* Cambridge University Press, New York, second edition, 2003, Chapter 1, pp. 1-13.
  - Edward W. Frees. *Longitudinal and Panel Data.* Cambridge University Press, Cambridge, 2004, Chapter 1, pp. 1-15.
  - Michenug Liu and Yiqing Xu. panelView. 2019. Available at: `https://cran.r-project.org/web/packages/panelView/panelView.pdf`.

September 4
No Class - Labor Day

- No class

September 11
Session 3 - Panel Balance and Missing Data

- Panel Balance.
- Data Imputation.
  Assigned readings:

  - Stefan Dahlberg, Sören Holmberg, Bo Rothstein, Natalia Alvarado Pachon, and Richard Svensson. *The Quality of Government Basic Dataset, version Jan19.* University of Gothenburg. The Quality of Government Institute, 2019. Download the codebook of the small database and get familiar with it.
  - Gary King, James Honaker, Anne Joseph, and Kenneth Scheve. Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation. *American Political Science Review*, 95(1):49–69, 2001.
  - James Honaker, Gary King, and Matthew Blackwell. Amelia II: A Program for Missing Data. 2018. Available at: `https://gking.harvard.edu/amelia`.

September 18
Session 4 - Review of Linear Regression

- **Assignment 1 is due**.

- OLS review

  Assigned readings:

    - Alexander Demos and Carlos Salas. Chapter 12: Regression: Basics, Assumptions, and Diagnostics, 2019.
    - Bengi Koseglu. Understanding Ordinary Least Square in Matrix Form with R, 2018.
    - Edward W. Frees. *Longitudinal and Panel Data*. Cambridge University Press, Cambridge, 2004, Chapter 2, 24-30. We will be using the Medicare data mentioned in those pages.

September 15
Session 5 - Fixed Effects Models

- Fixed Effects models.

  Assigned readings:

    - Cheng Hsiao. *Analysis of Panel Data*. Cambridge University Press, New York, second edition, 2003, Chapter 3, pp. 27-31, 53-54.
    - Edward W. Frees. *Longitudinal and Panel Data*. Cambridge University Press, Cambridge, 2004, Chapter 2, pp. 18-23.

October 2
Session 6 - Random Effects Models

- Random effects models.

  Assigned readings:

    - Edward W. Frees. *Longitudinal and Panel Data*. Cambridge University Press, Cambridge, 2004, Chapter 3, pp. 72-81.
    - Cheng Hsiao. *Analysis of Panel Data*. Cambridge University Press, New York, second edition, 2003, Chapter 3, pp. 34-38.

October 9
Session 7 - Tests for Fixed or Random Effects Models

- Model tests.

  Assigned readings:

    - Badi H. Baltagi. *Econometric Analysis of Panel Data*. Wiley, West Sussex, fifth edition, 2013, Chapter 4, pp. 76-82.
    - Yves Croissant and Giovanni Millo. *Panel Data Econometrics with R*. Wiley, Hoboken, NJ, fifth edition, 2019, Chapter 4, pp. 83-95.

October 16
Session 8 - Review of TSCS topics I

  Assigned readings:

    - Gary King, Michael Tomz, and Jason Wittenberg. Making the Most of Statistical Analyses: Improving Interpretation and Presentation. *American Journal of Political Science*, 44(2):347, 2000.
    - Nathaniel Beck and Jonathan N Katz. What to do (and not to do) with Time-Series Cross-Section Data. *American Political Science Review*, 89(3):634–647, 1995.

- Nathaniel Beck. From statistical nuisances to serious modeling: Changing how we think about the analysis of time-series-cross-section data. *Political Analysis*, 15(2):97–100, 2007.
- Kevin Arceneaux and David W. Nickerson. Modeling certainty with clustered data: A comparison of methods. *Political Analysis*, 17(2):177–190, 2009.

October 23
Session 9 - Dichotomous data

- **Assignment 2 is due**.

- Fixed and Random models with dichotomous dependent variables.
  Assigned readings:

  - Cheng Hsiao. *Analysis of Panel Data*. Cambridge University Press, New York, second edition, 2003, Chapter 7, pp. 188-202.
  - Edward W. Frees. *Longitudinal and Panel Data*. Cambridge University Press, Cambridge, 2004, Chapter 9, pp. 319-339.
  - Yves Croissant and Giovanni Millo. *Panel Data Econometrics with R*. Wiley, Hoboken, NJ, fifth edition, 2019, Chapter 11, pp. 211-217.
  - Nathaniel Beck, Jonathan N. Katz, and Richard Tucker. Taking Time Seriously : Time-Series-Cross-Section Analysis with a Binary Dependent Variable. *American Journal of Political Science*, 42(4):1260–1288, 1998

October 30
Session 10 - Count data

- Poisson Model.
- Negative Binomial Model.
  Assigned readings:

  - Yves Croissant and Giovanni Millo. *Panel Data Econometrics with R*. Wiley, Hoboken, NJ, fifth edition, 2019, Chapter 9, pp. 236-240.

November 6
Session 11 - Dynamic Models

- Dynamic models with lagged dependent variables.
- Generalized Method of Moments.
- Instrumental Variables Estimator.
  Assigned readings:

  - Yves Croissant and Giovanni Millo. *Panel Data Econometrics with R*. Wiley, Hoboken, NJ, fifth edition, 2019, Chapter 6, pp. 139-145 and Chapter 7, pp. 161-165, 168-181.
  - Luke Keele and Nathan J. Kelly. Dynamic models for dynamic theories: The ins and outs of lagged dependent variables. *Political Analysis*, 14(2):186–205, 2006.
  - Nathaniel Beck and Jonathan N. Katz. Modeling Dynamics in Time-SeriesCross-Section Political Economy Data. *Annual Review of Political Science*, 14(1):331–352, 2011.

November 13
Session 12 - Spatial Panel Data

- Spatial Lag Regression.
- Spatial Error Component.
  Assigned readings:

    – Yves Croissant and Giovanni Millo. *Panel Data Econometrics with R*. Wiley, Hoboken, NJ, fifth edition, 2019, Chapter 10.
    – Giovanni Millo and Gianfranco Piras. splm: Spatial Panel Data Models in R. *Journal of Statistical Software*, 47(1):1–38, 2012.
    – Robert J. Franzese and Jude C. Hays. Spatial econometric models of cross-sectional interdependence in political science panel and time-series-cross-section data. *Political Analysis*, 15(2):140–164, 2007.
    – Morgan Kelly. The Standard Errors of Persistence The Myth of Europe's Little Ice Age View project The Standard Errors of Persistence. (June), 2019.

November 20
Final Paper Workshop

- Students present their research paper.

November 27
Session 13 - Review of TSCS topics II

- We will assess the content of this session based on the progress of the class.
- In consequence, the activities and readings may change depending on the needs of students.
  Assigned readings:

    – Robert Reed. On the Practice of Lagging Variables to Avoid Simultaneity. *The Oxford Bulletin of Economics and Statistics*, 77(6):897–905, 2015.
    – Marc F. Bellemare, Takaaki Masaki, and Thomas B. Pepinsky. Lagged Explanatory Variables and the Estimation of Causal Effect. *The Journal of Politics*, 79(3):949–963, 2017.
    – Yu Wang and Marc F. Bellemare. Lagged Variables as Instruments. 2019, available at `http://marcfbellemare.com/wordpress/wp-content/uploads/2019/05/WangBellemareLaggedIVsMay2019.pdf`.

December 4
Session 14 - Review of TSCS topics III

- We will assess the content of this session based on the progress of the class.
- In consequence, the activities and readings may change depending on the needs of students.
  Assigned readings:

    – Aviv Nevo and Adam M. Rosen. Identification with imperfect instruments. *Review of Economics and Statistics*, 94(3):659–671, 2012.
    – Matthew Blackwell and Adam N. Glynn. How to Make Causal Inferences with Time-Series Cross-Sectional Data under Selection on Observables. *American Political Science Review*, 112(4):1036–1049, 2018.
    – Kosuke Imai and In Song Kim. When Should We Use Unit Fixed Effects Regression Models for Causal Inference with Longitudinal Data? *American Journal of Political Science*, 63(2):467–490, 2019.

December 11
Final paper due

No readings assigned.