Longitudinal Survey Research: Methodological Challenges

Institute of Philosophy and Sociology, Polish Academy of Sciences
Warsaw, Poland
December 15-18, 2015

Preliminary Abstracts

Conference:
The Present and Future of
Longitudinal Cross-sectional and Panel Survey Research

Day One, Tuesday, December 15

Plenary Lecture

Christof Wolf, GESIS, "Challenges of Survey Research"

Today, social surveys are the basis of most quantitative social science research. At the same time, concerns over the quality of survey data and the cost to produce high quality surveys, mainly conducted face-to-face, has sharply increased. In Germany, and I assume the situation is similar in other countries, the increase of survey costs partly results from higher demands by the research community, e.g. more effective supervision of field work, collection of para-data, geocoding etc., and it partly results from significantly dropping response rates leading to increased difficulties of survey organizations to reach respondents and obtain interviews.

What can we do? My answers to this question fall into three categories: a) improve existing surveys methods to increase response rates and decrease nonresponse bias; b) move to new survey methods, that are less expensive while still yielding high quality data; c) turn to other data sources that can at least partly replace survey data.

a) Improve existing surveys: We can improve current survey operations in many ways to increase response rates and decrease nonresponse bias. Examples are: unconditional incentives; better training of interviewers, closer monitoring of field work, real time information from field, flexible adaption of field work. Though these measures may counteract falling response rates they will certainly further increase costs.

b) New survey methods: One way to lower survey costs is moving away from the face-to-face method to less expensive methods. Currently it seems that either sequential mix-methods approaches or self-completion (either postal or online) offer the best trade-offs between cost and quality. Challenges here are in particular the selection of respondents or the sampling frame and the length of a survey which typically is lower for self-completion than for face-to-face.

c) Alternative data sources: Some of our research questions may be investigated with the help of what today is called “big data”. Examples include administrative data, for example, health insurance or pension fund data; other examples include data “found” online such as Twitter tweets, Facebook posts or Wikipedia entries.

Session One: Longitudinal Cross-sectional Survey Research

Rory Fitzgerald, Director of ESS ERIC and Coordinator of SERISS, City University London, UK

“Facing Up to the Challenges and Future of Repeat Cross-sectional Cross-national Social Surveys. The Synergies for Europe’s Research Infrastructures in the Social Sciences Initiative”

SERISS (The Synergies for Europe’s Research Infrastructures in the Social Sciences) is a new initiative that aims to address the most troublesome methodological challenges that cross-national surveys face and help to set new state of the art standards for cross-sectional and longitudinal survey methodology – from survey design to data access and storage.

Focused around three key themes – key challenges facing cross-national data collection, breaking down barriers between research infrastructures, and embracing the future of social sciences – SERISS will address issues relating to survey design and data collection, data management and curation from a collaborative, cross-national perspective. The project will better equip Europe’s social science data infrastructures to play a major role in addressing the key societal challenges facing Europe today and ensure that national and European policymaking is built on a solid base of the highest-quality socio-economic evidence. Bringing together ESS, SHARE, CESSDA and other leading cross-national European infrastructures the project has many synergies with the CONSIRT mission.

The presentation will draw on experiences from the European Social Survey and SERISS to illustrate how large scale infrastructures in Europe are facing up the future.

Melanie Revilla, Pompeu Fabra University, Spain

“Quality of Survey Data: How to Estimate It and Why It Matters”

There is no measurement without error. However, the size of the error can vary depending on the measure used. In particular in social sciences survey data, the size of the error can be very large: on average, 50% of the observed variance in answers to survey questions is error (Alwin, 2007).

Besides, the size of the error can vary a lot depending on the exact formulation of the survey questions used to measure the concepts of interest (Saris and Gallhofer, 2014). The size of the error can also vary across languages or across time. Thus, one of the main challenges for cross-sectional and longitudinal surveys, in order to make meaningful comparisons across groups or time, is to be able to estimate the size of the measurement error, and to correct for it. In this presentation, we will show how this can be done and what is the effect on the substantive results. However, even if the tools are available, in practice, these techniques are not implemented by most researchers. We want to argue that for the future of cross-sectional longitudinal survey methodology, this issue needs to be given more attention.

Alwin, D.F. (2007). *Margins of error: A study of reliability in survey measurement*. Hoboken, Wiley.

Saris, W.E. and I.N. Gallhofer (2014). *Design, evaluation and analysis of questionnaires for Survey research*, Hoboken, Wiley (second edition).

Peter Granda, Inter-university Consortium for Political and Social Research (ICPSR), University of Michigan

"Survey Data Documentation: The Disjunction between Description and Assessment"

When survey datasets appear for public use, some sort of documentation almost always trails along to tell researchers how to use and interpret the data. Interestingly, we often still call these documents "codebooks", a term that harks back to the very beginning of survey research.

This presentation will describe how standards for the content of these documents have evolved and how the introduction and development of the Data Documentation Initiative (DDI) over the last two decades has produced many new types of information and resources that researchers can currently access. It will highlight the role DDI has played in improving the comprehensiveness of the technical characteristics of variables and the important contribution it has made in facilitating data harmonization efforts. But this emphasis on improving variable-level metadata is only one side of the greater need to standardize data documentation in order to adequately address issues relating to data quality. The presentation will focus on such questions as:

- What is the difference between "descriptive" documentation and "assessment" documentation?
- Why has one developed much more quickly than the other?

- How can we encourage the greater use of documentation that assesses the QUALITY, not only of the variables in the public-use files, but of the entire lifecycle of the survey itself?
- What are quality profiles and what role can they play in standardizing survey documentation in the future?
- How does all of this relate to the Total Survey Error (TSE) paradigm?

Mitchell Seligson, LAPOP, Vanderbilt University, USA

"The AmericasBarometer by LAPOP: Challenges in Cross-National Longitudinal Surveys"

The AmericasBarometer is the largest, most comprehensive survey of democratic values and behaviors in the region. With roots that go back to the 1970s in the formation of the Latin American Public Opinion Project (LAPOP), since 2004 the surveys have grown to cover every independent mainland country in North, Central and South America, as well as almost all of the Caribbean. In the most recent round, over 50,000 people were interviewed in over 15 languages in more than 30 countries. We have learned a great deal in this process, with some of the most important lessons and challenges being:

- In terms of questionnaire quality, there is no substitute for conducting hundreds of test interviews (*not* standard pretests), covering the major cultural divisions of the region.
- Absent strict central supervision of sample design and execution, sampling errors doom comparability.
- A variety of problems in census bureaus throughout the hemisphere have increasingly weakened census mapping availability, thus undermining the key resource for efficient, reliable sample frames. The LAPOP newly developed LASSO© 2 system, however, drawing on remote sensing data, is increasingly becoming a viable alternative.
- Latin America has the highest levels of criminal violence in the world, a factor that drives up risks and costs, and pushes us to search for viable alternatives to the face-to-face interview.
- Funding constraints grow as the costs of fieldwork increase. Common solutions need to be found.

Session Two: Panel Survey Research

Xiaogang Wu, Division of Social Science, Center for Applied Social and Economic Research (CASER), Hong Kong University of Science and Technology (HKUST), Clear Water Bay Kowloon, Hong Kong SAR

"Household Panel Surveys in the Greater China: Current Situations and Future Challenges"

In this presentation, I will first introduce several household panel surveys in Hong Kong, mainland China, and Taiwan, i.e., the Hong Kong Panel Study of Social Dynamics (HKPSSD), the China Family Panel Studies (CFPS), and Panel Study of Family Dynamics (PSFD) in Taiwan, and some common theme for comparative studies. I then specifically focus on the research design and operational experience in three waves of HKPSSD data collections, respectively, in 2011, 2013, and 2015, to illustrate challenges in conducting longitudinal household surveys in modern society with increasingly mobile population and great emphasis on individuals' privacy, and the implications for procedures of data quality control and assessment. Finally, I argue that the future panel survey need to take advantage of the new technology and to combine government administrative records with individuals' survey data, to track the individuals more effectively with the limited cost.

Oliver Lipps, FORS, Switzerland

"Methodological Challenges of Panel Surveys Now and in Ten Years –A Swiss Perspective"

Panel surveys face most challenges in two areas: Data collection (including supplementary data such as register data) and data analysis. While there are improvements in data analysis (hardware, software availability and knowledge, scholar's longitudinal modeling skills), data collection is an increasing issue, especially keeping costs low and quality high. I therefore focus on data collection.

Panels are characterized by path dependencies which need a long-term governing strategy. An easy example is the modularization of questions (high frequency of few questions vs. low frequency of many questions, which questions in the same waves to allow for inclusion in multivariate models, etc.). It is difficult to adapt question (wordings), to change institutions (team, survey firm), the mode(s), and issues may change over time or are no longer appropriate, etc.

In view of this, the presentation will discuss the following challenges and suggestions:

- Improving institutional framework to getting funds to start panel. Easier if panels do not last eternally from the onset (flexible duration)?
- Reducing costs: mixing modes, more web. Tests needed to analyze breaks in the time series, including selection versus measurement errors, etc.
- Reducing initial under-coverage and nonresponse (bias): Incentives? Adapted modes/fieldwork? Nonresponse follow-up?
- Reducing (selective) attrition: Analyze underestimated change. Tailored incentives? Adapted modes/fieldwork?
- Increasing standardization and international comparability: ex-ante versus ex-post harmonization.
- Keeping a representative (cross-sectional / longitudinal) sample due to migration and demographic change: additional samples (costs!), adapt following rules (changed weights!).

Day Two, Wednesday, December 16

Małgorzata Mikucka, Centre for Demographic Research, Université catholique de Louvain, Belgium

"What Affects Subjective Evaluation of Health?"

The subjective evaluation of health is considered a legitimate indicator of overall health status, providing a valid and cost-effective measure of health. POLPAN contains two subjective measures of health: the evaluation of own mood (measured on a scale from 1 – very good to 4 – very bad) and the evaluation of own physical health compared to people of respondent's age (from 1- much better to 5 – much worse).

This paper discusses the meaning and the limitations of these measures. I focus on three aspects. First, I investigate correlation of these measures with measures of specific health problems included in the Nottingham Health Profile. Second, I discuss the propensity and ability of respondents to compare their health with the health of others. Finally, I explore if other measures included in POLPAN allow measuring respondents' optimism and I discuss the implications for the subjective measures of health.

I expect that the study will provide information on how to use and interpret the measures of health included in POLPAN. As similar measures are included in other surveys (e.g. European Social Survey), the study will also have implications for using these data. Finally, I expect that better understanding of the measures of subjective evaluation of health will be useful for framing health questions in the subsequent edition of POLPAN.

Sandra T. Marquart-Pyatt, Department of Sociology, Michigan State University

"A Cross-Lagged Panel Analysis of Democratic Values and Views of the State in Poland"

Previous research suggests that political attitudes are taking shape in unique trajectories in former state socialist countries compared with previous waves of democratic transitions. However, many of these investigations focus on cross-sectional analyses and are thus unable to disentangle contemporaneous and causal effects among a set of political attitudes. This research utilizes POLPAN, a panel dataset collected every five years since 1988, to address the question of how political attitudes develop in synergistic fashion in Poland. Using a cross-lagged panel design, this research examines the relation between democratic values and support for the welfare state or attitudes toward redistribution during a period of rapid social change. In comparing factors driving the formation of these attitudes, cleavage, emerging political context, and generational hypotheses are examined. Results point to some support for each of the aforementioned hypotheses. Implications of this research are discussed.

Anna Kiersztyn, Institute of Sociology at Warsaw University, and IFiS PAN

"Over-education in Poland, 1988-2013: Driving Factors and Consequences for Workers"

The last decades were a period of rapid educational expansion, giving raise to questions regarding the extent to which labor markets can accommodate the growing number of college graduates, and concerns that this change, rather than improving the labor market chances of individuals, may foster over-education. This paper seeks to assess three hypothetical explanations of over-education offered by the literature. The first assumes that the job structure is unresponsive to changes in the supply of workers with varying levels of schooling, and educated workers compete for a limited number of high skilled jobs. Educational expansion fosters an increase in inequality, as some graduates, particularly those from unfavorable backgrounds, are pushed into jobs with lower skill requirements. According to hypothesis 2, people voluntarily accept jobs for which they are overqualified in order to gain the experience and training necessary for career development, and over-education is a short-term phenomenon occurring at the beginning of a working life. Hypothesis 3 attributes over-education to the fact that people with the same educational credentials differ with respect to their actual skills, which are also determined by experience, cognitive ability, and other factors. These issues are examined on the basis of data from the Polish Panel Survey POLPAN, 1988-2013. First, I analyze the extent to which over-education is associated with individual socio-demographic characteristics, job mobility or cognitive capacity. Second, I assess whether being overeducated acts as a stepping stone or a dead-end for workers, using random-effects logistic regression models relating over-education to its lagged value.