



## **SO5041 Unit 2: Surveys, Questionnaires and Sampling**

---

Brendan Halpin, Sociology

Autumn 2023/4

## **SO5041 Unit 2**

---

### **Survey research**

# Surveys and Survey Research

- Social survey research is very widespread:
  - political opinion polls and market research
  - EU-wide Labour Force Survey & CSO's Quarterly National Household Survey (LFS plus)
  - EU Eurobarometer
  - European Social Survey, World Values Survey, International Social Survey Programme
  - Household Budget Surveys
  - Growing up in Ireland, TILDA
  - Slán, Irish Study of Sexual Health and Relationships
  - surveys of business opinion, of inventories etc.
  - many emanating from ESRI or government

# Surveys: representativity

- The key principle of survey research is representativity: because the sample is random, summaries of the sample's characteristics can be imputed to the relevant population
- Sometimes we end up with too few cases of a subgroup to analyse – e.g., ethnic minorities; over-sampling or specially targetted surveys may help

# Longitudinal surveys

- Longitudinal surveys are a special case
  - Panel surveys the same sample at regular intervals (e.g., European Community Household Panel, US Panel Study of Income Dynamics, German Socio-Economic Panel, British 'Understanding Society' Panel Study)
  - Retrospective studies ask respondents to report complete life histories retrospectively (Irish Mobility Study, UK Family and Working Lives Survey, German Life History Study, etc.)
  - Cohort studies take a group of subjects and follow them forward (e.g., the Growing Up in Ireland study, The Irish Longitudinal Study on Ageing)
- Taking time into account makes these in many ways much richer data sources

## Questionnaire design 1/3

- The questionnaire is the linchpin of the survey
- Must elicit right information with minimum of ambiguity or suggestion, minimum inconvenience to the respondent
- Question design is a black art, since small changes of phrasing may cause different results

## Questionnaire design 2/3

- Extensive reliance on standardised questions, or standardised forms of questions (e.g., the typical five point answer scale: strongly agree, agree, neutral, disagree, strongly disagree)
- Standard schedules exist for certain purposes, e.g., the General Health Questionnaire
- See e.g.,  
[https://www.understandingsociety.ac.uk/documentation/mainstage/variables/?s=ghq&post\\_type=variable\\_mainstage&submit=Search](https://www.understandingsociety.ac.uk/documentation/mainstage/variables/?s=ghq&post_type=variable_mainstage&submit=Search)

## Questionnaire design 3/3

- V important to minimise “open” questions: much cheaper to pre-code answers (but allow an “other, specify” answer)
- Very important to test questionnaires in a pilot survey, to trap ambiguities and other problems, and to help pre-code questions



# Access

- Lots of survey data is available to the public, or to researchers
- Via data archives (e.g., the Irish Social Science Data Archive:  
<http://www.ucd.ie/issda>)
- Via government, EU, organisations like OECD
- Via website like European Social Survey:  
<http://www.europeansocialsurvey.org/>

## **SO5041 Unit 2**

---

### **Other forms of data**

## Other forms of data

- Administrative data
- Online data (e.g., Twitter)
- Topic-specific data such as about Covid-19
- "Big data": commercial data, online activity trackers, mobile phone records, Fitbit, traffic records

# Admin data

- Huge amounts of relevant administrative data is available
- Not survey: collected as a byproduct of the operation of the state
- Vital statistics: Births, marriages, deaths
- Censuses
- Tax, employment/unemployment, benefits, education, business
- Irish Central Statistics Office puts lots online at <https://statbank.cso.ie>
- See also OECD, Eurostat et alia.

## "Local" social media research example

- Recent research by a former UL MA Sociology student
- How affect (positive or negative emotions) in politicians' tweets affects readers' response
- "Affect" is judged by a complex software setup created by IBM but made available to researchers
  - This sort of machine-learning or AI system is increasingly important
  - Not necessarily as accurate as human raters but can cover much more data
- <https://www.nature.com/articles/s41599-021-00987-4>

# "Big data" increasingly important

- Matters more and more
- Requires different skills
- Sometimes threatens to replace conventional sources
  - Quicker, cheaper
  - But as accurate??
- But really big problems of representativity