RISIS

RESEARCH INFRASTRUCTURE FOR SCIENCE AND INNOVATION POLICY STUDIES

DOCUMENTATION OF RISIS DATASETS

Mobility Survey of the Higher Education Sector- MORE 3

Asgeir Skelholt, Eric Iversen - NIFU

Outline

1   Introduction

MORE (Mobility Survey of the Higher Education Sector) is arguably the most comprehensive empirical study of researcher mobility available. This document introduces key dimensions of the latest member of the MORE family, the MORE3 survey of Higher Education Institutions (HEIs). It is aimed at researchers who intend to use the MORE3 dataseteither alone or together with the MORE2 datasetto study some aspect of the 'mobility and career paths of researchers in Europe'. The MORE surveys were carried out by a team lead by IDEA Consult (Belgium).[1] The data owner is the European Commission.

From 2019, the RISIS infrastructure provides access to the MORE3 on behalf of the data-owner. Access to earlier rounds of data are available for relevant projects.

The purpose of the current document is to provide a concise overview of the dataset and to point out where the reader may find more information that will help to use the survey results to help improve our understanding of Mobility and Career Paths of researchers in Europe. The following sections will answer questions about what the dataset contains, how it is put together, and where further information can be found in the extensive documentation that the MORE3 Team elaborated. As such, it builds uponand complementssimilar documents that are available for MORE1 and MORE2.

2   Background and basic characteristics

Name: Mobility Survey of the Higher Education Sector: Mobility and Career Paths of Researchers in Europe, 2016 (MORE3)

The MORE3 survey was designed to 'update, improve and develop the indicators of the MORE2 study'. It should be seen in terms of the EU Commission's aim to provide an empirical basis that can reliably be used to assess the impact of policy measures to open labor market for researchers'. To understand the basic characteristics of the MORE3 dataset, this document starts by providing some necessary background.

The MORE dataset consists of three independent waves of extensive, Europe-wide surveys: the pan-European MORE3 and MORE2 survey as well as its more limited predecessor MORE1. The MORE1 survey is not consistent with MORE2 and MORE3 and can be seen more as a pilot study for the subsequent rounds. MORE2 and MORE3 are based on the same basic design and constitute the main studies. The latest two waves share the same sampling design and provide comparable results despite the fact that the questionnaire has evolved in the third wave and despite the fact that this is NOT a panel study. As a consequence, the MORE documents stress that the latest two rounds provide the basis to study the evolution of main indicators at the aggregate level across two points in time but that they do not provide the basis to analyze the evolution of small subgroups (see Interim report, p 32).

The three surveys were carried out on behalf of the EU Commission among large samples of researchers at European universities in 2009, 2012 and 2016, respectively. Each wave consisted of a main survey (the "EU HE Survey" provided here) and a secondary survey on a special topic.

---

[1] IDEA Consult worked together with WIFO (AT) and Technopolis (BE) in setting up and presenting MORE3.

The RISIS focuses only on primary surveys among HES researchers in Europe, where it primarily focuses on the EU28 countries but where it also includes associated countries.

MORE provides a unique lens on mobility patterns and career paths in Europe. Based on the three rounds currently available:

MORE1 yielded 4,538 valid responses from researchers in EU27 countries (2009)

MORE2 yielded 10,547 valid responses from researchers in 33 European countries (EU28 plus candidate and EEA countries). An estimated population frame of 1.241 million researchers

MORE3 yielded 10,394 valid responses from researchers in 31 European countries (EU28 plus Switzerland, Iceland and Norway). An estimated population frame of 1.373 million researchers for EU28 and 1.439 million in total for the 31 countries.

The latest two waves provide measures of flows of international (and sector) mobility, measures of factors that influence mobility, and measures of effects that can be linked to researcher mobility. The survey responses provide promising avenues for a range of studies to better understand the career development paths of European researchers as well as insights into how careers are changing and what role (geographic and sectoral) mobility is playing in each. This links into a range of issues that EU policy has focused on[2]. The policy background as well as the academic discourse that underpins it, are laid out well in the MORE study documents (see references). The reader should first peruse this material before using the database.

3 Database content

The MORE3 database can be seen to consist of two elements. The first and most obvious element involves the responses to the questionnaire. To understand the responses, the reader will need to become acquainted with the questions and the way they were posed. A truncated version of the English version of the questionnaire is found in the Annex: the scroll-downs for years and for countries have been removed, as has piping and other formatting.

This survey instrument includes a total of 106 questions which, according to the project report (see 2016a), is intended to cover seven main dimensions of research careers and mobility:

1. Human resources: numbers and training

2. Career paths

3. Working conditions

4. Mobility and collaboration:

5. International mobility and collaboration

6. Interdisciplinary mobility and collaboration

7. Intersectoral mobility and collaboration

A second important element of the database involves the underlying sampling design. We find this information in the form of a set of weights associated with the responses. This element is crucial if we are to interpret the responses correctly against the population. In this section we briefly review these two constituent elements in more detail.

---

[2] See European Commissions communication Towards a European Framework for Research Careers (European Commission 2011; A renewed EU agenda for higher education COM(2017) 247 final. Also relevant is The EU Education and Training Monitor.

3.1 Definition and description of observations

The survey covers a representative sample of researchers currently working at institutions of higher education in the EU. It targets researchers at HES and includes a definition that is designed to exclude non-researching employees from responding to the survey. The definition is informed by the Frascati definition and is the same definition used in the previous waves of the MORE survey. The box illustrates how the MORE team framed the inclusion question.

We specifically target researchers (including doctoral candidates) 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: In case you do not fall into one of the above categories, the survey will not be relevant to you.

The respondent was then asked to tick one of the following boxes:
1. consider myself a researcher
2. do not fall into one of the above categories.

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 Commissions communication Towards a European Framework for Research Careers[3]. Researchers in the MORE2 and MORE3 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).

An important innovation of the MORE3 dataset is that the respondents reported when they experienced moving from one career stage to the next during their career. A full definition of the Career Stage Classification is found in the Annex.

Questionnaire design

The conceptual framework of the MORE study is presented in report (2016: 18ff) with reference to earlier waves. The MORE3 questionnaire extends and improves on the MORE2 instrument in a few ways. For example, it includes more self-reporting information about progress through

---

[3] http://ec.europa.eu/euraxess/pdf/research_policies/Towards_a_European_Framework_for_Research_Careers_final.pdf

career stages. The following box provides a notional comparison of the evolution of the types of information found in the MORE surveys. Points in bold indicate new or substantially different information in MORE3 compared to earlier waves):

MORE1	MORE2	MORE3
1. The researcher (country of birth, citizenship(s), gender, age, children),	1. Socio-demographics	1. Socio-demographics
2. Education (degrees, graduation year, country, field of highest degree)	2. Education 3. PhD and **doctoral training**	2. Education 3. PhD **and doctoral training** a. **Experience of current PhD students**

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. Current employment activities and working conditions
		    a. **Previous career stages, timing**
		    b. **Recruitment, work placement, internships, skills, dual positions**

4. Mobility events (up to five mobility events, countr(ies), duration, type) 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)	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. 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)	8. Awareness of EU policy and **EURAXESS** 9. Comparison research environments (EU non-EU; EU countries)

3.2 Data collection, sampling and weighting

The sample design of MORE3 and the data editing approach both follow the approach of MORE2 very closely. The approach is designed to provide best estimates on researchers in the HE-sector at both the EU and the country level for the EU28+ associated countries. MORE3 included three associated countries (Switzerland, Norway, and Iceland) while MORE2 operated with 6. In other words, three countries were dropped from the exercise.

Due to improvements during data-collection, the MORE team concludes that the MORE3 results are better than the last round (see 2016:32). Data for the MORE3 survey was collected using CAWI (Computer-assisted web interviewing) and, to a lesser degree, CATI (Computer-assisted telephone interviewing) techniques.

RISIS
RESEARCH INFRASTRUCTURE FOR SCIENCE
AND INNOVATION POLICY STUDIES

This project is funded by the European Union
under Horizon2020 Research and Innovation
Programme Grant Agreement n°824091

Researcher population

The population of researchers, which determines the frame of the study, was estimated using Eurostats headcounts, supplemented where necessary by estimates (based on earlier years or from other statistics). The study estimated that the population frame to consist of 1.373 million researchers. 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 if/when the data-owner provides this information. These can be used to improve the accuracy of the responses via post-stratification techniques.

MORE3 and MORE2 both employed a two-stage stratified random sampling strategy. The study lays out the sampling design (see Annex for details). Based on this material, we note that the sampling matrix consisted of 93 final cluster strata: i.e. 31 countries x 3 Fields of Science (FOS). An individual Higher Education Institution (university) makes up the individual cluster, from which at least a minimum number of researchers were surveyed.

The survey employs stratified random sampling based on FOS within countries so as to ensure that the sample of researchers is allocated proportionally to the FOS in each country: this makes estimates consistent with the country number of researchers in each FOS. However, there are also poststratification based on gender and career stage. It is important that the available weights are used to get representative results. Please see the attached Guidelines for the data analysis of the EU HE survey data. The sampling methods used in the survey are elaborated in the reporting from the MORE3 study project.[4]

3.3   Coverage of MORE3

The MORE3 EU HE survey was carried out in 2016. The survey was administered in 31 European countries: the 28 Member States of the European Union and Iceland, Switzerland and Norway. Three candidate countries that were covered in MORE2, were not included in this wave: Croatia, Turkey, the former Yugoslav Republic of Macedonia. It should also be noted that specific questions were directed to foreign nationals currently working at European HEIs (see Questions 45ff). A separate survey (not included here) also attempted to address a similar survey to European nationals working in non-European countries. The focal questions on episodes of mobility (e.g. Q64) generally have a 10-year reference period.

In addition to country, the survey is designed to provide correct representations of fields of science. 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. MORE3 follows the same criteria as MORE1 and MORE2 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)

---

[4] https://www.more3.eu/deliverables, see page 316 in
https://cdn1.euraxess.org/sites/default/files/policy_library/survey_on_researchers_in_european_higher_education_in
stitutions.pdf

- 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)

3.4 Quality and accuracy of data

The measures presented in this database are based on the self-reporting of a sample. The sample is limited in size (the 10,394 responses of MORE3 represent a tenth of a percent of the total population of researchers in Europe in 2016) but it is targeted. The way in which it is targeted is important to understand to appreciate how accurate the measures are likely to be and at which level they are most reliable.

The basic issues are laid out in the MORE documentation. In particular, the MORE3 Second Interim Report provides important information about how the survey was administered and how that may affect the reliability of results. The database user should specifically look at Section 4 ("Intepretations of the results") as well as Annex 1 to that report.

Sampling Design

Off the top, it should be noted that the measures are designed to estimate as closely as possible the true value of the various measures it provides (primarily, the share of different types of mobility) at the time of the survey. The true value can only be estimated. In this section, we review different types of error that can affect the accuracy of the database and how they are dealt with.

The first observation is that the accuracy of measures can differ depending on what level responses are aggregated at. This follows from the way the survey was targeted. In short, the estimates are most accurate when they are presented at the country level (or at the level of the EU as a whole). Using Eurostat data as the population frame, the MORE study (specifically, MORE3 and MORE2 waves) was designed to maximize the accuracy of estimates at the level of the country (and for EU countries as a whole) where field of science was also a target.

This also means that estimates at other levels of aggregation may be less accurate. In particular, this will likely affect estimates made at the level of gender and, more so, at the level of career stages where reliable population frames do not already exist for all countries. The MORE3 study however attempts to reconcile ex ante and ex post stratification procedures to circle in on best estimates. The MORE3 Second Interim Report provides some guidance here (see 2016: 30ff).[5]

Sampling error

The MORE3 team designed the sampling frame so that the country-level estimates would be as accurate as possible. In practice, they operated with a margin of error of 5% at the 95% confidence level (see 2016: 318). In other words, the estimates for the major questions were

---

[5] See chapter 4 Interpretations of the results and the Annex*.

intended to accurately reflect the real value (within this margin of error). This means that the measure of the rate of long-term mobility (based on Q64), for example, was intended to accurately reflect the real value at the level of the country (within this margin of error).

The MORE3 team reports that the sampling error is better (lower and more uniform) than in MORE2. The error differs however from country to country, with the highest value above 6.5% for a single country. The survey provides individual weights to improve the accuracy based either on the sampling design or on the calculation of non-response depending on the focus. In addition to ensuring accuracy at the level of the country, the dataset provides other weights to make sure that the responses accurately represent other levels of analysis (e.g. field of science, gender, career stages, FTE).

Table 1 provides an overview of the seven weights that are furnished in the dataset. The dataset includes sampling weights for four different analytical levels: field of science, gender, career stages, and FTE. It should however be noted that levels of analysis that focus on these levels are likely to be less accurate than those conducted at the country or the EU levels. In general, the MORE3 team does not guarantee the same accuracy of the sampling weights for these extra levels of analysis as they do for estimates at the country level (2016: 31).

Table 1: The seven weights used in the MORE3 study

Weights	Name	Description
1	weiFOS	sampling weight that differs per country and per field of science
2	weisex	weight differs per country and per gender
3	weifteFOS	as weifos with full-time/part-time is taken into account.
4	weiCS	sampling weight that differs per country and per career stage
5	calwq64	non-response bias-robust sampling weights q64
6	calwq79	non-response bias-robust sampling weights q79
7	calwq88	non-response bias-robust sampling weights q88

Career stage weights are a specific concern, since this is where current population frames provide least guidance. The errors of estimates for career stage will be substantially higher. Particularly if there were few responses in the strata (career stage within countries). The interested user will find a comparison of results based on ex ante and ex post weights in section 1.4 of the Annex (2016).

The dataset also provides a second type of weight that are directly connected to the three types of mobility that MORE focuses on: long-term (over 3 months), short-term (three months or less), and sector-mobility (other than HES). These weights seek to correct for the possible bias of non-respondents (unit non-response). These weights are calibrated on the basis of a non-response study that was carried out among a subset of researchers who did not respond to the original study.

There are some limitations to keep in mind in interpreting the Career Stage weight (weiCS). Some countries (Estonia, Cyprus and Luxembourg) do not publish career stage statistics, while others (Malta and Slovakia) only provide annual statistics about R1 researchers. As a result, the MORE3 Team has not been able to calculate post-stratification weights for careers that are not covered for these countries. This means that the aggregate career stage stratified indicator is only available for 25 of the 28 EU Member States.

Item Non-Response

The accuracy of the dataset is also affected by incomplete responses. In addition to unit non-response, the incidence of item non-response can be a problem. MORE3 reports that it has sought to impute values for some measures using a donor method (see documentation) but that this only involved about 2 percent of the responses.

In addition, the user will find that the (Stata version) dataset distinguishes between two types of missing values: plain missing values are coded differently than missing values attributable to responses of the type not applicable. The number of missing will vary between different variables. Plain missing are treated as .a in the dataset, while values of .b indicate not applicable/available as defined by the MORE3 team. In the latter case, a label is provided about the type of missing value.

4 Technical Specifications

4.1 Information on the data base system

The dataset was acquired as a flat file from the EU Commission via the Era More Studies Team. The data was then cleaned, labelled and set up using the commonly used Stata package. The dataset is available either as a flat file (comma separated) or as a Stata file. The Stata version is recommended as it maintains the labels and provides other useful formatting (e.g. for missing).

By default, the dataset includes the latest (MORE3) wave of the MORE survey. Data from MORE 1 and MORE2 are also available and can be provided upon request. It is important to emphasize that due to different sampling strategies between MORE 1 and MORE 2, it is not possible to compare MORE 1 with MORE 2 or MORE 3. It is, on the other hand, possible to compare MORE 2 and MORE 3 noting the different country coverage and the differing questions. The project documentation (see MORE3 Second Interim Report, EU Higher Education survey results (D2 part 1). pp 30ff:) provides further details about comparing MORE3 and MORE2.

4.2 Technical variable definition

The database includes the responses from the 106 main questions plus the identifier and the 7 weights. Some questions consist of more than one part, and their responses will therefore be found in more than one variable. As a result of this, there is a total of about 550 variables in the MORE3 dataset.

All variables have been labelled according to their formulation in the questionnaire (see Annex below). Due to restrictions in character length, no variable label is more than 80 characters. Please use the questionnaire for the complete wording for each question. Most variables are by default provided as float and given labels if applicable. However, some open-ended question are string variables.

5   Scientific use

Data from the MORE studies have been around for almost a decade. There has thus emerged a

---

Box: Differences in earlier waves of the MORE surveys

The differences in coverage between MORE3 and MORE2 have been spelled out in more detail in this report.   The differences between MORE2 and the pilot effort of MORE1 are, as noted, more formidable. Although the gist of the first wave was broadly similar to MORE2, the two waves are incompatible due to differences in the coverage of the two waves, in the sample design, and in the questionnaire itself. These issues are spelled out in the report for MORE2.

The responses to MORE1 (4550) were significantly smaller for the EU than MORE2 (9440) and more skewed towards Natural Sciences, with a relative underrepresentation of the Health area. In addition, there were significant problems in responses for individual countries such as France. Improvements in the population frame that came on line is one factor that helped to improve the design here.

---

range of studies that utilize MORE data. This material can be used to identify interesting and relevant use cases.   In closing we emphasize again that there are differences between the different waves and that only MORE2 and MORE3 are compatible for research purposes. See the box for information about differences with MORE1.

From 2019, the RISIS infrastructure provides access to the MORE3 on behalf of the data-owner. Access to earlier rounds of data are available for relevant projects. In putting together a request, the potential user should become acquainted with the extensive material already available in the scientific literature.

The MORE3 project team has generated extensive documentation. This documentation refers to examples of scientific output that build on the MORE family of datasets. The use cases noted in this material particularly relates to developments in EU policy.

Therefore, a good starting point to develop new use cases or to improve on existing ones is to be found in the project reports. See https://www.more3.eu/deliverables .

References

Jьrgen Janger & Agnes Kьgler & Kathrin Hofmann & Anna Strauss & Fabian Unterlass & Miriam Van Hoed & Annelies Wastyn & Lidia Nucez Lopez, 2017. "MORE3 Support Data Collection and Analysis Concerning Mobility Patterns and Career Paths of Researchers. EU Higher Education Survey Results," WIFO Studies, WIFO, number 60982, August.

Jьrgen Janger & Agnes Kьgler & Kathrin Hofmann & Anna Strauss & Annelies Wastyn & Lidia Nucez Lopez, 2017. "MORE3 Support Data Collection and Analysis Concerning Mobility Patterns and Career Paths of Researchers. Global Survey Results," WIFO Studies, WIFO, number 60983, August

Annex 1: Classification career stages

According to the definitions given in the ECs communication Towards a European Framework for Research Careers[6] 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.

---

[6] http://ec.europa.eu/euraxess/pdf/research_policies/Towards_a_European_Framework_for_Research_Careers_final.pdf

- 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.

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