

Life Course Research (MPhil)

University of Oxford
2022-2023 (Hilary Term)

Module Convener: Dr. Charles Rahal

Guest Discussants: Eva Akimova, Andrea Tilstra, Daniel Valdenegro

1 Teaching Arrangements

1.1 Aims

The course introduces students to theories, perspectives, and methodological approaches to life course research in industrialising and post-industrialised societies. Whereas in Demographic Analysis (DA) we focused on more ‘macro’ levels of analysis, we now turn to the individual level, primarily in the form of survey, life history, and events data, as well as the appropriate statistical toolkits. This might otherwise be termed the ‘micro’ level of analysis. More broadly, the ‘Life Course’ (approach/analysis/research/theory) refers to the analysis of people’s lives within structural, social, and cultural contexts. In this course we will have four ‘topics’ mostly concentrating on theory, and then spend two ‘topics’ talking about heavily applied, empirical topics. The course will run for eight weeks.

1.2 Topics

Rather than being a siloed and specific pedagogical niche, the Life Course approach is designed to be a somewhat holistic way of thinking about everything inbetween birth and death. For example, how do early events influence or affect decisions and events such as marriage, the number of children to have, whether to engage in crime, or otherwise? How do earlier choices in life (e.g. to smoke) effect health in later life? We begin with a brief general introduction to life-course history, theory, data, and methods more broadly (Topic One), before turning to an introduction to social-epidemiology and health (Topic Two), and fertility and family demography (Topic Three). Topic Four focuses on predicting prospective events across the life-course, and the philosophical and ethical implications of doing so. Topic Five gives an introduction to sequence analysis, and Topic Six will deal with Sociogenomics given that the Sociogenomics module is not running in 2022/2023.

1.3 Content and Structure

The course runs for eight weeks. For all six of the topics, we will have a Lecture which represents the core of the teaching for that topic. All six topics will also have a Discussion Group, and for three of the topics, we will have a Lab. The final session of the term will represent an ‘open office hours’ session where you can come to ask questions about your summative assessments. All of the content will be delivered in person in the LCDS Conference Room – **apart from February 3rd, February 17th, and February 24th** – which will be held in the Department of Sociology Seminar Room. Each session will aim to be between 90-120 minutes. The two slots of the week that we’ll meet on are Tuesdays and Fridays at 15:30-17:30. If you are feeling unwell, be sure to send an email to charles.rahall@sociology.ox.ac.uk to be invited to join remotely via MS Teams. Further details on the timings are discussed below, as are – most importantly –

details on the **two** presentations **each** to be given in the discussion groups which represent the ‘Formative Assessments’ for the module.

1.4 Empirical Focus in R

As with DA, the module (in terms of labs and the programmatic assignment) is taught exclusively in R. Students are *strongly encouraged* to take the excellent and preparatory ‘Learn R’ sequence on Codecademy, which should take about 15-20 hours. As always, we recommend R for Data Science (‘R4DS’) for an introduction to R.

1.5 Core Textbooks

There is no formal course textbook which covers all of the topics (and there are indeed not many – if any – definitive/essential texts on the subject), but we recommend the following books for a general theoretical background:

- Giele, J and Elder, G.(1998), ‘Methods of Life Course Research: Qualitative and Quantitative Approaches’, SAGE Publications.
- Elder, G. and Giele, J., (2009), ‘The Craft of Life Course Research’, Guilford Press, 1st edition.
- Hunt, S., (2017) ‘The Life Course: A Sociological Introduction’, Palgrave; 2nd ed.

1.6 Other Texts

There are also many other books applicable to various parts of this module, and to the study of life course more generally. Lets recommend six books, perhaps, with one pertaining to each topic (the enumeration pertains to topic number):

1. Elder, G. (1974), ‘Children Of The Great Depression’, Routledge.
2. Pickles, A., Maughan, B., and Wadsworth, M., (2007), ‘Epidemiological Methods in life Course Research’, Oxford University Press.
3. Price, S., McKenry, P, and Murphey, M., (2000), ‘Families Across Time: A Life Course Perspective’, Oxford University Press.
4. James, G., Witten, D., Hastie, T., Tibirishani, R. (2013), ‘An Introduction to Statistical Learning: with Applications in R’, Springer Texts in Statistics.
5. Cornwell, B., (2015), ‘Social Sequence Analysis, Methods and Applications’, Cambridge University Press.
6. Mills, M., Barban, N., and Tropf, C. (2020), ‘An Introduction to Statistical Genetic Data Analysis - The MIT Press’, MIT Press.

1.7 Learning Outcomes

On successfully completing the course, students should:

- Be able to understand important concepts and perspectives in Life Course Research;
- Be able to undertake their own work (in R) in the field of Life Course Research;
- Be able to critically assess research at the frontier of the field;
- Be able to understand broad critical issues in contemporary and advanced life course analysis;
- Be able to comprehend modern methods in the field of Life Course Research.

1.8 Datasets

There are two key sources of data which we are going to repeatedly make use of within this module, and which you will have to register for.

1.8.1 Fragile Families Challenge

First of all, you are going to need a copy of the Fragile Families Child and Well Being Study (FFCWBS). Please sign up for a copy **now**! Steps to do this:

1. Please create an account at Princeton University's Office of Population Research data archive.
2. Read the Overview of FFCWS.
3. Click to 'Sign Up'.
4. In the box for justification, enter something like:

"I am requesting the Fragile Families Challenge data (specifically: `FFChallenge_v5.zip` and `ffchallenge_papers_replication_materials.zip`) in order to participate in Life Course Research taught as part of the MPhil in Sociology and Demography at the University of Oxford, convened by Dr. Charles Rahal."

Patiently wait. But sign up now, as access can take several weeks to be approved! *They are expecting your application!*

1.8.2 United Kingdom Household Longitudinal Study (UKHLS)

In addition to this, you'll also need to sign up for two teaching datasets provided by Understanding Society:

1. 'Study Number 8465 - Understanding Society: Ethnicity and Health Teaching Dataset Wave 1, 2009-2010'
2. 'Study Number 8715 - Understanding Society: Longitudinal Teaching Dataset, Waves 1-9, 2009-2018'

These are both available via the UK Data Service. Please note: This is 'Safeguarded' data: safeguards include knowing who is using the data and for what purpose, so to access data at this level you need to complete a simple registration and authentication process. Safeguarded data can be used for non-commercial, commercial and teaching projects. It is my responsibility to make sure that you register with the UK Data Service and that *you are aware of your obligations regarding data handling and licensing*; please read the requirements and agree to the EULA **as soon as possible**!

2 Assessments

2.1 Formative Assessment/Discussion Groups

The student-led presentations in our discussion groups will be framed as formative assessments. They will involve two/three students presenting a ‘Peer Review’ of a published paper (with or without slides, as you prefer) pertaining to the substantive topic of the week at hand. The ‘Peer Review’ presentation should last anywhere between 10-20 minutes depending on how much you feel necessary to say about the paper. It should be presented in a **critical way** where you:

1. Summarize the paper and the broader area;
2. Critically analyze the paper;
3. Describe the unique contributions of the paper which have resulted in it getting published (and ultimately chosen for this session)

All students should read – as an absolute minimum – the two/three papers to be presented each week before class. You can think about this as a non-optional homework! There will be three ‘starter’ discussion questions per paper provided a minimum of one week in advance on Canvas. Note there are seven students and six discussion groups which discuss two or three papers each: each student will present two papers across the course. Needless to say, these discussion groups form a crucial part of the course. The random allocator for the 14 papers can be executed as follows, with the person to paper allocations mapped in the Detailed Course outline (Section 2.4, below, with names in red). Note; the allocator randomly shuffles people into an ordering of 1-7, and then uses this same ordering twice to ensure an even (time-)split between each presenter.

```
1  set.seed(22122022)
2  Student.Name <- c("sant6264", "scat9237", "wolf6641",
3                    "sant6259", "kell6611", "kell6547",
4                    "scro4438")
5  Student.Name <- sample(Student.Name)
6  df <- data.frame(Paper = c(1:14))
7  counter = 0
8  for (paper in c(1:14)){
9    counter = counter + 1
10   df$Name[paper] <- Student.Name[counter]
11   if (counter==7){
12     counter = 0
13   }}
14  print(df, row.names = FALSE)
```

2.2 Summative Assignment

Similar to DA, we split the assignment into a qualitative (essay based, 3000 words, three options) and a quantitative assignment. The essay questions will be disseminated in **Week Five**, to be submitted by **Monday of Week One of Trinity Term (24th April 2023)**. The quantitative assignment is to be disseminated in **Week Eight**, and submitted by noon on **Friday of Week Four of Trinity Term (19th May 2023)**. The quantitative assignment should be submitted online via Inspira as PDF files. Open Office Hours: CR will be holding an ‘Optional Office Hours Drop-in Session’ in the final session of the term (Friday 10th March, 15:30-17:30), where students can voluntarily drop by either individually or as a group to ask questions pertaining to either the essay or the quantitative part of the summative assessment. These sessions could involve the individual discussion of the project/programming problem, or a ‘Q&A’ style session if the demand from students is there (please send programming queries in advance if possible).

2.3 Class Outline

Week One

Lecture One: ‘Introducing History, Theory, Data and Classical Methods for Life Course Analysis’

Reading Group One: ‘An Introduction to Life Course Research’

Week Two

Lecture Two: ‘Life Course Perspectives on Health’

Reading Group Two: ‘Topics in Health and the Life Course’

Reading Group Two Guests: Andrea Tilstra

Week Three

Computer Lab One: ‘Model Uncertainty within observational Life Course research’

Lecture Three: ‘Family and the Transition to Adulthood’

Week Four

Reading Group Three: ‘Topics in Family Demography’

Lecture Four: ‘An Introduction to Predictive Methods for the Life Course’

Week Five

Distribution of Essay Questions

Computer Lab Two: ‘Replicating and Extending the Fragile Families Challenge’

Reading Group Four: ‘Predicting the Life Course Out Of Sample’

Reading Group Four Guests: Daniel Valdenegro

Week Six

Lecture Five: “An Introduction to Sequence Analysis with Applications in R”

Reading Group Five: ‘Topics in Sequence Analysis’

Week Seven

Computer Lab Three: ‘An Application to Sequence Analysis with TraMineR’

Lecture Six: ‘An Introduction to Sociogenomics and the LifeCourse

Week Eight

Distribution of Quantitative Assignment

Reading Group Six: ‘Topics in Sociogenomics and the Life Course’

Reading Group Six Guests: Evelina Akimova

Open Office Hours: CR, Monday at 15:30 BST

————— Fin. —————

January				
Mon	Tue	Wed	Thu	Fri
16	17 Lecture 1 (15:30-17:30)	18	19	20 Discussion Group 1 (15:30-17:30)
23	24 Lecture 2 (15:30-17:30)	25	26	27 Discussion Group 2 (15:30-17:30)
30	31 Computer Lab 1 (15:30-17:30)			

February				
Mon	Tue	Wed	Thu	Fri
		1	2	3 Lecture 3 (15:30-17:30)
6	7 Discussion Group 3 (15:30-17:30)	8	9	10 Lecture 4 (15:30-17:30)
13 Distribution of Assignment One	14 Discussion Group 4 (15:30-17:30)	15	16	17 Computer Lab 2 (15:30-17:30)
20	21 Lecture 5 (15:30-17:30)	22	23	24 Discussion Group 5 (15:30-17:30)
27	28 Computer Lab 3 (15:30-17:30)			

March				
Mon	Tue	Wed	Thu	Fri
		1	2	3 Lecture 6 (15:30-17:30)
6 Distribution of Assignment Two	7 Discussion Group 6 (15:30-17:30)	8	9	10 Open Office Hours (15:30-17:30)

2.4 Detailed Course Outline with Readings and ‘Peer Reviews’

2.4.1 Topic One

Lecture Title: ‘Introducing History, Theory, Data and Classical Methods for Life Course Analysis’

Topics Covered:

1. An introduction to the module
2. Theory and history of life course research
3. Remembering life-tables and time in the lifecourse.
4. Types of and example applications of data
 - (a) Records and registry,
 - (b) Genetic,
 - (c) Life history
5. The rise of large longitudinal datasets (+overview of key datasets)
6. Three ‘walk through’ empirical examples of lifecourse data and analysis (in R)
 - (a) A panel data example using the UKHLS
 - (b) An example of the Biography package for life history analysis
 - (c) Inferential analysis with cross-sectional data (Angrist and Krueger, 1991)

Discussion Group Title: ‘An Introduction to Life Course Research’

Peer Review Papers

1. [sant6264](#): Billari, F. (2015), ‘Integrating macro- and micro-level approaches in the explanation of population change’, *Population Studies*, pp. 11-20.
2. [sant6259](#): Mayer, K. (2009), ‘New Directions in Life Course Research’, *Annual Review of Sociology*, 35, pp. 413-433.
3. [kell6547](#): Bernardi, L., Huinink, J., and Settersten Jr, R., (2019), ‘The life course cube: A tool for studying lives’, *Advanced in Life Course Research*, 41.

Additional Readings

- Elder, G.H. Jr. & Liker, J. K. 1982. ‘Hard Times in Women’s Lives: Historical Influences Across Forty Years’. *American Journal of Sociology*. 88(2): pp. 241-269.
- Tu, Y., Krämer, N. and Lee, W., (2011), ‘Addressing the Identification Problem in Age-period-cohort Analysis’, *Epidemiology*, 23(4) pp. 583-593.
- Becker, G. S. (1965), ‘A Theory of the Allocation of Time’, *Economic Journal*, 75, pp. 493-517.
- Newman, S.J., (2018), ‘Errors as a primary cause of late-life mortality deceleration and plateaus’. *PLoS biology*, 16(12), p.e2006776.
- Lee, R.D., (2001), ‘Demography abandons its core’, Remarks made flatpanel discussion on Micro-Macro Issues at the annual meetings of the Population Association of America.

2.4.2 Topic Two

Lecture Title: ‘Life Course Perspectives on Health’

Topics Covered:

1. Social and biological links between early life conditions and later life health.
2. Research design and inference challenges in life course research.
3. Inequalities in Health across the Life Course.
4. The fetal origins hypothesis.
5. Biomarkers across the Life Course.

Discussion Group Title: ‘Topics in Health and the Life Course’

Computer Lab Title: ‘Model Uncertainty within observational Life Course research’

Peer Review Papers

4. [kell6611](#): Lynch, J., and Davey-Smith, G. (2005): ‘A life course approach to chronic disease epidemiology’, *Annual Review of Public Health*, **49** (26): pp. 1-35.
 5. [scro4438](#): Hayward, M., and Gorman, B., (2004), ‘The long arm of childhood: the influence of early-life social conditions on men’s mortality’, *Demography* **41** (1):87-107.
- Tilstra, A., Gutin, I., Dollar, N., Rogers, R., and Hummer, R., (Accepted) “‘Outside the Skin’”: The Persistence of Racial Disparities in Early Life Mortality’, *Demography*

Additional Readings

- Riley, M. W., (1987), ‘On the significance of age in sociology’, *American Sociological Review*: 1-14.
- Lu, M. and Halfon, N., (2003), ‘Racial and ethnic disparities in birth outcomes: a lifecourse perspective’, *Maternal and Child Health Journal*, 7(1), pp. 13-30.
- Berkman, L., (2009), ‘Social epidemiology: social determinants of health in the United States: are we losing ground?’, *Annual Review of Public Health* 30, pp. 27-41.
- Almond, D., and Currie, J. (2011), ‘Killing me softly: The fetal origins hypothesis’, *Journal of Economic Perspectives*, 25 (3), pp. 153-72.
- Pearlin, L. I., Schieman, S., Fazio, E. M., and Meersman, S., (2005), ‘Stress, health, and the life course: Some conceptual perspectives’, *Journal of health and Social Behavior*, 46(2), pp. 205-219.
- Umberson, D., Crosnoe, R., and Reczek, C., (2010), ‘Social relationships and health behavior across the life course’, *Annual Review of Sociology*, 36, 139-157.
- Glymour, M., and Manly, J., (2008), ‘Lifecourse social conditions and racial and ethnic patterns of cognitive aging’, *Neuropsychology Review*, 18(3), pp. 223-254.
- Ploubidis, G. B., Benova, L., Grundy, E., Laydon, D., and DeStavola, B., (2014), ‘Lifelong socio economic position and biomarkers of later life health: Testing the contribution of competing hypotheses’, *Social Science and Medicine*, 119, pp. 258-265.
- Benson, R. et al., (2017), ‘Do work and family care histories predict health in older women?’, *The European Journal of Public Health*, 27(6), pp. 1010-1015.
- Phelan, J., Link, B., and Tehranifar, P., (2010), ‘Social Conditions as Fundamental Causes of Health Inequalities: Theory, Evidence, and Policy Implications’, *Journal of Health and Social Behavior*, 51, pp. 28-S40.

2.4.3 Topic Three

Lecture Title: ‘Family and the Transition to Adulthood’

Topics Covered:

1. Transitions to adulthood
2. Family dynamics
3. The second demographic transition and its critiques

Discussion Group Title: ‘Topics in Family Demography’

Peer Review Papers

6. [wolf6641](#): Beaujouan, E. and Berghammer, C., (2019), ‘The gap between lifetime fertility intentions and completed fertility in Europe and the United States: A cohort approach’, *Population Research and Policy Review*, pp. 38:507–535.
7. [scat9237](#): Kalmijn, M. and Monden, C. (2006), ‘Are the negative effects of divorce on well-being dependent on marital quality?’, *Journal of Marriage and Family*, 5 (68), pp. 1197-1213.

Additional Readings

- Billari, F., and Liefbroer, A., (2010), ‘Towards a new pattern of transition to adulthood?’, *Advances in life course research* 15, pp. 59-75.
- Lesthaeghe, R., (2010), ‘The unfolding story of the second demographic transition’, *Population and Development Review*, 36 (2), pp. 211-251.
- Seltzer, J., (2019), ‘Family change and changing family demography’, *Demography* 56, no. 2, pp. 405-426.
- Sobotka, Tomáš, (2008), ‘The diverse faces of the second demographic transition in Europe’, *Demographic Research* 19, pp. 171-224.
- Thornton, A. (2001), ‘The developmental paradigm, reading history sideways, and family change’, *Demography* 38 (4), pp. 449-465.
- Zaidi, B., and Philip-Morgan, S., (2017), ‘The second demographic transition theory: A review and appraisal’, *Annual Review of Sociology*, 43, pp. 473-492.
- Billari, F., (2001), ‘The analysis of early life courses: Complex descriptions of the transition to adulthood’, *Journal of Population Research*, 18 (2), pp. 119-142.
- Shanahan, M., (2000), ‘Pathways to adulthood in changing societies: Variability and mechanisms in life course perspective’, *Annual Review of Sociology*, 26 (1), pp. 667-692.

2.4.4 Topic Four

Lecture Title: ‘Predictive Methods in Prospective Life Course Research’

Topics Covered:

1. Moving from insample to out-of-sample; partitioning into training and test sets
2. Fundamentals of feature engineering
3. An introduction to cross-validation and model evaluation
4. Linear model selection and regularization.
5. Tree-based methods
6. Support vector machines
7. Parameter Hypertuning
8. Basic example: the Titanic dataset

Discussion Group Title: ‘Topics in Predicting the Life Course’

Peer Review Papers

8. **sant6264:** Yang, L. , Wang, S., Altman, R., (2022), ‘POPDx: an automated framework for patient phenotyping across 392,246 individuals in the UK Biobank study’, Journal of the American Medical Informatics Association, 10.1093/jamia/ocac226.
9. **sant6259:** Seligmana, B., Tuljapurkar, S. and Rehkopf, D., (2017), ‘Machine learning approaches to the social determinants of health in the Health and Retirement Study’, Social Science and Medicine: Population Health, 4, p. 95-99.
10. **kell6547:** Wei, L. et al. (2015), ‘Is Demography Destiny? Application of Machine Learning Techniques to Accurately Predict Population Health Outcomes from a Minimal Demographic Dataset’, PloS One.

Computer Lab Title: ‘Replicating and Expanding upon the Fragile Families Challenge’

Additional Readings

- Billari, F., Fürnkranz J. and Prskawetz, A., (2016), ‘Timing, Sequencing, and Quantum of Life Course Events: A Machine Learning Approach’, European Journal of Population, Vol. 22, No. 1, pp. 37-65.
- Verhagen, M. (2021), ‘The Pragmatists Guide to Prediction’, available on SocArXiv: <https://osf.io/preprints/socarxiv/tjkcy/>.
- Lundberg, I., Narayanan, A., Levy, K. and Salganik, M. 2018, ‘Privacy, ethics, and data access: A case study of the Fragile Families Challenge’, arXiv:1809.00103.
- Salganik et al., (2020), ‘Measuring the predictability of life outcomes with a scientific mass collaboration’, PNAS, 117 (15) 8398-8403.
- Garip, F. (2020), ‘What failure to predict life outcomes can teach us’, PNAS, 117 (15), pp. 8234-8235.
- Mullainathan, S., and Spiess, J., (2017), ‘Machine learning: an applied econometric approach’, Journal of Economic Perspectives, 31, pp. 87–106.
- Layard, R., Clark, A., Cornaglia, F., Powdthavee, N., Vernoit, J., (2014), ‘What Predicts a Successful Life? A Life-Course Model of Well-Being’, Economic Journal, 124 (580).

2.4.5 Topic Five

Lecture Title: ‘An Introduction to Sequence Analysis with Applications in R’

Topics Covered:

1. Introduction and Theoretical Background
2. Concepts and Data
3. Detecting Sequence Structure
4. Whole Sequence Comparison
5. An Introduction to the TraMineR Package

Discussion Group Title: ‘Topics in Sequence Analysis’

Computer Lab Title: ‘An Application to Sequence Analysis with TraMineR’

Peer Review Papers

11. [kell6611](#): Van Winkle, Z., (2018), ‘Family Trajectories across Time and Space: Increasing Complexity in Family Life Courses in Europe?’, *Demography*. 55 (1):135-64.
12. [scro4438](#): Barban, N., (2013), ‘Family trajectories and health: A life course perspective’, *European Journal of Population* 29 (4), pp. 357-385.

Additional Readings

- Sironi, M., Barban, N. and Impicciatore, R.(2015), ‘Parental social class and the transition to adulthood in Italy and the United States’, *Advances in Life Course Research*, 26: pp. 89-104.
- Ritschard and Studer (2018), ‘Sequence Analysis and Related Approaches: Innovative Methods and Applications’, Springer Open
- Marini, M. (1984), ‘The order of events in the transition to adulthood’, *Sociology of Education*, pp. 63-84.

2.5 Topic Six

Lecture Title: ‘An Introduction to Sociogenomics’

Topics Covered:

1. TBA

Discussion Group Title: Topics in Sociogeomics

Peer Review Papers

13. [wolf6641](#): Mills, M.C. and Tropf, F. (2020), ‘Sociology, Genetics, and the Coming of Age of Sociogenomics’, Annual Review of Sociology, 46, pp. 553-581.
 14. [scat9237](#): Kong, A., et al. (2018), ‘The nature of nurture: Effects of parental genotypes’, Science, 1359 (6374), pp. 413-433.
- Akimova, E., Breen, R., Brazel, D., and Mills, M.C., (2021), ‘Gene-environment dependencies lead to collider bias in models with polygenic scores’, Scientific Reports.

Additional Readings

- Belsky, D. et al (2018), ‘Genetic analysis of social-class mobility in five longitudinal studies’, PNAS, 115 (31) , pp. 413-433.
- Okbay et al. (2022), ‘Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals’, Nature Genetics volume 54, pp. 437–449.
- Mills, M., Barban, N., and Tropf, C. (2020), ‘An Introduction to Statistical Genetic Data Analysis - The MIT Press’, MIT Press.