GRAD-E1322: Applied Longitudinal Data Analysis

1.   General information

Course Format	Onsite. The course is organized in onsite lectures. (3 hours per 8 sessions)
Instructor(s)	Michaela Kreyenfeld
Instructor's E-mail	kreyenfeld@hertie-school.org
Assistant (if applicable)	Alwine Hoppe hoppe@hertie-school.org William Fernandez W.Fernandez-Tinoco@phd.hertie-school.org
Instructor's Office Hours	Personal appointment (for setting a time, send email to instructor)

Link to Study, Examination and Admission Rules and MIA, MDS and MPP Module Handbooks

For information on **course room, times and session dates,** please consult the Course Plan on *MyStudies*.

Instructor Information:

Prof. Dr. Michaela Kreyenfeld is a professor of Sociology at the Hertie School in Berlin. Before coming to the Hertie School, she led the research group "Life Course, Social Policy, and the Family" at the Max Planck Institute for Demographic Research in Rostock. Between 2005 and 2012, she was a Junior Professor of Demography at Rostock University. She is co-director of the Einstein Center Population Diversity. She is a quantitative sociologist/demographer working in family, gender, aging, migration, and health research.

2.   Course Contents and Learning Objectives

Course contents:

This course provides an applied introduction to longitudinal data analysis. While this course teaches statistical methods, the focus is always on substantive research questions. Typical questions in this course are: How long does it take to leave poverty or unemployment? When do individuals leave the parental home, enter the labor market, get married, or have a child? Do processes differ by socio-demographic characteristics (such as gender, education, family context, and migration background)? How are they influenced by welfare state regulations and policy changes? How do life satisfaction, economic well-being, and earnings develop across the life course? What is the (causal) effect of health on economic outcomes? To answer questions of this kind, we use longitudinal micro-level data from various sources and regions of the world. We will employ different types of longitudinal data analysis techniques, such as event history models, sequence analysis, difference-in-difference techniques, fixed effects models, event studies, and related techniques.

Main learning objectives:
The course equips students with statistical skills to analyse longitudinal data. Students will be introduced to descriptive statistical methods, such as Kaplan-Meier survival functions. They will also get familiar with sequence analysis, a rather explorative and graphical technique to map life course patterns. The course also includes estimating and interpreting event history regression, fixed-effects models, DiD models, and event studies. This course will also enable students to deepen their R-skills.

Target group:

The course is designed for MA- and PhD-students interested in quantitative policy analysis.

Teaching style:

The course is organized in lectures (onsite). There will be problem-sets each week to understand measures and concepts and practice skills in class. Students should bring their laptops to the sessions.

Prerequisites:

The course does not require any deep statistical or mathematical knowledge, apart from the skills that were taught in Statistics I. Some R-skills are useful. If students do not have any R-skills, we offer a short tutorial the week before the start of the course. Students should contact the instructor before the first lecture in these cases.

Diversity Statement:

Understanding and respect for all cultures and ethnicities is central to the teaching at Hertie. Being mindful of diversity is an important issue for policy professionals in the planning, implementation, and evaluation of programs designed for specific groups, populations, or communities. Diversity and cultural awareness will be integrated into the course content whenever possible.

3. Grading and Assignments

Composition of Final Grade:

Assignment 1: **Research Paper**	Deadline: 02-04-2024, 5 p.m.	Submit via Moodle	40%
Assignment 2: Oral Presentation of Research Paper	Presentations are due in Session 7 (27-03-2025). Slides must be uploaded 1 day before the presentation (draft paper is optional)	Submit via Moodle	15%
Assignment 3: **Weekly Exercises & Participation**	Deadline: Mondays, 5 p.m.	Submit via Moodle	45%

Assignment Details

Assignment 1

The research paper includes a piece of empirical investigation based on the methods taught in the course (2,500 words +/-10%, excluding references and abstract, R-code should be included in Appendix, Teamwork +1000 words per author). It is expected that the paper not only includes the empirical investigation but also that the students select topics that allow them to develop testable hypotheses. The students will develop topics together with the instructor.

Assignment 2

Oral presentation (about 10 minutes) of planned research project. It is expected that all students prepare slides for their presentations and upload the slides one day before the presentation. They may also upload an extended abstract or a draft paper (optional).

Assignment 3

Data exercises will be handed out during lectures. Parts of the exercise must be completed in class, the rest has to be finalized at home. The completed exercises are to be uploaded on Moodle in the respective upload area. Each exercise contains a small task to practice the techniques we learned in class.

Late submission of assignments:
The deadline for the research paper is hard. Late submissions will be sanctioned with a reduction in the grade as stipulated below. Deadline extension is only granted based on a medical certificate or equivalent that must be submitted to the examination office.
- Up to 3 hours past deadline: 2-point reduction.
- Between 3 hours and less than 24 hours late: 5-point reduction.
- Per day late: 10 points (24-48 hours late means a 20-point reduction, etc.).

Attendance: Students are expected to be present and prepared for every class session. Active participation during lectures and seminar discussions is essential. If unavoidable circumstances arise which prevent attendance or preparation, the instructor should be advised by email with as much advance notice as possible. Please note that students cannot miss more than two out of 12 course sessions. For further information please consult the Examination Rules §10.

Academic Integrity: The Hertie School is committed to the standards of good academic and ethical conduct. Any violation of these standards shall be subject to disciplinary action. Plagiarism, misuse of AI, free riding in group work, and other deceitful actions are not tolerated. See Examination Rules §16, the Hertie Plagiarism Policy, and the Hertie Guidelines for Artificial Intelligence Tools.

Compensation for Disadvantages: If a student furnishes evidence that he or she is not able to take an examination as required in whole or in part due to disability or permanent illness, the Examination Committee may upon written request approve learning accommodation(s). In this respect, the submission of adequate certificates may be required. See Examination Rules §14.

Extenuating circumstances: An extension can be granted due to extenuating circumstances (i.e., for reasons like illness, personal loss or hardship, or caring duties). In such cases, please contact the course instructor and Examination Office *in advance* of the assignment deadline.

4. Course Sessions and Readings

All course readings can be accessed on the course Moodle page.

Session 1: Introduction to the Life Course Approach	
Learning Objective	The first session introduces the students to the (sociological) life course perspective. We also learn key terms and concepts (such as states, transitions, censoring). We will furthermore use R to visualize transitions between two time points.
Readings Theory	Elder, G.H., Johnson, M.K. & Crosnoe, R. (2003). The emergence and development of life course theory. In: Mortimer, J.T. & Shanahan, M.J. (Eds.): Handbook of the Life Course. Springer: 3-19. Kreyenfeld, M. & Konietzka D. (2021). Life course sociology. In: Schneider, N. & Kreyenfeld, M. (ed.): Research Handbook Family Sociology. Elgar. Mayer, K.U. (2009). New directions in life course research. Annual Review of Sociology, 35: 413-433.

Session 2: The Kaplan-Meier Survival Function	
Learning Objective	This session introduces the students to basic descriptive methods in the field of duration analysis. The substantive question that we seek to answer in this session is how welfare state contexts affect the timing of major life course events. We get familiar with survival functions calculated based on the Kaplan-Meier method as well as the Log-Rank Test.
Readings Method	Kleinbaum, D. G. & Klein, M. (2012). Survival Analysis: A Self-Learning Text (3rd ed.). Springer, pp: 1-54. Broström, G. (2021). Event History Analysis with R. Taylor & Francis. (Chapter 1):
Readings Theory	Billari, F.C. et al. (2021). The timing of life: Topline results from round 9 of the European Social Survey. ESS Topline Results Series 11. Billari, F.C. & Liefbroer, A.C. (2010). Towards a new pattern of transition to adulthood? Advances in Life Course Research, 15: 59-75. Chevalier, T. (2021). Varieties of youth transitions? A review of the comparative literature on the entry to adulthood. In: Catren, A.-M., et al. (Eds.). The Palgrave Handbook of Family Sociology in Europe. Palgrave Macmillan, Cham. Van den Berg L, Kalmijn M & Leopold T. (2021). Explaining cross-national differences in leaving home. Population, Space and Place. 27: e2476.

Session 3: Modelling Life Course Transitions with Time-Constant Covariates	
Learning Objective	Students will be introduced to regression techniques for duration data. We will get familiar with the Cox-model, the constant as well as the piecewise constant model. The substantive question that we seek to answer is whether there is a hidden gender bias in society. We adopt different strategies to uncover this gender bias. For example, we estimate event history models to uncover a "boy preference". Data for this analysis comes from the DHS from different regions of the world (such as Peru, Guatemala, Nigeria, Afghanistan).
Readings Method	Kleinbaum, D.G., & Klein, M. (2012). Survival Analysis: A Self-Learning Text. 3rd edition. Springer, pp. 97-159.
Readings Theory	Yoo, S.H., Hayford, S.R. & Agadjanian, V. (2017). Old habits die hard? Lingering son preference in an era of normalizing sex ratios at birth in South Korea. Population Research Policy Review, 36: 25–54. Becquet, V., Sacco, N. & Pardo, I. (2022). Disparities in gender preference and fertility: Southeast Asia and Latin America in a Comparative Perspective. Population Research Policy Review 41, 1295–1323.

Session 4: Modelling Life Course Transitions with Time-Varying Covariates	
Learning Objective	This session introduces time-varying covariates to the analysis of duration data. The substantive question that we seek to analyze is whether certain subpopulations (such as migrants, women, elderly) are subject to particularly long periods of unemployment. We furthermore raise the question of whether disadvantage cumulates. In other words, are migrant women the most disadvantaged group in the labour market?
Readings Method	see lecture manuscript
Readings Theory	Choo, H.Y. & Ferree, M.M. (2010). Practicing intersectionality in sociological research: A critical analysis of inclusions, interactions, and institutions in the study of inequalities. Sociological Theory, 28: 129-149. De Jong, G.F. & Madamba, A.B. (2002). A double disadvantage? Minority group, immigrant status, and underemployment in the United States. Social Science Quarterly 82: 117-130. Jacob, M. & Kleinert, C. (2014). Marriage, gender, and class: The effects of partner resources on unemployment exit in Germany. Social Forces, 92: 839-871. Keita, S. & Valette, J. (2019). Natives' attitudes and immigrants' unemployment durations. Demography, 56: 1023-1050.

Session 5: Panel Data Analysis and Difference-in-Difference Models	
Learning Objective	This session introduces the students to panel data. We will also deepen our R-skills to work with more complex data. In substantive terms, we examine how certain life course events (such as childbirth or marriage) affect life course outcomes and whether there are short- or long-term effects.
Readings Method	Miller, D. (2023). An introductory guide to event study. Journal of Economic Perspectives, 37, 203–230.
	Lechner, M. (2011). The estimation of causal effects by Difference-in-Difference Methods. Foundations and Trends in Econometrics, 4: 165-122.
Readings Theory	Lucas, R., Clark, A., Georgellis, Y. & Diener, E. (2004). Unemployment alters the set point for life satisfaction. Psychological Science, 15: 8-13.
	Luo, J. (2020). A pecuniary explanation for the heterogeneous effects of unemployment on happiness. Journal of Happiness Studies, 21: 2603-2628.

Session 6: Fixed Effects Analysis	
Learning Objective	This session introduces students to the fixed-effects approach. This approach largely operates over "de-meaned" data. We will use this method to study the effect of life course events on depression at advanced ages.
Readings Method	Allison, P.D. (2009). Fixed Effects Regression Models. Sage.
	Brüderl, J., & Ludwig, V. (2015). Fixed-effects panel regression. In H. Best & C. Wolf (Eds.), The SAGE Handbook of Regression Analysis and Causal Inference (pp. 327-357). Los Angeles: Sage.
	Hill, T. D., Davis, A. P., Roos, J. M., & French, M. T. (2020). Limitations of fixed-effects models for panel data. Sociological Perspectives, 63: 357-369.
	Collischon, M., Eberl, A. (2020). Let's talk about fixed effects: Let's talk about all the good things and the bad things. Kölner Zeitschrift für Soziologie 72: 289–299.
Readings Theory	Croezen, S., Avendano, M., Burdorf, A. & Lenthe, F. (2015). Social participation and depression in old age: A fixed-effects analysis in 10 European Countries. American Journal of Epidemiology, 182: 168-176.
	Gallie, D., & Russell, H. (1998). Unemployment and life satisfaction: A cross-cultural comparison. European Journal of Sociology, 39: 248-280.
	Kassenboehmer, S.C. & Haisken-DeNew, J.P. (2009). You're fired! The causal negative effect of entry unemployment on life satisfaction. Economic Journal, 119: 448-462.
	Stutzer, A. & Frey, B.S. (2006). Does marriage make people happy, or do happy people get married? The Journal of Socio-Economics, 35: 326-347.

Session 7: Presentation of Student Projects (mini-workshop: 4 hours)	
Learning Objective	In a "mini-workshop", students will present their research projects and receive feedback. Each presenter will have 5-10 minutes to present, followed by 5-10 minutes of discussion.

Session 8: Sequence Analysis	
Learning Objective	The session will be devoted to a more explorative approach in the area of longitudinal data analysis. We will use sequence analysis to map sections of life courses. In small teams, we will address a variety of topics, such as: the stability of political attitudes, the variability of self-employment, variations in partisanship, and right-wing populism.
Readings Method	Gabadinho, A., Ritschard, G., Mueller, N.S. & Studer, M. (2011). Analyzing and visualizing state sequences in R with TraMineR. Journal of Statistical Software, 40: 1-37.
Readings Theory	Schmitt-Beck, R., Weick, S. & Christoph, B. (2006). Shaky attachments: Individual-level stability and change of partisanship among West German voters, 1984–2001. European Journal of Political Research, 45: 581-608.