

Data Management Plan

Existing and New Data

Across the range of projects in Phase Two, we will use a variety of different data sources. Where possible, we will use existing data sets, and where appropriate we will adjust research questions to extract extra value from these existing data sets. However, as we plan to conduct innovative, inter-disciplinary research on a number of cross-cutting themes, we envisage that we will have to generate new primary data (both quantitative and qualitative) in a range of instances (i.e. projects 1a, 1b, 2b, 3a, 4b, 4c, 5b and 5c).

Information on Data

We plan to use a wide range of research methods implying various ways to collect data, including: semi-structured individual interviews; workshops/focus groups; statistical analysis of secondary data; analysis of documents, e.g. official policy documents, historical documents and academic literature; and calibration of numerical simulation models using secondary data. Secondary data will be collected as appropriate, e.g. by downloading from open-access sources or by placing a request. Where requests will need to be placed, preliminary scoping has already been undertaken to confirm that the data will in all likelihood be released to us.

All data will be stored in electronic format. The file formats used will depend on the methods employed, for example qualitative data could be stored in .doc format, in file formats used to support specialist qualitative data analysis software like Nvivo or Atlas TI, or as digitised voice recordings. Quantitative data could be stored in .xls format, in file formats used to support statistical/econometric software such as Stata and R, or in formats used to support mathematical modelling software such as Matlab.

Quality Assurance

All data collection will be undertaken by qualified and experienced scientists or social scientists, or by research students/assistants under their supervision. It will be, as usual, the responsibility of individual investigators and their teams to ensure appropriate documentation of data etc. It is normal practice in most areas of the social sciences to make datasets freely available for others to replicate (see below) and it is certainly our policy to do so; this provides a strong incentive for researchers to collect and document their data carefully.

Backup and Security

Data will be stored on the servers of LSE/Leeds, thus being covered by the backup and security practices of the two universities. Individual users will be expected to back data up in line with university policy and to take extra precautions when transporting data on portable drives. Both universities have the infrastructure to support remote access to servers, thus enabling e.g. uploading of data from the field when an internet connection is available.

Ethical Issues

Where relevant, i.e. principally in the collection of primary data from individual or group interviews, participants will be informed beforehand of how the data they provide will be handled and used. It is expected that in most if not all cases individual contributions will be anonymised, in line with standard practice. Informed consent will be sought to archive/share/use data as appropriate on a project by project basis.

Data Sharing and Archiving

As indicated above, it is expected that, unless provided to us on the condition that they must not be shared, data will be made freely available. We expect to therefore be able to share data in the overwhelming majority of our projects, however one or two projects are likely to make use of commercial data, e.g. from 'Bloomberg New Energy Finance' on the energy sector. It will not be possible to share or archive these data. Archiving, when acceptable to data providers, will be via the ESRC data archive.

Expected Difficulties in Data Sharing

We will share all data sets collected wherever it is appropriate to do so. However, in some instances we expect to be only given access to data on a confidential or restricted basis, and in such instances we will of course respect the conditions of any non-disclosure or similar agreements that we enter into. This may mean that some data is not shared, or that it is shared on a limited or partially restricted basis.

Copyright and Intellectual Property Rights

We will of course respect any copyright restrictions and acknowledge the holders of any copyrighted data sources or other materials in all outputs and publications. Copyright of all publications emerging from CCCEP research will initially be retained by project researchers/authors, however we expect that it would normally be reassigned to publishers on publication of journal papers etc. Intellectual property rights relating to all research outputs will be governed by the contractual arrangements between the LSE and the University of Leeds and the project researchers.

Responsibilities

Responsibility for all issues relating to data access, generation, use and sharing and for the associated copyright issues will ultimately rest with the project managers. However, CCCEP – and the research support and IT offices in the LSE and the University of Leeds – will provide any guidance and support needed to ensure that these responsibilities are suitably discharged.