SCOTTISH GOVERNMENT SOCIAL RESEARCH GROUP
SOCIAL SCIENCE METHODS SERIES

Guide 8: Longitudinal research

What are longitudinal studies?

A longitudinal study is an observational research study that collects information on the same sample (usually of individuals or households) at repeated intervals over an extended period of time – sometimes many decades. The various data collection periods are often called 'sweeps' or 'waves'. The Scottish Government funds, or is a major partner in, a number of longitudinal studies, including Growing Up in Scotland, the Millennium Cohort Study and the UK Household Longitudinal Study ('Understanding Society').

Although longitudinal research can collect qualitative as well as quantitative data, the vast majority of UK longitudinal studies to date have been quantitative, and they are the focus of this methods guide.

There are three main types of longitudinal studies:

Cohort studies

A cohort is a group of individuals who have experienced the same significant life event (typically birth) within a given period. Birth cohort studies (e.g. Growing Up in Scotland, the Millennium Cohort Study) follow up all individuals born on a particular day or week, from birth through to adult life. Age cohort studies begin data collection at a particular stage in life, such as the age at which they are eligible to leave school (e.g. the Scottish School Leavers Survey) or late adulthood (e.g. the English Longitudinal Study of Ageing).

Panel surveys

Rather than sampling on the basis of a significant life event, panel surveys sample a cross-section of the population, and then follow them up at regular intervals. For example, the UK Household Longitudinal Study (also called 'Understanding Society') is an annual survey of a nationally representative sample of around 40,000 randomly selected UK households. This world-leading study is a major investment (with a budget of almost £50 million) by a number of partners across the UK, and covers socio-economic, demographic, health and neighbourhood circumstances. It builds on the 18 years of data collected by its predecessor, the British Household Panel Survey.

Longitudinal data linkage studies

Existing administrative and statistical records, such as census returns, can be linked together to provide longitudinal data on individuals' demographic, socio-economic and health characteristics through time. For example, the Scottish Longitudinal Study is a large-scale linkage study of a 5.3% representative sample of the Scottish population created using census data, vital events data (births, deaths, marriages), National Health Service Central Register data (migration in or out of Scotland) and NHS data (cancer registrations and hospital admissions).

Guide 8: Longitudinal research

Advantages

Some of the key advantages of longitudinal data are discussed below.

Understanding causality

A major strength of longitudinal data is that they can help to disentangle causes and effects. This is crucial when assessing which factors policy should attempt to influence in order to improve social outcomes.

For example, there is unequivocal evidence from cross-sectional data that, overall, the unemployed have poorer health compared with the employed. Using cross-sectional data, it is not possible to identify whether unemployment causes poor health or poor health leads to unemployment. Longitudinal data can be used to examine whether the onset of poor health tends to precede the loss of one's job, or vice versa. This ability to observe the order of events makes causal inference with longitudinal data more valid than with cross-sectional data. In addition, as with analysis of cross-sectional data, multivariate regression analysis can be used to control for potential confounders, where data on these are available.

Similarly, longitudinal data can be used to examine the effects of policy interventions themselves, particular when policies are not implemented universally – for example, the ban on smoking in public places, which was implemented in Scotland before England; or the introduction of the Standard Grade qualification in Scottish schools, which was staggered over an extended period in the 1980s and 1990s across different Scottish regions. In these cases, the opportunity for a 'natural experiment' is presented.

Even when a policy has been implemented universally, there can be advantages to using longitudinal data. Repeat cross-sectional samples may not be directly comparable to one another, particularly over the longer term, as the demographic make-up of the population may change. For example, if an area has an improved level of employment after a regeneration policy has been implemented this may mask three underlying causes of the increase: increased employability of the people living there throughout the regeneration; higher employability of people moving into the neighbourhood because of the regeneration; or the lower employability of people moving out as regeneration proceeds. Only individual level longitudinal data can differentiate between these causes, by allowing us to look at how the circumstances of the people living in the area before the regeneration have changed or stayed the same.

However, as longitudinal studies are observational rather than experimental (e.g. in the case of the staggered introduction of the Standard Grade, children cannot be randomly allocated to live in different parts of the country), there is still the possibility that unmeasured confounding variables, such as cultural differences or differences between geographical areas, are responsible for observed effects. Even so, if we can assume the unmeasured variables are fixed over time, advanced statistical analysis techniques such as fixed effects and random effects modelling can be used in the analysis of longitudinal data to account for their influence. There is no way of accounting for the influence of unmeasured explanatory variables in cross sectional analysis.

Examining individual-level change

Longitudinal approaches are essential when the phenomenon of interest is directly concerned with individual level change, such as the dynamics of poverty or employment instability.

For example, repeated cross-sectional surveys can be used to estimate the proportion of people who are in poverty at different points of time. However, suppose that the proportion is estimated to be the same at two points in time - for example, the poverty rate in Scotland has remained around 17% since 2003. Longitudinal data allow us to examine whether it is the same people who are in poverty on both occasions, or whether there are equal and opposite flows into and out of poverty. This type of analysis can be carried out by simply by cross-tabulating the characteristics of respondents in one wave of a survey with their characteristics at a subsequent wave. Regression or event history analysis (see below) can also be used to examine characteristics associated with moves into or out of a particular state, such as poverty.

This type of analysis can provide more robust baseline data on a number of outcomes, allowing us to more fully understand the policy problem and tailor policy more effectively to improve outcomes.

Study of durations

As well as understanding flows of people into and out of particular states, we may also be interested in how long people are spending in a particular state. Event history analysis (also known as duration analysis, survival analysis, failure time analysis, duration economics and hazard modelling) is a range of techniques that allows us to measure the duration spent in a particular state (e.g. the time between becoming unemployed and finding a job, or being releases from prison and re-offending) and the factors that influence that duration (e.g. whether prisoners who have received correctional education have longer average times to re-offence than those who do not, when other factors that would predict re-offending are accounted for). This requires longitudinal data.

Disadvantages

Resources

Longitudinal studies can take a long follow-up time to generate useful data. Following up a large group of individuals over a long period of time can be costly and time consuming.

Representativeness and attrition

There may be changes in the general population over time, resulting in the longitudinal study sample no longer being representative of the wider population. In addition, there is invariably some amount of attrition (i.e. people dropping out of the study) from wave to wave, which can also result in the sample not being representative, e.g. people with lower incomes tend to be more likely to drop out.

Longitudinal studies are most useful for understanding relationships within individuals over time. However, this advantage diminishes over time if a substantial proportion of the original participants drop out of the study.

Panel conditioning

The experience of having taken part in previous interviews may affect the answers of respondents in subsequent interviews on the same topic, such that their answers differ systematically from the answers of individuals who are interviewed for the first time. This is particularly the case for questions assessing respondents' knowledge, but can also influence their reported attitudes and behaviours – although these effects have generally been found to be fairly modest.

Changing research agendas

During the course of a longitudinal study, new predictor variables may have been recognised as important, or better ways identified of measuring particular variables. A retrospective approach to questioning may be used to complete any gaps in data from earlier sweeps, or to fill in details of events that have occurred between successive sweeps. However, this can only be used with reliability for certain types of question.

Complexity

Although traditional, cross-sectional analysis techniques can be used with longitudinal data, more advanced statistical expertise and specialist software (e.g. Stata, MLWin) are required to fully exploit the potential of longitudinal datasets.

Guide 8: Longitudinal research

Examples of longitudinal data analysis

Women, lifelong learning and transitions into employment – National Child Development Study (NCDS)
http://cee.lse.ac.uk/cee%20dps/ceedp39.pdf

Aims
To examine whether lifelong learning that results in qualifications has any effect on transitions into employment of women in their 30s who are not in work.

Methods
Event history analysis was used to examine the factors that influenced transitions into employment between 1991 (when the NCDS cohort member were aged 33) and 2000 (when they were aged 42) of female cohort members who were not employed at the time of the fifth sweep of NCDS in 1991.

Results and policy implications
Among the 1,443 women in the NCDS who were out of work in 1991, 72% made a transition into employment between 1991 and 2000. After controlling for many other factors that could influence such transitions, including the presence of a partner and children, health status and prior education, adult learning leading to a qualification substantially increased the likelihood that a woman would return to employment.

Criminal Justice Transitions – Edinburgh Study of Youth Transitions and Crime
http://www.law.ed.ac.uk/cls/esytc/findings/digest14.pdf

Aim
To examine which factors predict having a criminal record at age 19.

Methods
Binary logistic regression was used to examine the characteristics and institutional histories of cohort members with a criminal record at age 19, as compared with youngsters with no such record.

Results and policy implications
Experience of school exclusion by third year of secondary education and leaving school at or before age 16 are important predictors of criminal record status by age 19. Therefore, reviewing education policy as it relates to school exclusion (in particular, whether such children might be better dealt with in alternatives to mainstream education) and maximising the number of youngsters retained in full-time education at school after age 16 may have a longer-term pay-off in criminal justice (as well as educational) terms.

Evaluation of the Working For Families Fund (WFF) - British Household Panel Survey (BHPS)

Aims
To compare the actual impacts of the WFF, which invested in initiatives to improve the employability of disadvantaged parents in Scotland, with what might have happened if there had been no WFF programme.

Methods
WFF recipients were matched with similar people in the BHPS, using a technique called propensity score matching, and their outcomes compared. They were matched according to their probability of renting accommodation rather than owning, which was ascertained using logistic regression. Gender, age, qualifications and disability were used as explanatory variables. The idea behind this is that similar propensity scores imply similar characteristics of the subjects from which the scores were generated.

Results and policy implications
The transition rates into employment from non-workforce active status over a one-year period were 40.6% for the BHPS sample and 58.3% for the WFF recipients, suggesting that the WFF is an effective policy.

Guide 8: Longitudinal research

Impact of children's early activities on cognitive development – Growing Up in Scotland (GUS)
http://www.scotland.gov.uk/Publications/2009/03/16101519/0

Aims
To assess whether children's early activities moderate the effect of socio-demographic factors on cognitive development

Methods
When the GUS birth cohort children were aged 10 months and 22 months, parents were asked about activities they carried out with their children, such as reading or other educational or play activities. Children's language development and problem solving skills were assessed at age 34 months. Regression analysis was undertaken to explore whether any of the activity measures were independently associated with cognitive ability scores once a range of socio-demographic factors were controlled for.

Results and policy implications
The number of activities carried out with children increased as affluence and advantage increased. However, children's early activities were found to have a strong impact on later cognitive development even when household income and area deprivation were controlled for. The same was found when the analysis focused on the less advantaged children only, suggesting that children's early activities can moderate the effect of socio-demographic disadvantage.

The dynamics of poverty - British Household Panel Survey (BHPS)
http://www.iser.essex.ac.uk/files/in-praise-of-panel-surveys.pdf

Aims
- To describe the extent and nature of persistent poverty in Britain, the main events ('routes') associated with movements into and out of poverty
- To examine what factors predict the length of poverty spells and the time between poverty spells.

Methods
Data from the first nine waves (1991-9) of the BHPS were examined. Event history analysis was used to examine how the length of poverty spells / the time between poverty spells varies with personal characteristics and with the length of time that individuals have already been poor / length of time since their last poverty spell.

Results and policy implications
Individuals' experience of poverty over a period of time more commonly reflected repeated short spells of poverty than a single long spell of poverty. Over a six-year period, about one-third of individuals were poor at least once, compared to the cross-sectional poverty rate of 18%. Only about 2% of individuals were persistently poor. This highlighted the need for measures to reduce the poverty entry rate (e.g. measures promoting job retention), not just those aiming to increase exits from poverty.

Relatively long poverty spells were experienced by pensioners and lone parents and their children, and the latter also became poor again relatively quickly following an exit from poverty. The chances of leaving poverty fell the longer the person had been poor, and the chances of falling back into poverty declined the longer the individual had stayed out of poverty. This underlined the value of early identification of those people likely to have long poverty spells and short recurrence times. Without early intervention, the problem to be addressed gets harder.

Guide 8: Longitudinal research

Further reading and resources

Inventory of UK longitudinal surveys. http://www.scotland.gov.uk/Topics/Statistics/About/longitudinal.
Provides information about UK longitudinal studies of relevance to Scottish policy makers.

Bynner J et al (2008). Use of Longitudinal Research in the Evaluation of the Scottish Government's National Outcomes. http://www.scotland.gov.uk/Publications/2008/11/25160140/0
Provides details of a wide range of longitudinal datasets, including their relevance to the National Performance Framework.

Economic and Social Data Service longitudinal website http://www.esds.ac.uk/longitudinal/
Provides access to longitudinal datasets and other useful resources.

Centre for Longitudinal Studies (CLS) http://www.cls.ioe.ac.uk/
An ESRC Resource Centre based at the Institution of Education, housing three British birth cohort studies: the 1958 National Child Development Study, the 1970 British Cohort Study and the Millennium Cohort Study. Provides support and facilities for those using data from the three studies, as well as conducting research using the cohort data.

UK Longitudinal Studies Centre http://www.iser.essex.ac.uk/survey/ulsc/
Established by the ESRC to promote longitudinal research and to design, manage and support longitudinal surveys.

Longitudinal and Life Course Studies http://www.journal.longviewuk.com/index.php/llcs/index
Open access journal providing up-to-date information on analysis techniques and application.

Longitudinal Data Analysis for Social Science Researchers http://www.longitudinal.stir.ac.uk/
Includes online training materials and an annotated reading list:
http://www.longitudinal.stir.ac.uk/refs/reading_lda_08.pdf

This guide was prepared by Alison Stout and Peter Matthews in the Office of the Chief Researcher.