Plan Overview

A Data Management Plan created using DMPonline

Title: Kvinnor inom vetenskap och teknik: karriärhinder och kostnader för barn

Creator: Olof Ejermo

Principal Investigator: Olof Ejermo

Data Manager: Peter Holmström, Olof Ejermo

Affiliation: Lund University

Funder: Forte (Swedish Research Council for Health, Working Life and Welfare)

Template: LU standard template for data management plans

ORCID iD: 0000-0002-4604-5912

Project abstract:

Women face difficult trade-offs in juggling the demands of working life and family life. The Nordic countries have responded to this through the introduction of ‘family-friendly’ policies whereby women could stay on paid parental leave. These policies have contributed to high labor force participation rates for women, and lower so-called
“child penalties” in earnings. There remains, however, potential adverse consequences of parental leave in high-skilled jobs that build upon the continuous accumulation of knowledge: workers suffer from a “burden of
knowledge” in keeping up with an ever-expanding knowledge frontier in their field, and time away from the job can lead to quick depreciation of knowledge. Thus, parental leave may favor women’s participation in the labor market generally, but jobs characterized by rapid knowledge accumulation may suffer.
Our proposed research program address this paradox by investigating the career paths of women and men in two particularly knowledge-intensive areas of key importance to societal progress, science and technology. We
will provide causal evidence on the impact of child-bearing on academic careers and careers in highly innovative private firms and how this impact is related to the roll-out of parental leave schemes. Going beyond the focus on earnings inequalities, we will examine the potential adverse effects on knowledge output in terms of publications and patents, which might have long-lasting, negative implications for societal growth and welfare. Our study also has implications for the determinants and consequences of women’s continued underrepresentation in science
and technology professions, addressing the fairness of access for both genders to high-earning, knowledge-intensive jobs.
We present a proposal for comparative research across the Nordic countries with an international and gender-
diversified team of established researchers and a plan for further educating younger scholars in this important
area.

ID: 144482

Start date: 01-11-2021

End date: 31-10-2027

Last modified: 02-04-2025

Grant number / URL: 2021-01552

Copyright information:

Kvinnor inom vetenskap och teknik: karriärhinder och kostnader för barn

1. Data description

1.1 What method of data collection will you be using?

- collecting material or generate/produce data (e.g. interviews, measurement data)
- using existing data (archival data, data from previous project – your own or other researchers')

Data will be analyzed on Statistics Sweden's database platform MONA, where registry data will be used. To this platform will be added data on inventors (who patent) and scientists (who publish). Data on scientists are taken from an earlier data collection. Inventor data will draw both on earlier collected data and newly collected data. All the collected is only used on the MONA platform.

1.2 Describe how data will be collected, created or reused.

Own reused data referred to above will be reused by linking by Statistics Sweden to their registry data. These data are store on secure servers hosted by LDC.
New collected data on inventors will be assembled by sending data from PATSTAT on inventors to Statistics Sweden where an algorithm will search for the best link with their registries. A copy is stored on a secure server hosted by LDC containing information only on the individual in identified form (the personnummer) and the application number to which the person is linked .

1.3 What type of material (physical or digital) will you use (e.g. text, images, measurement data)? In which file formats will you save your data?

The data on inventors and scientists referred to above are stored in excel files or alternatively, ascii text files

1.4. According to your estimation, how large is the maximum storage capacity will you need throughout the project (primary data and revisions of processed data)?

- <100 GB

2. Documentation and data quality

2.1 How will your data/your material be documented and described with metadata, take collection method, content, structure, standards and formats in consideration; in order for you and other researchers or computer software to read and be able to interpret the data correctly?

No metadata standard has been followed. The collected are documented in two forms. First, a technical document exists which describes the collection and processes undertaken for publication data. Second, the algorithm used for linking inventors with personal id has been published on github.

2.2 How will the quality of the research data be ensured and documented (for example by repeated measurements, data entry validation etc.)?

see above.

3. Storage and backup

3.1 How will you ensure integrity of storage and backup of data and metadata during the research process?

Local IT-support (LDC) has been consulted to set up secure servers. No data will be analyzed locally, all data will be analyzed on Statistics Sweden servers on MONA.

3.2 How is information security and access to data controlled, for example in reference to sensitive data and personal data?

There are no sensitive data involved. Individual data which can be used to identify individuals exist in the dataset.
For inventors, these consist of:
Personnummer, address, patent applicant
For staff lists- publications:
Personnummer, department, faculty, email
Under the construction of these databases project members must have access to personal information in order to identify who-is-who in publications and inventor information. After completion, only the project manager Olof Ejermo can access it.

4. Legal and ethical requirements

4.1 Will the project be processing personal data?

- Yes – in that case you must report this in Pulu (https://pulu.adm.lu.se)

4.2. How will you ensure that data is processed according to the regulations concerning for example personal record handling, confidentiality and intellectual property rights?

Data are stored on secure password-protected servers but they are not anonymized. The identity of individuals are necessary to preserve for research to be undertaken now and in the future at Statistics Sweden systems where all analysis is undertaken.

4.3. In what way will you ensure that data is handled correctly from an ethical standpoint?

No sensitive data from an ethical perspective are used in the project.

4.4. Collaborative research projects involving external parties, may require an agreement between participants/principals of the study regarding processing, storage, ownership and aspects of intellectual property rights. Is this the case in your study?

- Yes, it does involve external parties

At the request of Statistics Sweden LU has Data processing agreements (PUBA) with external partners who use data on their servers for the purpose of the project.

5. Data sharing and long-term preservation

5.1 Will research data and/or information on data (metadata) be made publicly available?

- No, restrictions do exist

Data are not made publically available and the data have not been documented under an official metadata standard.

5.2 If so, how, when and where will data and/or metadata be made available? Are there any limitations (legal and/or ethical) that prevents sharing or reuse of it?

-

5.3 If you plan to make data/metadata publicly available, will you use a unique and persistent identifier (PID) such as a DOI?

-

5.4 If data has been created or collected, is there a reason for keeping these forever or may they be destroyed after 10-20 years? What would the reasons be for preservation?

- My current assessment is that data probably may be destroyed.

6. Responsibilities and resources

6.1 Who is responsible for the data management and assists with the data management during the project? Who is responsible for data management, keeping of records and long-term preservation, after the project finished?

Responsible for all these aspects is the project PI Olof Ejermo.

6.2 What resources (cost, labor and miscellaneous costs) will be allocated to data management (including storage, backup, data sharing and long-time preservation preparation) within the project?

Costs consist mainly of storage (server space) costs which are negligible. Data management costs during the project are paid by PI research time in the project.