9th GESIS Summer School in Survey Methodology
Cologne, August 2020

Syllabus for Course 7:
"Designing, Implementing, and Analyzing Longitudinal Surveys"

Instructors:	Dr. Tarek Al Baghal	Dr. Alexandru Cernat
E-mail:	talbag@essex.ac.uk	alexandru.cernat@manchester.ac.uk
Homepage:	www.iser.essex.ac.uk	alexcernat.com

Date: 17-21 August 2020
Time: 10:00-13:00 + 14:30-17:30
Time zone: CEST, course starts on Monday at 10:00
Venue: Online via Zoom

About the Instructors:

Dr. Tarek Al Baghal is a Research Fellow at the Institute of Social and Economic Research, University of Essex, UK. He is the lead questionnaire designer for Understanding Society: The United Kingdom Longitudinal Household Study, one of the largest longitudinal studies in the world. His PhD is in Survey Research and Methodology from the University of Nebraska, and he has a Master's in the same field from the University of Maryland.

Dr. Alexandru Cernat is a lecturer in Social Statistics at the University of Manchester. Previously he was a Research Associate at the Cathie Marsh Institute for Social Research and the National Centre for Research Methods, University of Manchester where he investigated non-response in longitudinal studies with a special focus on biomarker data. He has received a PhD in survey methodology from the University of Essex working on the topic of mixed mode designs in longitudinal studies.

Selected Publications:

- Cernat, A., Sakshaug, J. W., & Castillo, J. (2019). The Impact of Interviewer Effects on Skin Color Assessment in a Cross-National Context. International Journal of Public Opinion Research. https://doi.org/10.1093/ijpor/edy030
- Antoun, C., & Cernat, A. (2019). Factors Affecting Completion Times: A Comparative Analysis of Smartphone and PC Web Surveys: Social Science Computer Review. https://doi.org/10.1177/0894439318823703
- Sakshaug, J. W., Cernat, A., & Raghunathan, T. E. (2019). Do Sequential Mixed-Mode Surveys Decrease Nonresponse Bias, Measurement Error Bias, and Total Bias? An Experimental Study. Journal of Survey Statistics and Methodology, 1–27. https://doi.org/10.1093/jssam/smy024
- Wenz, A., Al Baghal, T. and Gaia, A. (2020) The Impact of Language Ability on Longitudinal Data Quality. Online first at Journal of Survey Statistics and Methodology
- doi.org/10.1093/jssam/smz045
- Sloan, L. Jessop, C., Al Baghal, T., and Williams, M. (2019). Linking Survey and Twitter Data: Informed Consent, Disclosure, Security, and Archiving. Journal of Empirical Research on Human Research Ethics. doi.org/10.1177/1556264619853447
- Al Baghal, T., Sloan, L., Jessop, C., Williams, M., and Burnap, P. (2019). Linking Twitter and Survey Data: The Impact of Survey Mode and Demographics on Consent Rates Across Three UK Studies. Online first at Social Science Computer Review doi.org/10.1177/0894439319828011

Short Course Description:

The course will provide an overview of those aspects of survey design, implementation, and analysis that are unique to longitudinal surveys or that have distinct features in the longitudinal context. The course will specifically cover:

- A review of the advantages and limitations of longitudinal surveys and an outline of some of the uses to which longitudinal surveys are put.
- Key aspects of longitudinal survey design such as the sampling design, interval between waves, and data collection modes.
- Important aspects of designing a questionnaire and measurement for longitudinal studies, particularly for capturing micro-level change
- The impact of non-response and attrition in a panel, and adjustments such as weighting and imputation given that missing data patterns differ between waves.
- An introduction to important analysis considerations and methods when using a longitudinal survey design.
- In addition, each morning time will be available in a computer lab for exercises for practical application of these methods and an opportunity for more interaction with the instructors.

Keywords:

Longitudinal surveys, data collection, measurement, attrition, weighting

Course Prerequisites:

- participants should have basic knowledge of survey methodology from a cross-sectional perspective, in particular with respect to survey design, instrument development, and survey implementation
- basic knowledge of statistics and statistical modelling (i.e., regression) and of a statistical software (Stata or R).

Target Group:

Participants will find the course useful if:

- they are interested in the uses and importance of longitudinal surveys ;
- they are involved in planning or conducting a longitudinal survey;
- they are interested aspects of longitudinal data collection;
- they are in the processes involved in measurement of phenomena over time;
- they want to use appropriate methods to analyze longitudinal data.

Course and Learning Objectives:

By the end of the course participants will:

- be familiar with the central design issues of longitudinal surveys;
- understand different strategies on how to collect longitudinal data;
- be able to design questions that meet research objectives for longitudinal surveys and minimize error;
- be able to identify and implement features that should help to prevent panel attrition;
- be able to study the nature of non-response and attrition in a longitudinal survey;
- understand generally the methods used in weighting and imputation in a panel survey.

Organizational Structure of the Course:

This is a five-day course with a total amount of 30 hours of virtual class time. Participants can expect a mix of interactive teaching and exercises. Exercises (in Stata and/or R) are designed to deepen their understanding of the course material.

- In these meetings, practical applications will be given for the students to work through, with the opportunity to ask instructors questions one-on-one about the course or their own projects.
- The practical application will relate to topics in longitudinal survey design covered in the lectures, including the impact of question wording on measurement, the identification of attrition, the development of weights, and how these can impact analytic findings.
- After the individual work the lecturer will go through the solutions with the entire class. Practical sessions will use Stata or R.

Software and Hardware Requirements:

We will use Stata and R for practical exercises. Participants who do not own a copy of Stata will be provided with access to a full Stata licence by GESIS for the duration of the course. Stata will be installed and activated prior to the course by GESIS staff through remote access on the participants' machines. Participants who (also) wish to use R should have a recent version of R (https://cran.r-project.org/) and Rstudio installed (https://www.rstudio.com/). Both programs are free and open source. For R the following packages should be installed: tidyverse, haven, sample, lme4, plm, broom.

Long Course Description:

Longitudinal surveys collect data from the same sample units over multiple time points, and are important in charting the changes that occur in a population. There are many benefits of longitudinal surveys compared to a cross-sectional study. One important benefit is the ability to measure change both at the macro- and micro-level. Others benefits include the possibility of identifying causality and efficiencies in data collection. As a longitudinal survey provides different benefits compared to a cross-sectional design, it also has unique design features that must be considered. This course aims at covering the main stages of designing and implementing longitudinal surveys, at revealing the main challenges faced when conducting longitudinal surveys and offering solutions for these. The lectures will be interactive, with students encouraged to provide questions and background on their relevant research. Practical examples will be used to illustrate how longitudinal surveys have been and are currently being conducted and used. The daily time for practical applications will specifically be focussed on interaction between students and the instructors, focusing on the lecture material.

The course will start by defining a longitudinal survey, and outlining the types of such survey that are carried out currently. Subsequently, basic longitudinal designs will be introduced and their strengths and weaknesses in terms of data collection and analysis potential will be discussed. We will discuss the main stages and decisions involved in planning a longitudinal study. The focus will be on design features and implications that are specific to longitudinal surveys.

Second, the course will focus on measurement in longitudinal surveys. In longitudinal designs, decisions should take into account analysis objectives, which are typically distinct from those of cross-sectional survey. Discussion will focus on the strategy of designing questions for longitudinal surveys, with special attention given to autobiographic memory processes. In panel studies, measurement errors can affect both inter- and intra-wave change indicators, and questionnaire design that can reduce these will be covered, specifically dependent interviewing and calendar interviews. Another possible measurement problem that will be covered is panel conditioning and its potential detection.

Next, the course will discuss nonresponse in panel surveys. In particular, we will identify and discuss unique causes of non-response in the longitudinal survey context and we will outline unique opportunities for non-response prevention. The discussion will focus on panel attrition and we will describe methods to reduce this, dealing separately with locating sample members, making contact with them, and obtaining their co-operation. These methods include targeted field strategies, respondent incentives, targeted letters and other communications, providing a preferred mode to respond, and tailoring questions to engage respondents.

Fourth, in part relating to issues of nonresponse, we will introduce the importance of and methods for weighting longitudinal survey data and imputation of missing data. Given changes in nonresponse and missing data across

waves, considerations should be taken for how these methods should be handled over time and how they can take advantage of longitudinal information. This session will provide a general understanding of why and how these methods are used, as well as specific detail to begin implementation in other surveys.

The last course session aims to link what we learned about longitudinal data collection with data analysis. We are going to talk about why traditional methods such as regression modelling are not appropriate for longitudinal data. We will then look at multilevel models of change with a focus on fixed and random effects and their use with real data.

In addition to the lectures, the instructors will be available in the practical application sessions. These sessions will be interactive in nature, fostering communication between the students and instructors using short applied exercises, designed to further the understanding of issues in survey design and collection of longitudinal data. During this time, participants will have the possibility to discuss challenges of their own longitudinal surveys during the course and individually with the instructors. The exercises will focus on application of materials, such the writing of survey questions related to their research interests and the calculation and application of weights to longitudinal data. Some of these assignments will be based on data from Understanding Society: The United Kingdom Longitudinal Household Study (for more information, see www.understandingsociety.ac.uk).

For computer-based assignments we will make use of the statistical package Stata and of R (students can choose which one to use).

Day-to-day Schedule and Literature:

Day	Topic(s)
1	**Introduction to Longitudinal Surveys; Research Questions; Sample Design; Survey Modes**
	Compulsory reading: ▪ Lynn, P. (2009). Methods for longitudinal surveys. Pp. 1-19 in: P. Lynn (Ed.), Methodology of Longitudinal Surveys. Chichester: Wiley. ▪ Groves, R. M., & Lyberg, L. (2010). Total Survey Error: Past, Present, and Future. Public Opinion Quarterly, 74(5), 849–879. http://doi.org/10.1093/poq/nfq065 ▪ Smith, P., Lynn, P. & Elliot, D. (2009). Sample design for longitudinal surveys. Pp. 21-33 in: P. Lynn (Ed.), Methodology of Longitudinal Surveys. Chichester: Wiley Suggested reading: ▪ Sikkel, D. & A. Hoogendorn (2008): Panel surveys. Pp. 479-499 in: E.D. de Leeuw, J.J. Hox & D.A. Dillman (Eds.), International Handbook of Survey Methodology. New York, London: Lawrence Erlbaum Associates. ▪ Lynn, P. (2011). Maintaining cross-sectional representativeness in a longitudinal general population survey. Understanding Society Working Paper 2011-04. https://www.iser.essex.ac.uk/research/publications/working-papers/understanding-society/2011-04. ▪ Buck, N.H., Ermisch, J.F. and Jenkins, S.P. (1995) Choosing a Longitudinal Survey Design: The Issues. Occasional Paper 96-1, Colchester: University of Essex. (Pages 1-20, 33-41) https://www.iser.essex.ac.uk/files/occasional_papers/pdf/op96-1.pdf ▪ Rose, D. (1995) Household Panel Studies: An Overview. Innovation: The European Journal of Social Sciences, 8(1)7-24. ▪
2	**Measurement in Longitudinal Surveys: What to measure, cognitive processes, problems, and questionnaire design**
	Compulsory reading: ▪ Jäckle, A. (2009). Dependent Interviewing: A Framework and Application to Current Research. Pp. 93-112 in: P. Lynn (Ed.), Methodology of Longitudinal Surveys. Chichester: Wiley ▪ Belli, R. F., Smith, L., Andreski, P., & Agrawal, S. (2007). Methodological comparisons between CATI event history calendar and conventional questionnaire instruments. Public Opinion Quarterly. 71, 603-622.

		Suggested reading:
		- Hox, J. J. (1997). From theoretical concept to survey question. In Lyberg, Biemer, - Collins, de Leeuw, Dippo, Schwarz, & Trewin (Eds.)., Survey measurement and - process quality. - Callegaro, M. (2008). Seam Effects in Longitudinal Studies. Journal of Official Statistics, 24: 387-409. - Chapter 4 in Tourangeau, R., Rips, L. & Rasinski, K. (2000): The Psychology of Survey Response Cambridge, MA: Cambridge University Press. - Sturgis, P., Allum, and Brunton-Smith, I. (2009). Attitudes Over Time: The Psychology of Panel Conditioning. Pp. 113-126 in: P. Lynn (Ed.), Methodology of Longitudinal Surveys. Chichester: Wiley. - Uhrig, SC. N. (2012) Understanding panel conditioning: An examination of social desirability bias in self-reported height and weight in panel surveys using experimental data". Journal of Longitudinal and Life Course Studies, 3: 120-136
3		**Attrition in Longitudinal Surveys: Specific issues and methods to counter nonresponse**
		Compulsory reading: - Watson, M. & Wooden, M. (2009). Identifying factors affecting longitudinal survey response. Pp. 157-181 in P. Lynn (Ed.), Methodology of Longitudinal Surveys. Chichester: Wiley. - Lepkowski, J.M. & Couper, M.P. (2002). Nonresponse in the second wave of longitudinal household surveys. Pp. 259-272 in R.M. Groves, D. Dillman, J.L. Eltinge & R.J. Little (Ed.), Survey Nonresponse. New York: Wiley. Suggested reading: - Couper, M.P. & Ofstedal, M.B. (2009). Keeping in contact with mobile sample members. Pp. 183-203 in P. Lynn (Ed.), Methodology of Longitudinal Surveys. Chichester: Wiley. - Laurie, H., Smith, R. & Scott, L. (1999). Strategies for reducing nonresponse in a longitudinal panel survey. Journal of Official Statistics, 15: 269-282. - Lynn, P. (2014). Targeted response inducement strategies on longitudinal surveys. Pp. 322-338 in U. Engel et al (Ed.s), Improving Survey Methods: Lessons from Recent Research. New York: Routledge. - Laurie, H. & Lynn, P. (2009). The use of respondent incentives on longitudinal surveys. Pp. 205-233 in P. Lynn (Ed.), Methodology of Longitudinal Surveys. Chichester: Wiley. - Banks, J., Muriel, A., & Smith, J. P. (2011). Attrition and health in ageing studies: evidence from ELSA and HRS. Longitudinal and Life Course Studies, 2(2), 101–126. - Groves, R.M. (2006). Nonresponse Rates and Nonresponse Bias in Household Surveys. Public Opinion Quarterly, 70: 646-675. - Groves, R.M, Singer, E. & Corning, A. (2000). Leverage-Saliency Theory of Survey Participation: Description and an Illustration. Public Opinion Quarterly, 64: 299-308.
4		**Weighting and Imputation in a Longitudinal Survey**
		Compulsory reading: - Lynn, P. (1996). Weighting for nonresponse. Pp 205-214 in R. Banks et al (Ed.s), Survey and Statistical Computing. Amersham: ASC. Chapter can be downloaded from http://iserwww.essex.ac.uk/home/plynn/downloads/. - Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. Psychological Methods, 7(2), 147. Suggested reading: - Seaman, S. R., & White, I. R. (2013). Review of inverse probability weighting for dealing with missing data. Statistical Methods in Medical Research, 22(3), 278–295. http://doi.org/10.1177/0962280210395740 - Peytchev, A. (2012). Multiple Imputation for Unit Nonresponse and Measurement Error. Public Opinion Quarterly. http://doi.org/10.1093/poq/nfr065 - Enders, C. K. (2010). Applied Missing Data Analysis (1st ed.). New York: The Guilford Press.

5	Introduction of analysis issues in longitudinal surveys
	Compulsory reading: ▪ Chapters 3, 4, 5 in Singer, J., & Willett, J. (2003). Applied longitudinal data analysis: modeling change and event occurrence. Oxford University Press. Suggested reading: ▪ Chapters 6, 7 and 8 in Singer, J., & Willett, J. (2003). Applied longitudinal data analysis: modeling change and event occurrence. Oxford University Press.

Preparatory reading:

General introduction to survey methodology and survey data collection (alternatively you could take a course such as "Introduction to Survey Design" in the first week):

▪ Groves, R. M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E. and Tourangeau, R. (2009). Survey Methodology. Hoboken: Wiley.

Additional Recommended Literature:

General introduction to survey methodology and survey data collection:
▪ de Leeuw, E. D., Hox, J. J. and Dillman, D. A. (Eds.) (2008). International Handbook of Survey Methodology. New York: Lawrence Erlbaum Associates.
▪ Dillman, D. A., Smyth, J. D. and Christian, L. M. (2008). Internet, Mail, and Mixed-Mode Surveys. The Tailored Design Method. Hoboken: Wiley.

General longitudinal data collection:
▪ Lynn, P. (Ed.). (2009). Methodology of Longitudinal Surveys (1 edition). Chichester, UK: Wiley.

Mixed modes and longitudinal studies:
▪ Cernat, A. (2015). Impact of mixed modes on measurement errors and estimates of change in panel data. Survey Research Methods, 9(2), 83–99. http://doi.org/10.18148/srm/2015.v9i2.5851
▪ Cernat, A. (2015). The Impact of Mixing Modes on Reliability in Longitudinal Studies. Sociological Methods & Research, 44(3), 427–457. http://doi.org/10.1177/0049124114553802
▪ Lynn, P., Uhrig, S.C.N. and Burton, J. (2010) Lessons from a Randomised Experiment with Mixed-Mode Designs for a Household Panel Survey. Understanding Society Working Paper 2010-03, Colchester: University of Essex. (Pages 1-6) http://research.understandingsociety.org.uk/publications/working-paper/2010-03.pdf

Online panels:
▪ Blom, A. G., Bosnjak, M., Cornilleau, A., Cousteaux, A.-S., Das, M., Douhou, S., & Krieger, U. (2016). A Comparison of Four Probability-Based Online and Mixed-Mode Panels in Europe. Social Science Computer Review, 34(1), 8–25. http://doi.org/10.1177/0894439315574825
▪ AAPOR Standards Committee. (2010). AAPOR Report on Online Panels (pp. 1–82). AAPOR Executive Council.
https://www.aapor.org/AAPORKentico/AAPOR_Main/media/MainSiteFiles/AAPOROnlinePanelsTFReportFinalRevised1.pdf]