ESRC-DFID Example Data Management Plan

Existing data

The research objectives require qualitative data that are not available from other sources. Some data exist that can be used to situate and triangulate the findings of the proposed research (eg, surveys of poverty impacts; opinion polls), and which will supplement data collected as part of the proposed research. However, qualitative and attitudinal data are generally rare or of insufficiently high quality to address the research questions. The research objectives also require quantitative analysis of public data. Some quantitative data are available, but they are insufficiently detailed. In their current form, they would not permit as full a comparison across the cases as is desirable.

Information on data

For these reasons, the research project involves primary data collection: 1) public data; 2) semi-structured interviews; and 3) focus group discussions with people identified through profiling techniques

1. Public data

Where possible, we will use online and/or electronic archives. This will involve extracting and processing quantitative data, includingparticipants, objectives and outcomes. Key search terms and their translation into the relevant languages, inclusion and exclusion criteria for items, variable codes and metadata will be refined and agreed in the inception phase of the project. Preliminary searches indicate that a sufficiently detailed dataset can be generated. The junior researchers will log their progress, documenting potentially contentious categorising decisions, difficulties faced in categorising items, and qualitative insights which do not fit the spreadsheet format. Data will be inputted and stored in a widely available spreadsheet format (eg Excel or SSPS), to ensure accessibility to Southern researchers.

2. Semi-structured interviews with individuals

The team anticipates undertaking 25-40 semi-structured interviews in each country from a sample frame to be developed in Phase 2. Data will be collected and stored using digital audio recording (eg MP3) where interviewees permit. In case they do not, interviews will be undertaken in pairs to enable detailed note-taking. Interview notes will be typed up according to agreed formats and standards. Where interviews are taped and in English, the UK research assistant will assist with transcription.

3. Focus group discussions matched to profiles

The sample frame for the focus group participants will be derived from public data. Numbers of focus groups will depend on geographical and other variations in patterns; how quickly a robust pattern of findings emerges; and the scope for identifying and convening the appropriate groups. Focus groups will involve two researchers, and be conducted in the vernacular. Whether recorded or not, the event will be transcribed or documented using agreed formats and standards for handling the issue of multiple voices, interruptions, labelling of participatory and visual activities, and so on.

All transcripts will be in Microsoft Word. All the researchers (except the UK research assistant) will be reasonably fluent in both English and the main language in which interviews and focus groups will be conducted, so that transcriptions will be translated into English only where the researcher is fluent in both languages and better able to transcribe in English, or to enable analysis of particular sections of the text. This will avoid unnecessary cost.

During the inception Phase 2, the metadata, procedures and file formats for note-taking, recording, transcribing, storing visual data from participatory techniques, and anonymising semi-structured interview and focus group discussion data will be developed and agreed. Focus group and interview transcripts will be coded in NVivo or a qualitative software suited to the different languages; the most appropriate software for a comparative multi-language study has not yet been identified.

Quality assurance

The PI will be responsible for overall quality assurance, with lead country researchers and the UK research assistant undertaking specific activities to ensure quality control. Detailed protocols for extracting data from secondary sourceswill be developed, piloted, refined and agreed in Phase 2. Quality will be assured through routine monitoring by the lead country researcher, and periodic cross-checks against the protocols by the UK-based research assistant. While interview and focus group protocols are being developed in Phase 2, standards and systems for note-taking, recording (if possible), transcribing and storing visual data from participatory techniques such as drawings, photographs and video, use of metadata, systems for downloading and storing SMS data (a potential follow-up research tool) will also be defined. Focus groups and interviews will always involve two researchers. Quality control for the qualitative data collection will be assured through refresher focus group discussion training during research design workshops and to junior researchers, where appropriate. Either the UK Institution or lead country researcher will check through each transcript for consistency with agreed standards. Where translations are undertaken, quality will be assured by one other researcher fluent in that language checking against the original recording or notes.

Backup and security

Our data will need to be backed up regularly; because of likely problems with viruses and hardware in developing countries, this will include regular email sharing with the UK research assistant, so that up-to-date versions are stored on the UK Institutions server. Qualitative data will be backed up and secured by the lead country researcher on a regular basis and metadata will include clear labelling of versions and dates. There are some potential sensitivities around some of the data being collected, so the project will establish a system for protecting data while it is being processed, including use of passwords and safe back-up hardware.

Ethical issues

A letter explaining the purpose, approach and dissemination strategy (including plans to share data) of the research, and an accompanying consent form (including to share data) will be prepared and translated into the relevant languages. A clear verbal explanation will also be provided to each interviewee and focus group participant. Commitments to ensure confidentiality will be maintained by ensuring recordings are not shared; that transcripts are anonymised and details that can be used to identify participants are removed from transcripts or concealed in write-ups. As the highly-focused nature of the research means that many participants may be easily identifiable

despite efforts to ensure anonymity or confidentiality, where there is such a risk, participants will be shown sections of transcript and/or report text to ensure they are satisfied that no unnecessary risks are being taken with their interview data. Interviews with elite policymakers will not guarantee confidentiality unless this is requested, as interviewees will be expected to speak in their official capacities or institutional roles. However, as is often the case, interviewees may be more comfortable if some sections of their interview are not recorded or made public. In such circumstances, recording will be paused or sections of text will be expunged from shared transcripts, and an indication made that this is the case.

Expected difficulties in data sharing

Not all of the transcripts will be translated into English (see above), which will limit the accessibility of the data.

Copyright/Intellectual Property Right

The institutional partners will jointly own the data generated. Online and archival sources will be cited and clearly acknowledged in the database and research outputs. Permission will be sought from secondary sources to share the findings of the research on public websites.

Responsibilities

The PI will direct the data management process overall, with the UK research assistant responsible for ensuring metadata production, day-to-day cross-checks, back-up and other quality control activities are maintained. The lead country researchers will be responsible for routine supervision of the dataset development. Data extraction, processing and inputting for the dataset will be undertaken by the in-country junior researchers. The UK Institution, lead country and junior researchers will share responsibilities for collecting and transcribing focus group and interview data, with the UK research assistant supporting as necessary. The PI will be finally responsible for dealing with quality and sharing and archiving of data.

Preparation of data for sharing and archiving

The most appropriate means of sharing the data generated through the project will be online, through institutional websites. The project will have a dedicated space on the UK Institutional website to facilitate this, and all other involved institutions will also be encouraged to host the data on their websites.