# MORE2 Protocol Outline

## The MORE2 project

The MORE2 project was carried out by a consortium led by IDEA Consult (BE) with support from European Commission DG Research. The RISIS project complements the original study to present and open the data as one of set of other ERA datasets. It is presented by NIFU (Norway) with support from IDEA Consult.

## Basic characteristics

The Mobility Survey of the Higher Education Sector (MORE) survey is a study of researcher mobility patterns, the factors that shape them and the effects they have. It builds on MORE I (2009) survey, which was the first empirical study to focus on the mobility patterns of European researchers and their career paths. It was followed up in 2012 by the MORE2 study in order to support further evidence-based policy development on the research profession at European and national level. MORE2 targeted researchers working in different fields and career stages at higher education institutions (universities and other higher education institutions) in the EU27 member states, associated countries (Iceland, Norway, Switzerland) and candidate countries (Croatia, Turkey, the former Yugoslav Republic of Macedonia). In total, 33 countries were included in the MORE2 study.

The group of EU27 member states consists of Austria, Belgium, Bulgaria, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, and United Kingdom.

## The data and analysis methodology employed

The population-frame for the MORE2 study was primarily estimated on the basis of Eurostat headcounts of researcher-populations and of universities across Europe. MORE2 employed a two stage stratified random sampling strategy based on the population figures in Eurostat, supplemented by other sources including ETER. HEI researchers in EU27+6 were selected based on the FRASCATI manual definition. MORE2 sampled nearly 50,000 researchers in roughly 2.500 clusters. The clusters consist of the individual departments of EU27+6 HEIs. The online survey was launched by email and telephone in May 2012 and closed at the end of July 2012. After cleaning, 10,547 valid responses were yielded, which constitute the total number of researchers in the sample.

## Information on the database system

The MORE2 questionnaire includes 9 types of information. Here is an overview.

| |
|---|
| 1. Socio-demographics |
| 2. Education |
| 3. PhD and doctoral training |
| 4. Current employment and working conditions (including contract, status, satisfaction, inter- and trans-disciplinary mobility) |

| 5. Academic mobility and career paths (including PhD, including past and current mobility, including motivations, barriers and effects of mobility) | | |
|---|---|---|
| | a. PhD mobility | |
| | b. Further career mobility | |
| | c. <3 month mobility | |
| | d. Non-mobility | |
| | e. Virtual mobility | |
| 6. Collaboration / Virtual mobility | | |
| 7. Intersectoral mobility | | |
| 8. Awareness of EU policy | | |
| 9. Comparison research environments (EU – non-EU; EU countries) | | |

**Four distinct career stages**

In the MORE2 data, the researchers are grouped into four distinct career stages[1]. The variable "q13careerstage" gives information about these stages. The four career stages are:

- R1 (first stage researcher: doctoral candidate stage or equivalent, without having undertaken a doctorate),
- R2 (recognized researcher: PhD holder or equivalent who is not yet fully independent; post-doctoral stage),
- R3 (established researcher: researcher who has developed a level of independence; research specialist or manager, senior lecturer, senior scientist), and
- R4 (leading researcher: researcher leading his/her research area or field; professor stage).

**The MORE2 questionnaire**

The MORE2 questionnaire consists of 97 questions. A large number of questions are related to the researchers' careers, working conditions and mobility of researchers. The questions are grouped into several themes:[2]

- Background (q2 to q7)
- Education and training (q8 to q11)
- Current employment as a researcher, including PhD (q12 to q29)
- About your PhD training and mobility experience (q30 to q46)
- Your geographical mobility experience as a researcher (q47 to q62)
- Non-mobility (q63 to q65)
- About collaboration with other researchers (q66 to q67)
- Involvement with non-university sector (q68 to q81)
- Other topics (q82 to q87)
- Choice of job attributes – early stage researcher (q88 to q91)
- Choice of job attributes – later stage researcher (q92 to q97)

---

[1] Based on http://ec.europa.eu/euraxess/pdf/research_policies/Towards_a_European_Framework_for_Research_Careers_final.pdf

[2] The first question (q1) is used to exclude persons that do not considered themselves as researchers from the MORE2 survey.

In the MORE2 questionnaire, there are three main questions related to mobility: long-term international mobility experience (q47), short-term international mobility experience (q61), and intersectoral mobility (q68). The three main questions related to mobility are only answered by the R2-R4 groups.

**The non-response follow-up survey**

Two types of non-response surveys were organized for the MORE2 study:

- During the first CATI interviews, a short telephone survey was conducted in case of an explicit refusal to participate to the MORE2 survey. Each time the CATI team received an explicit refusal, the researchers were asked if they had been long term mobile (q47) and/or intersectoral mobile (q68).

- A short web based survey focusing on CATI non-respondents was implemented after the initial data collection process. Researchers were asked if they had long-term mobile (q47), short-term mobile (q61) and/or intersectoral mobile (q68).

The non-response survey had as its goal to come up with coefficients to align the shares of in-sample responses with the ones of the population, taking into account non-response bias. These coefficients have been incorporated in the sample weights and have led to adjusted, non-response bias-robust sampling weights. These are included in the MORE2 data under the main names calwq47, calwq61 and calwq68.

**The structure of the MORE2 survey**

1. Survey Design: two stage stratified random sampling strategy
    a. Strata: the sample is stratified by country and by field of science. There are 99 strata (33 countries x 3 broad Fields of Science).
        i. Country strata 1: 27 + 6 countries
        ii. FOS strata 2: 3 field of science
    b. Cluster: "Department A of University B in Field of Science C in Country D".
    c. PSU: a faculty of a given university in field of science x in country y. The PSU depends on the size of the university/departments.
2. Population and response. Distinguish between the following levels:
    ⇨ Population (n=1,388,000), n=1,389,407 (Eurostat figure 2009)
    ⇨ Sample (n=49,056)
    ⇨ Response rate: 21.5%
        ○ Valid responses (n=10,547)
        ○ Unit non-response (n=38,517)
        ○ Gross sample (valid responses + unit non-response) (n=49,064[3]): See the Excel fil more2_gross_sample
    ⇨ Item-response for R2-R4 researchers (q41, q61 & q68) (n=8,357)
        ○ The R1 researchers (n=2,190) did not receive the questions concerning mobility (q47, q61 & q68).

---

[3] The gross sample number of 49,064 exceeds the sample number of 49.056 by eight units. This is due to eight units being excluded from the gross sample during additional (telephone) contacts as doubles, but not being registered as such in the system. Therefore we cannot ex post identify those eight units and they are still included in the gross sample.

3. Calculation of survey weight
    a. Design weights: adjust for probability to be included in the sample (inclusion probability):
        i. Probability that Department A of University B is selected in the strata
        ii. Probability that researcher from Department A of University B is selected in the strata
    b. Response rate per stratum:
        i. See the Excel file more2_sample_distribution.
4. Accounting for response probability and non-response
    a. Unit non-response: the probability that questionnaire is not responded to by the sampled researcher appears to be non-random.
        i. Dimension and distribution of unit non-response:
        ii. Steps to deal with selective non-response:
            1. Non-Response study
            2. Calibration: The editing for non-responses is meant to decrease a potential bias in the answering patterns by calibrating the given responses of all researchers to a proportion that corrects for this potential bias. In order to account for this potential bias one needs to apply the calibrated weights in the MORE2 data.
    b. Item non-response (probability that specific question is not answered by sampled researcher).
        i. Selective non-response: subgroups are correlated with target variables (q47, q61 & q68 mobility questions). Three factors (youngest quintile, part-time, junior positions) account for most of the non-response. This is entirely due to routing in the survey, since it was always the intention that certain questions were only asked to a subgroup of researchers.
        ii. Partial responses: Among the returned questionnaires, 757 were partially filled and they have been initially excluded from the sample.
5. The use of weights: The weight "weihc" is applied to the following questions: q1 to q46, q63 to q67, and q82 to q87. Concerning the calibrating: If one uses q47, then one needs to apply the weight "calwq47hc". If one uses q61, then one needs to apply the weight "calwq61hc". If one uses q68, then one needs to apply the weight "calwq68hc". All questions that follow up on q47 (q48 to q59), q61 (q62) and q68 (q69 to q81) do not need to include a calibration adjustment when the indicators that we are calculating are expressed as shares or percentages. However, when the indicator is an absolute number, it is necessary to apply the calibration adjustment.
6. The svyset command: If one uses the weight "weihc", the svyset command is: svyset university [pweight=weihc], fpc(FPC) strata(strata) || _n. If one uses the weight "calwq47hc", the svyset command is: svyset university [pweight=calwq47hc], fpc(FPC) strata(strata) || _n. If one uses the weight "calwq61hc", the svyset command is: svyset university [pweight=calwq61hc], fpc(FPC) strata(strata) || _n. If one uses the weight "calwq68hc", the svyset command is: svyset university [pweight=calwq68hc], fpc(FPC) strata(strata) || _n.
7. Singleton: When using the svyset command one may receive the following note: "Missing test statistics because of stratum with single sampling unit". This concerns subgroups of researchers with only one observation included.

## Compatibility with the MORE1 HEI survey

The compatibility of MORE1 and MORE2 HEI surveys is limited due to methodological changes in the definitions of mobility and in the survey strategy. We would therefore not recommend to use the results as 'time series' data, but rather as different and separate datasets.

The most important differences are listed in the following table:

|  | MORE1 | MORE2 |
|---|---|---|
| Definition of mobility | Reference country is country of previous education, where education includes PhD. | Reference country is country of highest educational attainment, excluding PhD. Expanded definition in terms of contract types (e.g. to include doctoral candidates under scholarship or stipends). |
| Sampling strategy | Aimed at representativity at EU level. | Aimed at representativity at country level. |
| Questionnaire | Singular questions to provide direct answers to the research questions. | Curriculum-approach where the researcher is asked to provide raw data on each education and career step. The data are consequently edited and analyzed by the study to provide answers to the research questions. |
| Analysis | No information on current research career. | Strongly based on information of current research career. |