RISIS RESEARCH INFRASTRUCTURE FOR RESEARCH AND INNOVATION POLICY STUDIES

Report on the content and technical structure of the

MORE II

Infrastructure

NIFU

Nordic Institute for Studies in Innovation, Research and Education

RISIS "Research infrastructure for research and innovation policy studies"

FP7, Grant agreement no: 313082

Task 6, Workpackage 6, coordinated by AITAustrianInstitute of TechnologyGmbh

SEVENTH FRAMEWORK PROGRAMME

Report on the content and technical structure of the *MORE2* infrastructure (Task 6 of WP6)

Table of contents

Contents

1 Basic characteristics

Name: Mobility Survey of the Higher Education Sector: Mobility and Career Paths of Researchers in Europe, 2012 (MORE2)

Improving the conditions for researcher mobility is (re)emerging as a central priority of European research and innovation policy. This priority is prominent in the Europe 2020 strategy and in current policy initiatives under the European Research Area. These forward-looking objectives build on an extensive tradition in Europe for the study of researcher mobility patterns, the factors shape them and the effects they have. This tradition traces at least back to the European Human Capital and Mobility program (1992).

The Mobility Survey of the Higher Education Sector (MORE) survey is a key study in this tradition. MORE I (2009) is the first in a family of wide-scale empirical studies that focused on the mobility patterns of European researchers and their career paths. It was followed up in 2012 by the MORE2 study to provide internationally comparable data, indicators and analysis in order to support further evidence-based policy development on the research profession at European and national level. The survey of researchers in higher education in Europe was the first work package of this study, complemented by a (limited) survey of researchers outside Europe, case studies on the working conditions and career paths of early stage researchers and on the remuneration of researchers and the collection of a set of internationally-comparable indicators in existing sources. This Report presents the MORE2 HEI survey in Europe and it builds on a parallel report detailing the MORE1 dataset.

MORE2 targeted researchers working in different fields and career stages at higher education institutes in all EU27 countries and 6 (at the time) Associated and Candidate countries[1]. The study provides measures of flows of international, interdisciplinary and intersectoral mobility, of factors that influence mobility and non-mobility (motivations and barriers), and of effects that can be linked to researcher mobility. Also topics such as structural PhD training and virtual mobility are addressed. The survey responses thus provide promising avenues for a range of studies to better understand the career development paths of researchers working in Europe.

Important to mention is that the collected data are accurate at country level (EU27, Candidate and Associated countries) thanks to the specific sampling design and implementation.

This Report provides the background information about the MORE2 to indicate the type of dataset it is, how it will be opened up, and how it might be combined with other data. In general, the infrastructure will be operated by NIFU[2]. Access is foreseen in three forms: On-site access at NIFU, thereafter on-line access, and access during approved RISIS training events.

All accessibility environments will respect privacy/confidentiality issues while attempting to provide maximum analytical possibilities/learning outcomes. The platform will be piloted with MORE I data. In time, this pilot will subsequently be complemented by the

[1] Candidate Countries: Croatia, Turkey, the former Yugoslav Republic of Macedonia
Associate Countries: Norway, Switzerland, Iceland

[2] NIFU's legal name is Nordisk institutt for studier av innovasjon, forskning og utdanning (NIFU). www.nifu.no

micro-data from the MORE2[3] survey, carried out by IDEA Consult (Belgium). Details of the MORE1 survey are provided in a preceding report on which underlying document builds.

2 Information on substantive content of MORE2

The MORE2 study was to provide internationally comparable data, indicators and analysis in order to support further evidence-based policy development on the research profession at European and national level. It thereby, like the MORE1 study, produced an accurate register of researcher-populations or of universities across Europe.

MORE2 employed a two stage stratified random sampling strategy. It sampled almost 50,000 researchers in around 2.500 clusters and yielded 10,547 valid responses (response rate of 21.5%) through a multichannel survey, combining both CATI (computer-assisted telephone interviews) and CAWI (computer-assisted web interviews).

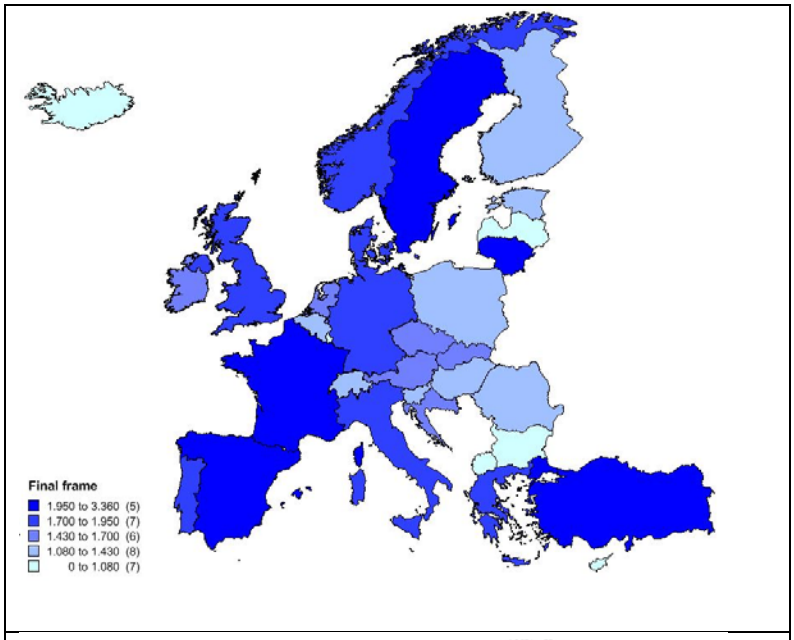
The responses combine information about (in **bold** information that is new or more elaborated in MORE2 compared to MORE1):

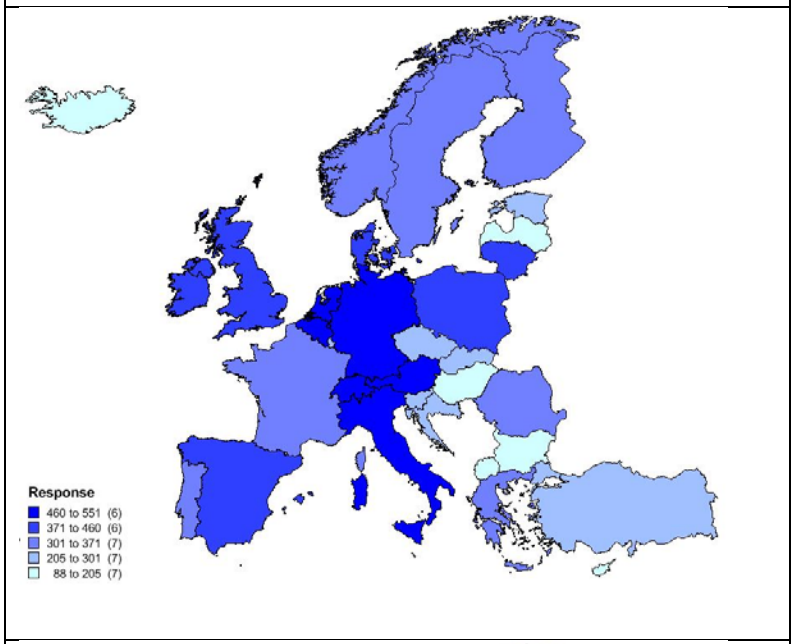
MORE1	MORE2
1. The researcher (country of birth, citizenship(s), gender, age, children),	1. Socio-demographics
2. Education (degrees, graduation year, country, field of highest degree)	2. Education 3. PhD and **doctoral training**
3. Current position (university/college, faculty, field, position level, seniority)	4. Current employment and **working conditions** (including contract, status, satisfaction, inter- and transdisciplinary mobility)
4. Mobility events (up to five mobility events, countr(ies), duration, type)	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
5. Assessment of mobility among mobile as distinct from non-mobile researchers: a. Detailed focus on most recent mobility event (motivations, push and pull factors, assessment) b. Plans/aspiration to work in another country: (country, rationale and background for choice of destination)	
	8. Awareness of EU policy 9. Comparison research environments (EU – non-EU; EU countries)

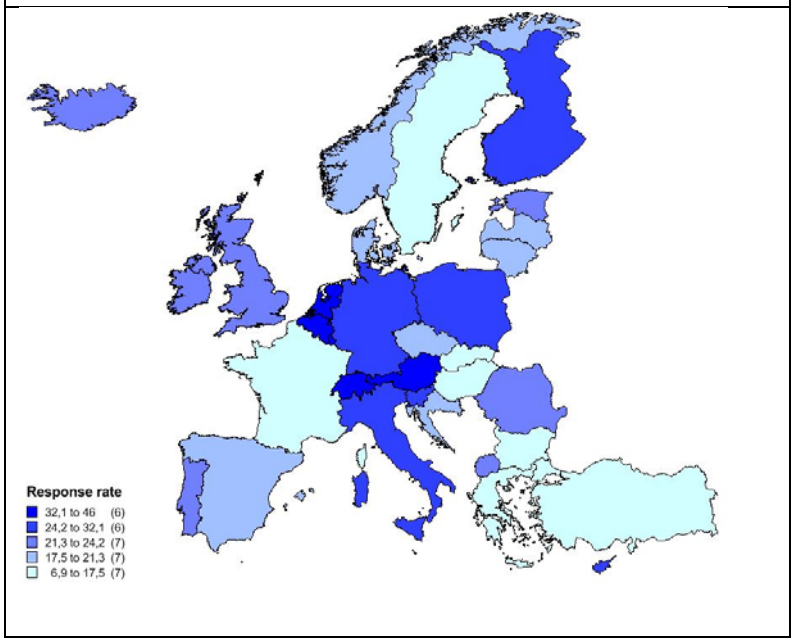
The responses provide a snapshot of mobility patterns and career paths of EU27 researchers (in 2012). The sample is distributed across EU27 countries and 6 (at the time) Associated and Candidate countries[4]. The geographical distributions of the final frame, response and response rate are presented below.

[3] http://www.more-2.eu

[4] Candidate Countries: Croatia, Turkey, the former Yugoslav Republic of Macedonia
Associate Countries: Norway, Switzerland, Iceland


Final frame
1.950 to 3.360 (5)
1.700 to 1.950 (7)
1.430 to 1.700 (6)
1.080 to 1.430 (8)
0 to 1.080 (7)


Response
460 to 551 (6)
371 to 460 (6)
301 to 371 (7)
205 to 301 (7)
88 to 205 (7)


Response rate
32,1 to 46 (6)
24,2 to 32,1 (6)
21,3 to 24,2 (7)
17,5 to 21,3 (7)
6,9 to 17,5 (7)

2.1 Definition and description of observations

It is important to mention here that the MORE I frame was built as a first in its kind, based on desk research and web search techniques. The MORE2 study did not fully repeat this intensive exercise but based its frame on the MORE I frame, merged with the EUMIDA database[5]. The applied definitions and principles in this process are described below.

Researchers

MORE2 employed a two stage stratified random sampling strategy based on the population figures in Eurostat. HEI researchers in EU27+6 were selected based on the FRASCATI manual definition[6]. The definition covers professionals with tertiary (or higher) education and not "technicians and equivalent staff" or "other supporting staff". The accuracy of the definition was confirmed by a question in the survey:

> We specifically target "researchers" within this survey, including people:
>
> - carrying out research OR
> - supervising research OR
> - improving or developing new products/processes/services OR
> - supervising the improvement or development of new products/processes/services.
>
> If you consider yourself to fall into one or more of the above categories, we kindly ask you to complete the questionnaire.
>
> I…
>
> ☐ consider myself a researcher
> ☐ do not fall in to one of the above categories

Researchers population

The population of researchers, which determines the frame of the study, was estimated using Eurostat's headcounts for 2009, supplemented where necessary by estimates (based on earlier years or from other statistics). The total population was estimated to be roughly 1.4 million researchers of this type in the EU27+6 countries (of which 1.2 million in EU27). The ETER effort (and its predecessor, Eumida) have since helped to refine the population, standardize the names and location, and provide a map of the distribution of researchers in Europe. As noted, HEI names in the MORE data will be linked to the ETER designations. These can be used to improve the accuracy of the responses via post-stratification techniques (see also sections 2.4 and 5, below).

Fields of Science

[5] 4 countries were not present in the MORE I frame or the EUMIDA database: Turkey, Croatia, Macedonia (FYROM) and Iceland. To complete this and other lacking information, an additional updating procedure was organised through web search routines on the most relevant directories of universities and on the official sources, i.e. research ministries.

[6] http://www.oecd.org/innovation/inno/frascatimanualproposedstandardpracticeforsurveysonresearchandexperimentaldevelopment6thedition.htm

The classification of the fields of science (FOS) is the main benchmark for the specialization of HEIs and of researchers and it is therefore also the stratification criterion for the sample within each country. MORE2 follows the same criteria as applied in MORE1 and selects 3 fields of science, fully compatible with official statistics and with the EUMIDA project database. The FOS classification is an aggregation of the six FOS classification proposed by the OECD in 2006 according to the following scheme:

- FOS 1 (Natural sciences) and FOS 2 (Engineering and technology) will fall in NATURAL (Abbreviation in the project)
- FOS 3 (Medical sciences) and FOS 4 (Agricultural sciences) will fall in HEALTH (Abbreviation in the project)
- FOS 5 (Social sciences) and FOS 6 (Humanities) will fall in SOCIAL (Abbreviation in the project)

Higher Education Institute (HEI) and HEI cluster

The clusters consist of the individual departments of EU27+6 HEIs. A university department is defined a part of the HEI (faculty or department) that is specialised into only one FOS, regardless of its formal nature within the structure of the institution and standardised in term of size within each country. The precise definition of a cluster is thus "Department A of University B in Country C and Field of Science D". This definition is equal to the one applied in MORE I, although a check led to merge a number of clusters into one according to the rule of one field of science.

All higher education institutes in the merged EUMIDA/MORE1 database, both of public or private nature, were therefore considered eligible for the sample frame under the condition that they included at least one ISCED 6 student (PhD level) and/or one researcher. The HEI cluster database was subsequently updated to update information on existing HEI clusters, include new clusters or delete those that were no longer active.

It is important to note that a two stage stratified cluster-sampling strategy is distinct from a standard random sample survey. Within the each country and field of science (the strata), the MORE2 sample grouped the population units not on researchers but on departments of HEIs. A primary sampling unit (psu) is thus a faculty of a given university in country x and field of science y (cluster).

Two-stage stratified random sampling

Using these definitions and data, the sampling strategy stratified by country (33) and broad fields of science (3 – Natural Sciences and Technology, Medical Sciences and Agriculture, and Social Science and Humanities). This resulted thus in 99 strata (33*3), to be sampled from the 'clusters' at the level of individual university departments. Almost 2,500 HEIs were sampled across EU27+6 countries.

A list of researchers within each cluster was collected and researchers were further selected into the sampling frame under the assumption of random sampling. University department websites were checked for researchers' email addresses (updates and new addresses compared to MORE1). After cleaning, 49,056 individual researchers were identified as targets.

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.[7]

2.2 Data acquisition and processing

a. Where are the data retrieved from:

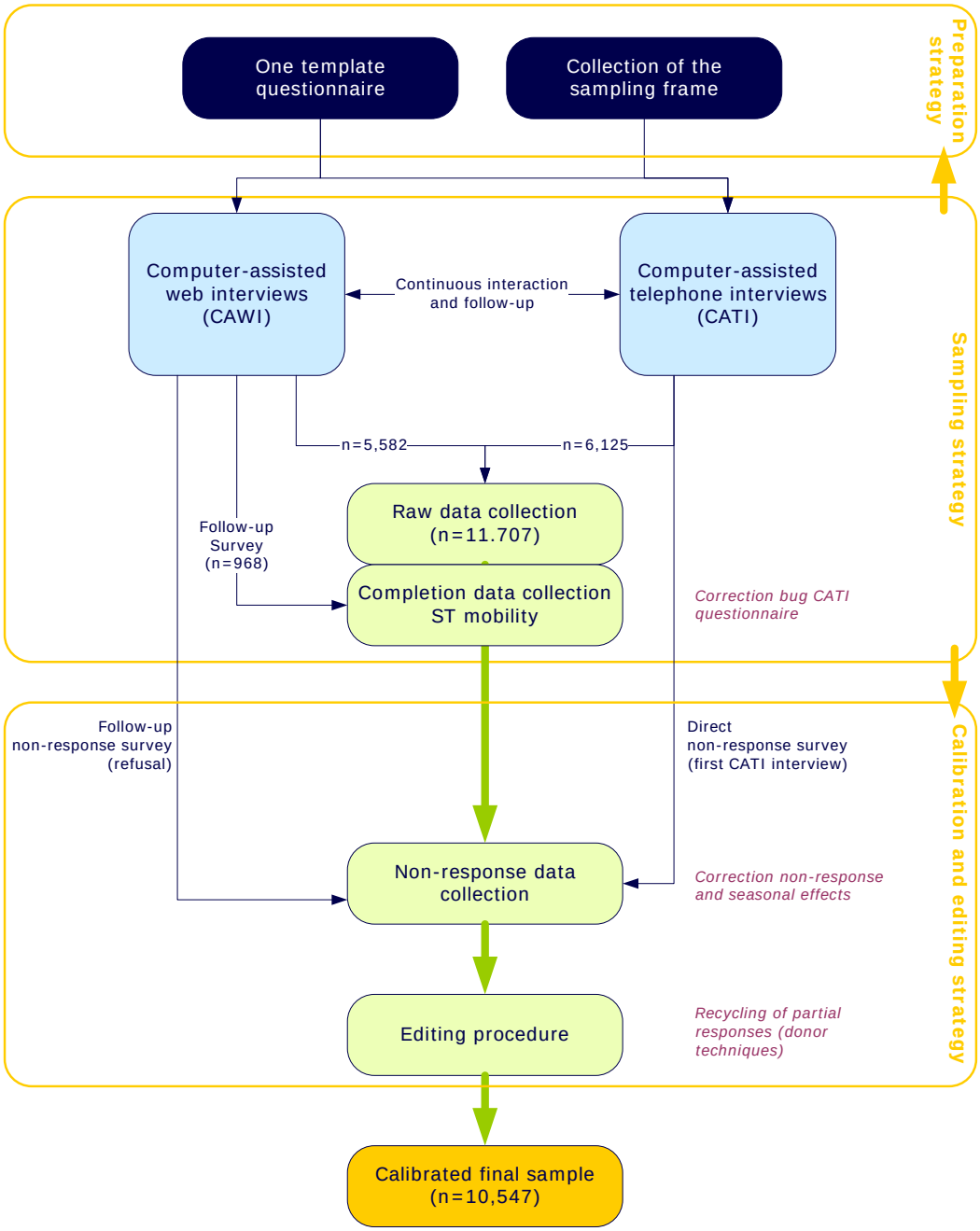
The MORE2 survey was carried out in 2012 under the aegis of the MORE2 project. IDEA Consult organised the multichannel survey combining CATI (computer assisted telephone interview) and CAWI (computer assisted web interview). The project documentation about the project is available online[8]. The process of data collection up to data cleaning and editing is outlined in the figure below.

The data-owner is the European Commission.

[7] For more information, consult MORE2 HEI report, 2012 on which this presentation is based: http://ec.europa.eu/euraxess/pdf/research_policies/more2/Report%20on%20survey%20of%20researchers%20in%20EU%20HEI.pdf.

[8] http://ec.europa.eu/euraxess/pdf/research_policies/more2/Final%20report.pdf

Figure 1: Data collection process



Source: IDEA Consult, MORE2 HEI report (2012)

b. How are the data processed in terms of data cleaning (e.g. harmonisation of organization names, etc.):

The dataset have been cleaned. Weights are included for the population and for the strata. The MORE HEI report (2012) describes the data and characteristics of the respondents. A methodological report with details on the sampling strategy and data was delivered, but is not publicly available. HEI names have been harmonized with ETER as indicated in section 5 of this report and further developed in the Activity Sheet 2.4.

2.3 Information on all variables/indicators

a. Description of all variables and/or indicators:

The variables of the MORE2 dataset cover the following topics (in **bold** information that is new or more elaborated in MORE2 compared to MORE1):

MORE1	MORE2
6. The researcher (country of birth, citizenship(s), gender, age, children),	1. Socio-demographics
7. Education (degrees, graduation year, country, field of highest degree)	2. Education 3. PhD and **doctoral training**
8. Current position (university/college, faculty, field, position level, seniority)	4. Current employment and **working conditions** (including contract, status, satisfaction, inter- and transdisciplinary mobility)
9. Mobility events (up to five mobility events, countr(ies), duration, type)	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
10. Assessment of mobility among mobile as distinct from non-mobile researchers: c. Detailed focus on most recent mobility event (motivations, push and pull factors, assessment) d. Plans/aspiration to work in another country: (country, rationale and background for choice of destination)	
	8. Awareness of EU policy 9. Comparison research environments (EU – non-EU; EU countries)

Next to the raw dataset with variables, one of the MORE2 deliverables is an indicator set, accessible online[9] and with the following structure:

Theme	*Subtheme*
HR of researchers	Stock of researchers
	Training phase
Employment situation of researchers	Sector of employment
	Characteristics, position, status, contract
	Work satisfaction
Mobility of researchers - Stock	Geographical mobility - current
	Geographical mobility - long term
	Geographical mobility - short term
	Geographical non-mobility
Mobility of researchers - Motives, Barriers, Effects	Geographical mobility - motives
	Geographical mobility - barriers
	Geographical mobility - effects
Mobility of researchers - Non-geographic mobility	Virtual mobility/Collaboration
	Intersectoral Mobility
	Intersectoral mobility - characteristics, position, status, contract

[9] http://www.more-2.eu/www/index.php?option=com_content&view=article&id=118&Itemid=125

	Intersectoral mobility - work satisfaction

b. Information on the sectorial classifications used (e.g., economic sectors, technological fields, organizations types, etc.), and listing of all categories for each classification scheme:

Next to the key definitions of researchers and fields of science (described under section 2.1 as they define the sampling strategy), other classifications are applied for career stage, type of mobility and reference countries.

Career stages R1 to R4

In order to allow for country comparisons in terms of functions and experience levels, the concept of specific career stages was introduced according to the four career stages outlined and defined in the European Commission's communication "Towards a European Framework for Research Careers"[10]. Researchers in the MORE2 surveys were asked to self-select into one of these four stages.

These four career stages are:

- R1: First Stage Researcher (up to the point of PhD),
- R2: Recognized Researcher (PhD holders or equivalent who are not yet fully independent),
- R3: Established Researcher (researchers who have developed a level of independence) and
- R4: Leading Researcher (researchers leading their research area or field).

More details are provided in Annex 3.

Mobility

Definitions of researchers' mobility used systematically in the indicator descriptions:

- International mobility versus intersectoral mobility:
  Moving to another country versus moving to another sector (though both can occur in the same move)
- PhD mobility versus post-PhD mobility:
  Mobility of researchers enrolled in a PhD programme during their R1 career stage versus
  mobility in any of the following research career stages and, even though the for simplicity selected terminology suggests otherwise, regardless of whether or not the researcher has obtained a PhD.
- PhD degree mobility versus >3 month mobility during PhD:
  Mobility with the purpose of obtaining the PhD in another country versus mobility of three months or more during the PhD while still obtaining the PhD in the home country
- >3 month mobility versus <3 month mobility:
  Mobility with duration of 3 months or more versus mobility with duration of less than 3 months
- Employer mobility:
  Mobility including a change of employer

10 http://ec.europa.eu/euraxess/pdf/research_policies/Towards_a_European_Framework_for_Research_Careers_final.pdf

- Virtual mobility:
  The use of web-based or virtual technology to collaborate internationally (cf. Annex 1 section 2)
- Non-mobility or never-mobile researchers:
  Having never been mobile to another country (not within the last ten years nor before)

Reference countries

In the MORE2 HEI data (2012), different countries of reference are collected:

- Panel country: the country identified as country of current employment during the collection of researcher contact details (before the survey)
- Country of current employment (as identified in the survey)
- Country of PhD: country where the researcher is currently enrolled in a PhD programme or has previously obtained his or her PhD.
- Country of citizenship
- Country of residence
- …

The panel country is the only one collected in preparation of the survey to be one of the dimensions in the stratified sampling strategy. Therefore, unless otherwise indicated in the description of the indicators, panel country is applied as reference and the indicators referring to panel country are accurate at country level.

Other references are sometimes applied, depending on the specific indicator. It is worth mentioning that the panel country and country of current employment are 98.4% of the cases the same.

c. Information on the temporal coverage used (e.g. annual data from 1990-2010, etc.):

The survey was carried out in the period May-July 2012. For 'researcher mobility' variables, the reference period was either the whole researcher career or the last three years 2009-2012.

d. Information on the geographical coverage and classifications used (e.g. EU-27 member states, regional breakdown using NUTS classification revision 2010, etc.):

- 27 EU Member States
- Candidate Countries: Croatia, Turkey, the former Yugoslav Republic of Macedonia
- Associate Countries: Norway, Switzerland, Iceland

2.4 Quality and accuracy of data

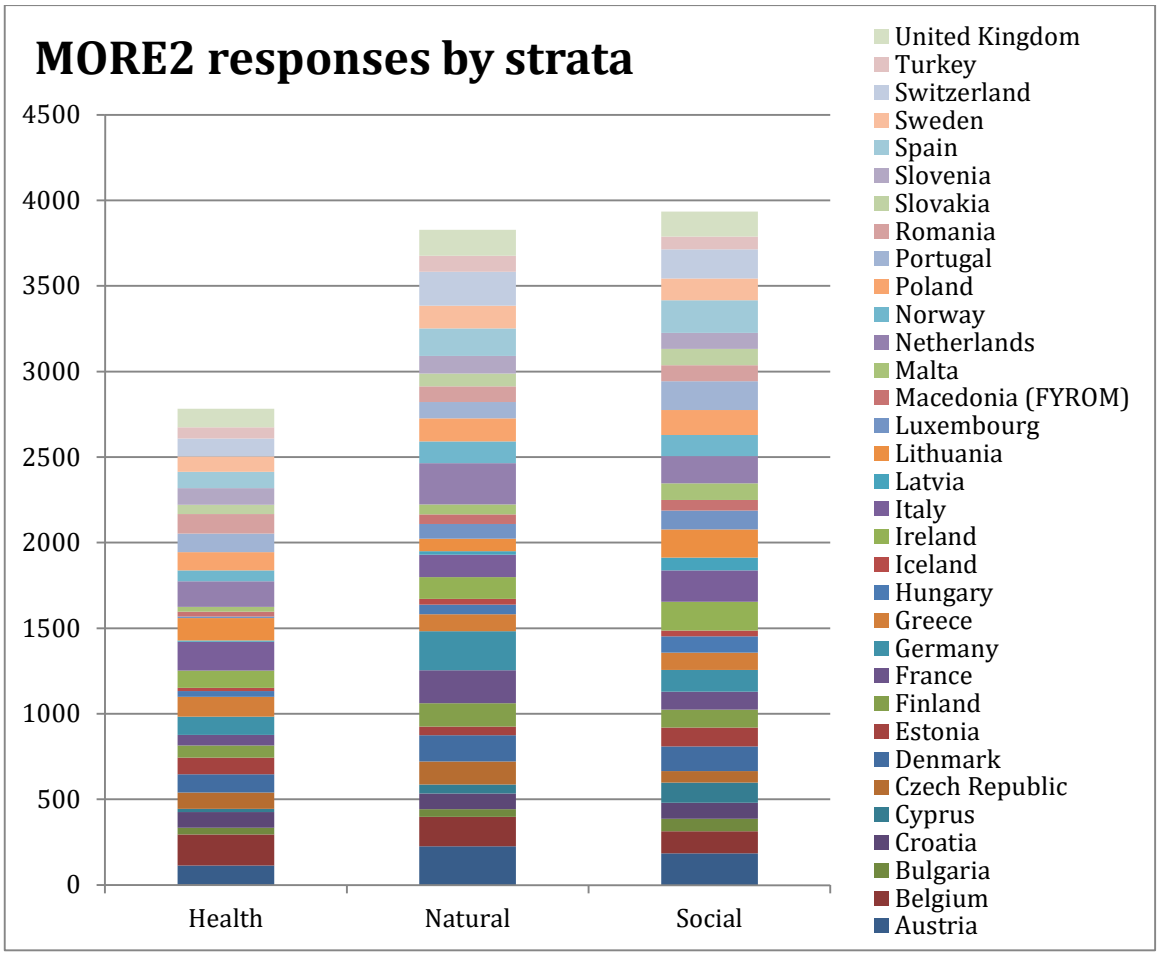
In general, the sampling and data collection strategy were designed to generate high-quality and accurate data. The aims was to provide representative data **at country level**, with a maximum error of 5% at a probability of 95%. This objective was reached in most countries.

a. Information on the number of missing values: details of the survey including the extent of missing values is provided in the project work.

Only complete responses were included in the final dataset. In part thanks to the CATI method, the number of partial responses was relatively low (less than 7.5%). Those partial responses with high degree of completion were edited using donor techniques, which allowed 97 responses to be saved for the final dataset.

The responses per strata are shown in Figure 3.

Figure 3. Distribution of responses by strata (country x field of science).



Source: IDEA Consult based on MORE2 HEI data (2012)

b. Estimation of data quality issues with respect to data acquisition, reliability of retrieving system:

Bug in <3 month mobility question

The combined CATI and CAWI data collection process lowered the risk of partial responses. However, a bug during the programming of the CATI system created an issue related to the <3 month mobility item that followed a different filtering from the CAWI for a subpopulation of respondents. This issue resulted in the lack of information on 2,787 units, a group too substantial not to have an impact on the analysis of <3 month mobility

indicators. Therefore a supplementary survey was launched during the month of September in order to increase the number of correct questionnaires. The 968 additional filled-in questionnaires were then used, together with the original correctly filled-in questionnaires, to implement donor techniques to 1802 units from a deck of 2720 units. The filtering bug has created an issue that has been partially recovered by the supplementary survey, but results will not be as accurate as when all respondents would have given their own information in the original survey. The editing however does allow preserving the minimum representativity requirements for <3 month mobility.

Response bias or seasonal effects

Whereas the editing for partial responses and on <3 month mobility items was meant to complement the information where missing, the editing for non-responses was 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. For example, if researchers that are willing to participate in a survey entitled 'mobility of researchers' are in turn the most mobile researchers in the population, our estimates of mobility would be higher than for the entire population. By surveying non-respondents and comparing their answering pattern to the original respondents' pattern, it is possible to identify and (to some extent) correct this bias. A similar logic applies in terms of seasonal effects due to timing of the survey just before and over summer, which lowered the response rates and could lead to a potential bias of people responding that were still on job during summer.

A non-response survey through both CATI and CAWI allowed to correct for this kind of bias. The estimates were calibrated based on this additional information.

Coverage

The return rate (including also partial responses) was 23%, the response rate (including only the responses used in the analysis) was 21.5%. Only in a few countries, estimates of missing information in the population or concentration of linguistic issues formed a risk to/lowered the accuracy of the data (Iceland, Latvia and Bulgaria). The maximal sampling errors per country are given in Table 1.

Table 1: Max Sampling errors in HC and FTE by country

country	HC
Austria	4.2
Belgium	4.4
Bulgaria	7.7
Croatia	5.8
Cyprus	6.4
Czech Republic	5.6
Denmark	4.8
Estonia	5.9
Finland	5.5
France	5.1
Germany	4.6
Greece	5.5
Hungary	7.2
Iceland	10.1
Ireland	4.9
Italy	4.4

Latvia	9.6
Lithuania	5.0
Luxembourg	5.4
Macedonia (FYROM)	7.6
Malta	6.1
Netherlands	4.1
Norway	5.5
Poland	5.0
Portugal	5.1
Romania	5.6
Slovakia	6.5
Slovenia	5.5
Spain	4.6
Sweden	5.2
Switzerland	4.5
Turkey	6.4
United Kingdom	4.8

Source: IDEA Consult, MORE2 HEI report (2012)

3 Legal issues encountered and access conditions

a. Legal issues concerning access of the database: Access to the micro-data is currently restricted to the MORE2 project group.
b. Owner of raw data (at the time of contract): European Commission DG Research (Directorate B – European Research Area)
c. Current practice for opening up of the database to external users: External users are directed to the project team-leader (IDEA Consult) and the Commission Scientific Officer (Peter Whitten)
d. Legal necessities for potential opening procedures: Permission from the data-owner (EU Commission) has been promised. Reference is made to the European Commission and the MORE2 study and consistency in terms of weighting methodology is guaranteed. Further clearance from the respective part of the Commission may be necessary.

Steps to ensure opening have already been made. We have got a green light from the Commission about the opening of MORE 1 and MORE 2.

4 Technical structure of *MORE2 (HEI)*

4.1 Information on the data base system

The data are stored at IDEA Consult in Excel and Stata (12) format as one integrated table. The weighting, clustering, and stratification of the survey design is described in a separate protocol.

4.2 Technical variable definition

a. The variables are named according to the numbering of the question. Information about the formatting are found in the Annex (below)
b. Data type of all variables: The data are saved in the appropriate format (e.g. byte for Likert scale). There are nine different types of information based on the 87

questions of the full survey. These nine types of information correspond to a large extent to the 5 types identified in MORE1 and are linked to the conceptual framework of the MORE2 study in the table below.

c. Current usage and definition of unique identifiers: The unique identifier currently used is a numeric value ("response") that has been associated with the response. A time-stamp is not associated. An overview of variables is provided in the Annex.

MORE2 survey structure	**Conceptual framework**
1. Socio-demographics 2. Education 3. PhD and doctoral training	Human resources of researchers • 'Stocks' of researchers • HRST, Scientists and Engineers, R&D personnel • Researchers in their training phase • Researchers working in the HEI sector in the EU • Researchers who have moved from the EU to non-EU countries
4. Current employment and working conditions (including inter- and transdisciplinary mobility)	Employment situation of researchers • Employment sector • Characteristics of employment contract • Position/status of the researcher • Contractual status • Work satisfaction in terms of different aspects of researchers' career
	Research environment as an attractiveness factor for researchers
5. Academic mobility and career paths (including PhD, including past and current 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	Mobility of researchers • Stocks of mobility o International mobility intra-EU o International mobility extra-EU o Intersectoral mobility o Virtual mobility • Flows of mobility o Quantification of movements • Influencing factors of mobility • Motivations for mobility • Effects of mobility
8. Awareness of EU policy 9. Comparison research environments (EU – non-EU; EU countries)	Research environment as an attractiveness factor for researchers

4.3 Description of the Entity Relationship Model of *MORE*

a. Definition of single tables: The dataset currently consists of 1 single table.

4.4 Interfaces for access and to other infrastructures

As with the MORE I dataset, the MORE2 survey results can be further enriched and complemented by linking them to each other and potentially to other datasets.

5 Further planning of the opening of *MORE*

There are currently two finalized survey in the MORE family. These are provided in the RISIS infrastructure. We emphasize again that MORE2 cannot be compare with MORE1 due to differences in the survey and sample designs. A third round (MORE3) is currently in the field: it is possible that this will be added to the RISIS infrastructure.

a. Document concrete steps towards opening of the respective dataset

In order to open the MORE2 dataset for on-site as well as on on-line visits, a number of steps must be made to accommodate visitors at NIFU. The steps are:

- Definition of basic terms and principles
- guidelines and terms of use of the dataset,
- Rules and procedures governing the access-infrastructure, including the definition of access conditions for RISIS researchers
- Preparation of the dataset
- Documentation of the dataset
- Tests and remedial efforts to address potential coverage and non-response bias
- Tests and remedial efforts to address potential design and misspecification effects.
- Integration with relevant activities, including linking information from ETER (see also below)
- The technical implementation for dissemination.

A first step will involve making the MORE2 accessible through NIFU and in-house. The next step is to make it available on-line. In subsequent steps, the infrastructure that will allow combined analysis with the previous MORE I data will be explored. MORE I and MORE2 are standalone datasets: integration of the datasets will be made where consistency of the survey design of the two surveys allows.

b. Necessary updates and/or technical changes

Together with MORE1, the sampling designs of the two studies were compared to assess the best form and modes of access to the datasets. We have consulted with the data-owner during this process. It was made clear that the differences in the two datasets are too large to allow them to be compared**. The data-owner (the EU Commission) and the responsible party for collecting the data (IDEA Consult) therefore strongly advise against comparing MORE2 with MORE1.**

The Commission has given RISIS the right to open the two rounds of MORE. To do so, we have worked with the responsible party for collecting the data in both cases (IDEA Consult).

Annex 1: Labels and values

Information on all variables/indicators

Structure:

MORE2
1. Socio-demographics
2. Education 3. PhD and doctoral training
4. Current employment and working conditions (including contract, status, satisfaction, inter- and transdisciplinary 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)

Annex 2: Classification career stages

According to the definitions given in the EC's communication Towards a European Framework for Research Careers"[11] the different stages are characterized as follows:

A first stage researcher (R1) will:

- "Carry out research under supervision;
- Have the ambition to develop knowledge of research methodologies and discipline;
- Have demonstrated a good understanding of a field of study;
- Have demonstrated the ability to produce data under supervision;
- Be capable of critical analysis, evaluation and synthesis of new and complex ideas and
- Be able to explain the outcome of research and value thereof to research colleagues."

Recognized researchers (R2) are PhD holders or researchers with an equivalent level of experience and competence who have not yet established a significant level of independence. In addition to the characteristics assigned to the profile of a first stage researcher a recognized researcher:

- "Has demonstrated a systematic understanding of a field of study and mastery of research associated with that field
- Has demonstrated the ability to conceive, design, implement and adapt a substantial program of research with integrity
- Has made a contribution through original research that extends the frontier of knowledge by developing a substantial body of work, innovation or application. This could merit national or international refereed publication or patent.
- Demonstrates critical analysis, evaluation and synthesis of new and complex ideas.
- Can communicate with his peers - be able to explain the outcome of his research and value thereof to the research community.
- Takes ownership for and manages own career progression, sets realistic and achievable career goals, identifies and develops ways to improve employability.
- Co-authors papers at workshop and conferences."

An established Researcher (R3) has developed a level of independence and, in addition to the characteristics assigned to the profile of a recognized researcher:

- "Has an established reputation based on research excellence in his field.
- Makes a positive contribution to the development of knowledge, research and development through co-operations and collaborations.
- Identifies research problems and opportunities within his area of expertise Identifies appropriate research methodologies and approaches.
- Conducts research independently which advances a research agenda.
- Can take the lead in executing collaborative research projects in cooperation with colleagues and project partners.
- Publishes papers as lead author, organizes workshops or conference sessions."

---

11 http://ec.europa.eu/euraxess/pdf/research_policies/Towards_a_European_Framework_for_Research_Careers_final.pdf

A leading researcher (R4) leads research in his area or field. He or she leads a team or a research group or is head of an industry R&D laboratory. “In particular disciplines as an exception, leading researchers may include individuals who operate as lone researchers.” A leading researcher, in addition to the characteristics assigned to the profile of an established researcher:

- “Has an international reputation based on research excellence in their field.
- Demonstrates critical judgment in the identification and execution of research activities.
- Makes a substantial contribution (breakthroughs) to their research field or spanning multiple areas.
- Develops a strategic vision on the future of the research field.
- Recognizes the broader implications and applications of their research.
- Publishes and presents influential papers and books, serves on workshop and conference organizing committees and delivers invited talks”

Annex 3. MORE2 Protocol

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[12]. 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:[13]

- 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)

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).

---

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

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

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[14]): 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. 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).

[14] 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.

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.