



Summary of longitudinal survey methodologies

The MYWEB project is assessing the feasibility of a European Longitudinal Study for Children and Young People (ELSCYP). This paper provides a summary of longitudinal survey methodologies.

1. Introduction

This document presents the essential features of longitudinal methodology. It provides a short description of types of longitudinal surveys, their methodological advantages, methodological challenges, and different methods of data collection and principles underpinning the design of longitudinal surveys. More detail about the potential impact of longitudinal surveys on policy is provided in a separate briefing.

2. What is a longitudinal survey?

Surveys collect data about the economic, social, political, and cultural shape of a country. Policy makers can use this information to make informed decisions about future policies and evaluate the effectiveness of past or current policies. Contrary to a census, surveys collect information on a small part of the population. Yet when using probability sampling, it becomes possible to make inferences about

the whole population of relevance (Bethlehem, 2009).

Surveys can be divided in two main types: cross-sectional and longitudinal. The key difference is that cross-sectional surveys occur once whereas longitudinal surveys take place on multiple occasions over time (Lynn, 2009: 1). Cross-sectional surveys tell us what people are thinking or doing at one point in time. If the same questions are asked in another, later, survey, we can learn how these results change for the population as a whole, but we do not know how individuals have changed their views or their behaviour, or why Buck (2008). The only way we can do this is if we ask questions of the same people at different times in a longitudinal survey. We can then start to understand change and stability at the level of the individual, rather than for the population as a whole (Buck 2008).

Longitudinal surveys are “concerned with illuminating social change and with improving the understanding of causal inferences over time” (Bryman, 2012: 63).

3. Advantages of longitudinal surveys

Transitions: Longitudinal surveys gather “much longer continuous histories of events and transitions (...) than could be collected retrospectively in a single interview” (Lynn, 2009: 6). The data collected can also be more accurate than data gathered during a single interview, which relies on memories subject to recall errors. Furthermore, longitudinal surveys have the unique advantage of being able to compile unbiased information about expectations that can be analysed against



measures of outcomes collected at a later stage.

Patterns of change over time: Longitudinal surveys allow for the measure of stability or instability and the identification of causality. Individual-level change can only be understood in the context of changes taking place over a considerable amount of time. This type of analysis enables researchers to identify patterns of change (e.g. steady growth, fluctuation around a low level, sudden decline followed by stability) (Lynn, 2009). For example, if the proportion of children and young people satisfied with their life is relatively stable over time there might be many of them starting to feel satisfied with their lives while others are not satisfied anymore. A small proportion of children and young people might be satisfied with their lives on a continuous basis, while the majority show strong variations in time. This insight provides greater information about the dynamics and the factors associated with children and young people being satisfied with their lives. Longitudinal surveys also capture characteristics such as frequency, timing or duration. As noted by Richardson (2012), longitudinal studies make time trend analysis possible, which in the case of children and youth well-being would be useful to identify priorities in child well-being measurements and to validate child well-being statistics.

4. Distinctive challenges

Longitudinal studies have a number of disadvantages / challenges in terms of data collection and analysis (Lynn, 2009):

Sample attrition: This refers to the continued loss of respondents from the sample due to nonresponses through deaths, moving, and subjects withdrawing from the research. Some surveys successfully re-engage with sample members that have been non-respondents at one or more wave. However, the general response rate of sample units that would have responded to each questionnaire may be lower than in cross-sectional surveys.

Panel conditioning: Responses from participants may be conditioned by their previous experience of taking the survey. Panel conditioning can affect the way respondents report events, or even change respondents' behaviour.

Coverage error: This error occurs when there is a gap between the sampling frame and total population. Longitudinal surveys are more likely to suffer from coverage error if the sample does not include, or under-represents, additions to the population since the sample was selected. This coverage error is likely to increase with time.

Time and cost: It generally takes longer to prepare a longitudinal survey than a cross-sectional one since questionnaire instruments cannot be designed independently from those used at a later stage. Overall, longitudinal surveys are more costly than cross-sectional surveys.

5. Methods of data collection

Surveys can use different methods of data collection (Bethlehem, 2009).

Paper questionnaires can deliver good quality data but are time consuming and costly due to increased fieldwork and data entry costs.



Telephone interviews are efficient in terms of time and cost but limit the type of participants (those with a phone) and cannot be too long or complicated.

Mail surveys are very cost efficient but response rates and data quality can be low.

All of these instruments can be computerised¹, which enhances the quality of data collected, takes less time, and makes the work of the interviewer easier (Bethlehem, 2009).

In the case of longitudinal studies, it is possible to combine different methods of data collection, switching between different waves, in addition to any combination that could take place within one wave. It is also possible that one wave of data collection opens new possibilities for future waves, such as if respondents provide their email to take part in the web survey in the next wave (Lynn, 2009).

6. Designing a longitudinal survey

It is common to distinguish two types of longitudinal design: the cohort study and the panel study.

A cohort study is defined as the aggregate of individuals who experienced the same life event within the same time interval. These are often birth cohorts, but they do not have to be. Entry at primary school or secondary school could define a cohort. However, the cohort is defined it is followed over time. The main characteristic of a cohort study design is

¹ Methods of Computer Assisted Interviewing (CAI) include: Computer Assisted Personal Interviewing (CAPI), Computer Assisted Telephone Interviewing (CATI), and Computer Assisted Self Interviewing (CASI) such as Computer Assisted Web Interviewing (CAWI).

that the cohort is closed against new entries because such entries are, by definition, impossible (Ruspini 2002). A series of parallel cohorts can be used in an 'accelerated longitudinal design' (Galbraith *et al.* 2014) to allow an analysis of both age and cohort effects.

Panel surveys generally sample the whole population rather than single years of age with the aim of understanding the dynamics of change of the whole population, and its evolution over the lifetime of the study (Buck 2002). These are often based on households and follow all the people living in the sample household, not just a reference individual. There are different panel designs including:

- Fixed panel: data is collected from the same units on multiple occasions. Once the survey has started, no additions are made to the sample. In principle, only 'deaths' (i.e. emigration) reduces the size of the eligible sample.
- Repeated panel: This involves a series of panel surveys, which may or may not overlap in time. They are usually designed to represent an 'equivalent' population (i.e. the same population definition applied at different point in time). For instance, a repeated panel could consist of samples of the age of 16-17 one-year cohorts selected in different years, each panel involving at least three waves over at least three years.
- Split panel: This involves a combination of cross-sectional and panel sample at each wave of the study. For instance, a study would involve one panel sample for which data is collected each time, as well as a supplemental cross-sectional sample on each occasion.



Longitudinal surveys feature a range of specific design characteristics. For instance, the definition of the **population** under study needs to include a time dimension. The population that filled the questionnaire at Wave 1 and Wave 2 might be different to the population which filled it at Wave 1, 2, and 3 (Bethlehem, 2009). **Weights** are used in surveys to make a sample representative of the population (i.e. to have a similar distribution of gender, age, marital status etc.). Longitudinal studies can use both cross-sectional weights and longitudinal weights (Lynn, 2009).

In the case of longitudinal surveys a few design issues have to be considered against the quality of data collected and cost incurred. This is the case for the **interval between waves** where important changes can be missed if the interval is too long, but unnecessary cost will occur if the interval is too short. There is a similar balance to find in relation to the **duration of the survey**. The longer the survey, the richer and more valuable the data but the greater the costs (Lynn, 2009).

Most longitudinal studies include some repeated **measures** – they measure the same variables in the same way on each wave of questionnaire. However, the proportion of repeated measure vary greatly. For instance, household panel surveys usually use repeated measures to capture histories of economic activities whereas longitudinal studies focusing on children have very few repeated measures as the questions appropriate to ask and how to ask them are age-specific (Lynn, 2009).

Administrative data can be used to supplement or validate survey data. Researchers can use the administrative records to reduce the number of questions that need to be asked and therefore diminish the burden of respondents. Administrative data can also reduce recall bias (e.g. remembering the exact date an event such as a surgery occurred). Administrative data can also be used as the sampling frame, which enables researchers to understand and adjust for attrition. However, administrative data is not always available for all survey respondents due to challenges such as consent to linkage, success of linkage, and completeness of administrative data.

References

- Bethlehem, J. (2009) *Applied Survey Method, a Statistical Perspective*. Hoboken: John Wiley & Sons, Inc., Publications.
- Bryman, A. (2012) *Social Research Methods*. 4th ed., New York: Oxford University Press.
- Buck, N. (2002) *National Strategy for Longitudinal Studies*, Essex: ISER
- Buck, N. (2008) Introducing panel surveys: the BHPS in *In Praise of Panel Surveys*, Essex: ISER
- Calderwook, L. and Lessof, C. (2009) 'Enhancing Longitudinal Surveys by Linking to Administrative Data.' In Lynn, P. (ed.) *Methodology of Longitudinal Surveys*. Chichester: John Wiley & Sons Ltd,
- Galbraith, S., Bowden, J. and Mander, A. (2014) 'Accelerated longitudinal designs: An overview of modelling, power, costs and handling missing data' *Stat Methods Med Res* published online 20 August 2014 DOI: 10.1177/0962280214547150
- Lynn, P. (2009) 'Methods for Longitudinal Surveys.' In Lynn, P. (ed.) *Methodology of Longitudinal Surveys*. Chichester: John Wiley & Sons Ltd, pp. 1-19.
- Richardson, D. (2012) *An Evaluation of International Surveys of Children*. Paper prepared for the DG EMPL of the European Commission.
- Ruspini, E. (2002) *Introduction to Longitudinal Research*. New York: Routledge.

