Improving the evidence base of research...
There are very few dissertations that won’t be enhanced by some use of quantitative survey data. Even where the main research is primarily using qualitative methods, survey data can provide background and give valuable context.

Rich data on a wide range of topics of interest to social science students...
Undergraduates have access to a vast range of data on the topics Sociology and Politics students are interested in, from simple tables to full survey datasets.

Developing skills for employment...
Bringing quantitative data into a research dissertation is an ideal way to practice and develop quantitative skills that are highly sought after by employers in today’s competitive jobs market.

Using quantitative data in dissertations. It’s not all or nothing.
While for some students a secondary analysis of a survey data may form the main focus of their research, for many more it is likely that the use of quantitative data will be as part of a mixed methods approach or even just involve the selective use of some statistics to provide context for a more qualitative study.

It all depends on the research question.
How to Manage Undergraduate Dissertations Which Use Empirical Survey Data in Sociology and Politics

‘DATA-LIGHT’ Using statistics to provide context or background for the research

Some reference to published statistics can be invaluable when introducing a topic and developing a research question. For example, a student might be interested in a dissertation on the phenomenon of living alone, using qualitative methodologies to investigate the reasons why people choose to do so. A useful first step might be to give some statistics to show a historical overview of the prevalence of living alone. This information is available from the UK’s General Household Survey.

Re-using published tables and graphs

Bringing data like this into a dissertation does not always require students to access and work with original sources. In this example, we use an existing table sourced from the Guardian newspaper’s online ‘datablog’. When looking for data outputs on a particular topic, a good source are the published reports that accompany many of the Government social surveys. E.g. British Social Attitudes (http://www.bsa-30.natcen.ac.uk/) and Health Survey for England (http://www.hscic.gov.uk/catalogue/PUB09300) - these reports are packed with themed tables and graphs that can be re-purposed to provide evidence for setting context and developing arguments in dissertations.

Another valuable source are the growing number of dedicated websites providing on-line access to statistics on themes like ethnicity (www.ethnicity.ac.uk) and religion (www.brin.ac.uk).

Encouraging students to engage more with these kind of data outputs as part of their research will almost always improve a dissertation by strengthening the empirical basis to background chapters and the literature review and does not require more than very basic quantitative and computing skills to do. As with any use of secondary data, students should always be encouraged to reflect on such sources critically, and to ensure they give full acknowledgment to authors and original data sources.

How to Manage Undergraduate Dissertations Which Use Empirical Survey Data in Sociology and Politics

Getting closer to the data: using on line data services to create your own bespoke tables

Advances in web-technology have led to a proliferation of new sites that offer students more flexible on-line access to original survey data, with the opportunity to make customised tables without the requirement of specialist training and licensed software. With these sources, students can move beyond ‘data as context’ to include some more targeted exploratory analysis of key relationships. Some examples include....

**Local**

![Local Example](http://www.neighbourhood.statistics.gov.uk)

Neighbourhood Statistics

http://www.neighbourhood.statistics.gov.uk

Collates statistics for local areas from a number of sources on a wide range of topics from housing to crime. For projects incorporating a local case study it’s an ideal way to bring in some statistical data to help provide context.

**National**

![National Example](http://www.britsocat.com/)

British Social Attitudes

http://www.britsocat.com/

Data from one of the UK’s flagship social surveys, including a vast number of mainly attitudinal measures, many repeated over time. Easy to generate your own crosstabulations from the full archive of BSA surveys from 1983.

**European**

![European Example](http://www.ccesd.ac.uk/)

Centre for Comparative European Survey Data Information System:

http://www.ccesd.ac.uk/

Generate comparative statistics from an archive of 100,000s of survey questions for European countries spanning over 50 years.

**World**

![World Example](http://data.worldbank.org/indicator)

World Bank Development Indicators

http://data.worldbank.org/indicator

Over 3000 indicators available for the period since 1960. Could be used to provide background statistics for a country study or as data for a cross country analysis.

How easy are they to use? Most of the sites provide on-line guides and are generally very straightforward to use. Our website includes further guides, including one specifically design for dissertations.

License is Creative Commons. Use it and cite: Mark Brown and the ESSTED team. See www.socialsciences.manchester.ac.uk/essted
How to Manage Undergraduate Dissertations Which Use Empirical Survey Data in Sociology and Politics

Making the Data Centre Stage: Supporting students to undertake a secondary analysis of survey data

So far we have considered ways students can incorporate data outputs into dissertations where the data plays a mainly supporting role, providing context and background and some simple exploratory analysis. Survey data can be used more centrally when exploring relationships, testing hypotheses and even developing theories as part of a secondary data analysis. This requires access to the individual level data and use of a data analysis package (such as SPSS).

Where to start? A suitable research question.

A successful secondary analysis is dependent on the formulation of clearly specified research questions to guide the research design and analysis. Students often struggle with this crucial stage in the research process.

A common problem is where questions are expressed in very general terms, sometimes more as topics of interest than researchable questions. Where this is the case, even with a rich survey dataset students will quickly get bogged down in the data and lose a sense of what the analysis is aiming to do.

Helping students narrow down their ideas into a suitably specific question(s) for a secondary analysis is a key challenge for supervisors. Encouraging students to take a position with regards to relevant theory can be helpful here, inspiring questions that are framed in a way that involves using the data to ‘test’ a specific theory or part of a theory, expressed as a hypothesis or series of mini hypotheses.

One advantage of formulating research questions as hypotheses is that it helps give clear focus to all subsequent stages of the research, including the search and evaluation of a suitable dataset and the design of the analysis itself.

The good ‘test’ of a well specified question for secondary analysis is that it should make explicit the data requirements for the research, including the sample and the required key variables. This will guide the process of searching and evaluating potential datasets.

Example: From a broad topic to research hypothesis

Research topic:
Forms of political participation among the young

Relevant theory/idea:
That the young are disaffected with politics and disengaging from conventional forms of political participation.

Research question:
How does political participation vary by age in the UK, and how is this changing over time?

Research hypotheses:
- that the young show lower levels of prevalence on conventional measures of participation than older groups (e.g. voting)
- that the young show greater engagement with new forms of participation than older groups (such as e-petitions and protest marches)

In our example, the hypotheses identify the key concepts that need to be defined and operationalised (most obviously the various measures of participation), and a sample that allows us to compare young and old people in the general population.
Searching for Datasets: The UK Data Service (www.ukdataservice.ac.uk)

The UK Data Service is the gateway to a vast archive of empirical data, including hundreds of survey datasets. The vast majority of these can be downloaded and used freely by students and academics in HE under a simple end user licence and a quick online registration process.

The starting point for a search of potential datasets for a secondary analysis is the 'Data Discovery' page of the UK Data Service. This offers a highly flexible search facility. An initial data search using the search term ‘political participation’ returns 755 data hits! (The search can subsequently be refined by data type/subject/date or simply by using additional keywords).

A word on student surveys
A well designed quantitative survey requires proper training and considerable resources. While there are times when it may be appropriate, student surveys are notoriously problematic with many suffering from poorly designed questionnaires, small and highly unscientific samples and low response rates. With so many excellent surveys available students should be encouraged to look first to the existing survey resources before considering doing their own.
Evaluating data: is it fit for purpose?

There is a temptation to download the first suitable looking dataset - but students should be encouraged to first carry out a careful evaluation to see that the data really is fit for purpose. This centres on two key criteria:

- **The sample**... Do the survey respondents match the population of interest? Was it a random sample? Are there sufficient cases in the groups I want to compare?

- **The variables** Are you able to operationalise the key concepts involved from the variables available?

This task will be made so much easier if guided by a clearly specified research question.

A key advantage of the UK Data Service is that all datasets are accompanied by detailed documentation that can be used to carry out a detailed data evaluation on-line.

**NESSTAR**

(http://nesstar.esds.ac.uk/webview/index.jsp)

For many of the major surveys data evaluation can be done using the on-line software NESSTAR.

NESSTAR enables students not only to navigate the relevant documentation on survey design (generally available in user guides) but to visualise the data itself in the form of frequency tables and bar charts for each variable.

Users can actually go further and use NESSTAR to carry out simple analysis on-line, though we would recommend that where possible the main analysis is carried out using a package like SPSS after downloading the dataset onto a PC.

Data issues

The use of teaching datasets in methods classes means it is easy to gloss over some of the challenges of working with ‘real survey data’. In sourcing their own datasets students may encounter hierarchical data structures, complex weighting schemes and large numbers of missing values with which they are unfamiliar, and easily be put off. Providing support at this stage is crucial. A weekly drop-in clinic where students can get one to one support in setting up their data can make all the difference in keeping a project on track.
Data Analysis

The type of analysis conducted will obviously depend on the research question. However, in the majority of cases a secondary analysis at undergraduate level will involve relatively simple techniques, typically involving crosstabulation, correlation and/or comparison of means with the use of some controls. Where the aim is to generalise from the sample data to the target population, the inclusion of some appropriate statistical tests would be strongly recommended. Far more important than the sophistication of techniques used is the extent to which they are used appropriately as part of a coherent analysis and interpreted with a suitable degree of critical reflection.

Getting Critical

As with any methodological approach, students undertaking dissertations based on the analysis of survey data should be encouraged to discuss and reflect critically on all aspects of their research design and analysis. This should include some engagement with the epistemological and ontological assumptions being made in survey research, and include a critical discussion of the way central concepts and definitions related to the student’s research question have been operationalized in the research design.

Statistical significance: In encouraging the use of statistical tests, students should be encouraged to reflect on the difference between a finding that is statistically significant and a finding that is of substantive interest in the context of their particular research question.

Association and causation: A common pitfall in secondary data analysis is to over-interpret a statistical association between two variables as evidence of a causal relationship. When hypotheses are set up to test a relationship e.g. between political participation and age students should be encouraged to think of confounding effects of other variables and control for these wherever possible.

In most cases students conducting secondary analysis of surveys are probably best advised to avoid directly framing their research questions and hypotheses in terms of causality, especially when using cross-sectional data sets. In survey research, a better understanding of causal sequence generally requires working with longitudinal data, which is beyond the skill set taught to most undergraduate students in the Social Sciences. In the current example students might be encouraged to reflect on alternative explanations for an age difference in political participation, whether differences may simply reflect a life cycle effect (an individual’s political participation simply changes as they age) or a genuine cohort effect (where the nature of participation varies for different generations).

Why should students consider a secondary analysis of survey data?

Currently, relatively few students in Sociology and Politics conduct this type of dissertation. However a number of factors make it an increasingly appealing and viable option:

- **Data access**: The UK is uniquely well provided with survey data sets which have never been easier to access, download and use
- **Skills**: Training in data analysis (using SPSS or similar) is now increasingly a core component of methods training in the Sociology and Politics degree programme.
- **Employability**: A dissertation using secondary analysis of a social survey demonstrates a range of skills that are highly sought after by employers.
How to Manage Undergraduate Dissertations Which Use Empirical Survey Data in Sociology and Politics

ESSTED

This briefing paper is the second in a series produced by the ESSTED project. The authors work at the University of Manchester, and include Wendy Olsen, Mark Brown, Jackie Carter and Jen Buckley.

The ESSTED project is based at the University of Manchester and is working to better integrate the teaching of quantitative methods into the wider undergraduate curriculum. Social science students can benefit from learning about quantitative data and methods by seeing them demonstrated as integral to understanding of the substantive topic areas that inspire them.

Our interdisciplinary team, involving colleagues in Social Statistics, Sociology, Politics and MIMAS, are developing subject specific teaching resources. We also organise workshops and write briefing papers on different aspects of using quantitative data in teaching.

For more information see the project website: www.socialsciences.manchester.ac.uk/essted

For updates about briefing papers, workshops and teaching resources join the ESSTED mailing list (see the above URL).

Join our mailing list to receive future briefings.

License is Creative Commons. Use it and cite: Mark Brown and the ESSTED team. See www.socialsciences.manchester.ac.uk/essted