



Data Management Plan

Deliverable: D7.3 Data Management Plan

Author: Florence Blandinieres

Version: 1.0

Quality review: Georg Licht

Date: 21 September, 2017

Grant Agreement number: 727073

Starting Date: 01/04/2017

Duration: 24 months

Coordinator: Dr. Georg Licht, ZEW

Email: licht@zew.de



Contents

1	Introduction	2
2	R&D parameters: use within FRAME, access and description of the data	3
2.1	Data description	3
2.2	Access of the data	4
3	Technologies adoption parameters	4
3.1	Description of the data	4
3.2	Access of the data	5
A	Agreement for the exploitation of the Fraunhofer dataset	6

1 Introduction

The purpose of the FRAME Data Management Plan is to provide an overview of the main elements constituting the data management policy in the light of the data sources involved. The Final Data Management Plan will provide a complete overview when most of the scripts and estimations will be available at the end of the project.

The report is organized in two main sections related to the estimations linked to the Work Package 6. This work package is the only one involving empirical data which is crucial to estimate the parameters involved in all of the work packages. The sections describe the name of the dataset, its sources, accessibility and brief description of its use within the FRAME project. The Data Management Plan has been designed to be the most user friendly as possible by providing access to most of the data involved in the calibration of the model (see the D7.2 and D9.1 deliverables for more details) but also to the scripts involved in each work package to study the behaviour of the model. The GitHub depository is available at <https://github.com/H2020FRAME>. The depository will be updated over the submission process of the different deliverables. The scripts will be available for replication after the acceptance of the deliverables and eventually, after the acceptance of submitted articles to scientific journals.

Broadly speaking, the FRAME project will ease the access to the data and scripts for replication within the scientific community. Most of the sources of data is publicly available, either by being in opened access already or by belonging to national statistical offices. Consequently, the use, archiving and, preservation of these datasets follow the rules of the respective institutions. However, due to confidential agreement and copyrights, two micro sources of information will not be available (i.e. PatStat and Fraunhofer Institutes licensing agreements). The latter do not have implications for the replication of the findings because the aggregated estimations to define the calibration of the model will be provided. The archiving and preservation policies of the data and scripts involved in the FRAME project will follow the ZEW Data Protection Policy. Consequently, the latter will be available for replication within 10 years. The access to the different datasets involved within the project is centralized on the website <http://www.h2020frame.eu/frame/scientific-materials/data.html>.

2 R&D parameters: use within FRAME, access and description of the data

The first set of estimations relates to R&D parameters. The first aspect of the WP6 is to disentangle between the sources of R&D conducted. The first deliverable will extend the approach developed by Botazzi and Peri (2007) who estimate the elasticity of the private knowledge stock Z through the total number of patent applications with respect to private R&D on a panel dataset at the country level. Doing so will allow to disentangle between public and private R&D efforts, which in turn, will be used as key elasticity parameters for the Work Package 1. The estimations will be also key parameters within the estimations involved in Work Package 4, in which the results from the baseline model are extended for different countries.

2.1 Data description

The OECD data is available online and covers a wide range of OECD and non-OECD countries. Data is available from 1981 onwards. RD data reported in this dataset have been collected according to 2002 guidelines of the Frascati Manual, which have been now superseded by the Manual's 2015 edition. Revised definitions are not expected to revise significantly the major indicators. Data are provided in million national currency (for the euro zone, pre-EMU euro or EUR), million current PPP USD and million constant USD (2010 prices and PPPs). This table contains research and development (R&D) expenditure statistics. Data include gross domestic R&D expenditure by sector of performance (business enterprise, government, higher education, private non-profit, and total intramural) and by source of funds (business enterprise, government - including public general university funds -, higher education, private non-profit and funds from abroad - including funds from enterprises and other funds from abroad). For more information, please see <https://stats.oecd.org/Index.aspx?DataSetCode=GERD_FUNDS>.

The Community Innovation Survey (CIS) based innovation statistics are part of the EU science and technology statistics. Surveys are carried out with two years' frequency by EU member states and number of ESS member countries. Compiling CIS data is voluntary to the countries, which means that in different surveys years different countries are involved. The CIS is a survey of innovation activity in enterprises. The harmonised survey is designed to provide information on the innovativeness of sectors by type of enterprises, on the different types of innovation and on various aspects of the development of an innovation, such as the objectives, the sources of information, the public funding, the innovation expenditures etc. The CIS provides statistics broken down by countries, type of innovators, economic activities and size classes. The CIS surveys cover the period 1998-2014 and most of European countries. For more information to get access to the data, http://ec.europa.eu/eurostat/documents/203647/771732/How_to_apply_for_microdata_access.pdf.

PATSTAT contains bibliographical and legal status patent data from leading industrialised and developing countries. This is extracted from the EPO's databases and is provided as raw data or online. The PATSTAT product line forms a unique basis for conducting sophisticated analyses of bibliographic and legal status data. It has become a standard in the field of patent intelligence and statistics. The PATSTAT product line consists of three individual databases. They are available in raw data format or via PATSTAT Online, a web-based interface to the databases. With PATSTAT Online, you can run queries in the databases, conduct statistical analyses, visualise the data and download it for offline use. For the conditions related

to different access modalities, see <https://www.epo.org/searching-for-patents/business/patstat.html#tab-2>.

2.2 Access of the data

The estimations will rely on the International Patent Classification and the NACE classification in order to assess the number of patents and R&D efforts at the sector and country levels. To do so, three main databases will be crossed in order to specify sector and country-specific estimates of the public and private R&D efforts: OECD data, CIS data from EUROSTAT, and PatStat. With the exception of PatStat, the other sources of data are available upon request for scientific purposes. The results of the estimations at a higher degree of aggregation (sectorial and country) will be available. The scripts used to define the amount of patents per sector and countries will be downloadable from our GitHub depository.

3 Technologies adoption parameters

The estimation of the adoption parameters relies and extends the current efforts made by Comin and Mestieri (2014). The average speed of adoption will be based on the CHAT dataset and will be involved in Work Packages 1, 2, and 3. The first extension is the assessment of the private investment in adoption. Its importance will be measured by the share of private R&D on GDP, and extended for each sector, and country. The second extension will focus on the role of applied public research efforts on the speed of adoption. The latter will involve the Government Budget Appropriation Of Research and Development (GBAORD) data from the OECD, and the result of a matching procedure between German CIS data and the Fraunhofer data about confidential agreements made between Fraunhofer institutes and firms over 20 years. Doing so allows to disentangle between the respective role of public and private R&D in diffusing and adopting technologies.

3.1 Description of the data

The Cross-country Historical Adoption of Technology (CHAT) dataset is an unbalanced panel with information on the adoption of over 100 technologies in more than 150 countries since 1800. It contains information on the diffusion of about 104 technologies in 161 countries during the last 200 years. It extends the data used in Comin and Hobijn (2004) and Comin, Hobijn, and Rovito (2006). Almost all of our source data is only available at an annual frequency. For some of the older technologies, like steamships, data go back until the early Nineteenth Century. The last year in the sample is 2003. Some data, especially in the earlier part of the sample, is not available at an annual frequency. The technology measures in CHAT capture a similar intuition. They are either: (i) the number of capital goods specifically related to accomplishing particular tasks, (ii) the amounts of particular tasks that have been accomplished, (iii) the number of users of a particular manner to accomplish a task. For more details about CHAT, see <http://www.nber.org/papers/w15319.pdf>.

The last database involved is a collection of the confidential agreements signed between the Fraunhofer institutes and German firms to license or to perform R&D. Fraunhofer Society consists of one of the largest applied research organization in Europe and hence, represents a large portion of German public research disseminated among German firms. Combined with the

CIS data, the sample represents around 50,000 identified firms collaborating with Fraunhofer institutes. This unique dataset allows to measure the importance of public and applied R&D.

3.2 Access of the data

As mentioned above, PATSTAT requires is not freely available but the results of the estimations will be available. The scripts used to request the right set of patents among sectors and countries will be also available for replication. The OECD data (see subsection 2.2) and the CHAT dataset is downloadable from <http://www.nber.org/data/chat>. The only exception to an opened access to the data concerns the micro-data about the confidential agreements between the Fraunhofer institutes and German firms to license or to perform R&D. The FRAME project could benefit from a specific allowance to use the data within the frame of the study only (see Appendix A for more details).

A Agreement for the exploitation of the Fraunhofer dataset

Project Agreement

between the companies and institutions shown below

ZEW, Zentrum für Europäische Wirtschaftsforschung GmbH
L7, 68161 Mannheim, Federal Republic of Germany.....

-hereinafter referred to as "ZEW" -

Prof. Diego Comin

Professor of Economics, Dartmouth College, 6106 Rockefeller Hall, Hanover, NH 03755-3514,
U.S.A.

-hereinafter referred to as "Prof. Comin" -

CIRCLE, Lund University

P.O.Box 117, 22100 Lund, Sweden

-hereinafter referred to as "CIRCLE" -

and

Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e. V.,

Hansastraße 27c, 80686 München, Germany

as legal entity for its

departments for customer and market information systems (A4) and Fraunhofer Academy (A9),

-hereinafter referred to as "FhG" -

ZEW, Prof. Comin, CIRCLE and FhG together hereinafter referred to as "Partners" -

for the joint implementation of the project

"Fraunhofer Impact Study"

080-P-315466/801161-080

B6-Fis

A handwritten signature in blue ink, appearing to be a stylized 'C' or similar character.

WHEREAS, all Parties, for their mutual benefit and with the purpose of identifying areas of common interest for joint collaborative activities, e.g. in the form of a joint research project to estimate the impact that collaborating with Fraunhofer has on the performance of German companies (hereafter referred to as the "Fraunhofer Impact Study"), intended to disclose to each other certain business and technical information and data in the following scientific fields:

- 1 data evaluation, data matching and data analysing
- 2 setting up a longitudinal database that combines information on the firms with information on the extent and timing of their collaboration with Fraunhofer

WHEREAS, all Parties have, in order to define the nature and content of the collaboration regarding the Fraunhofer Impact Study the Parties hereto wished to exchange technical and/or business information and data of a confidential nature presently in their possession and wished to ensure that the same remain confidential.

WHEREAS, all Parties have already signed a Non Disclosure Agreement (version January 20, 2015) at January 21, 30, 2015.

Now, therefore, in order to define the details of the project "Fraunhofer Impact Study", the Parties agree to the following regulations as set forth below:

1 Subject

Subject matter of this Agreement shall be the Partners' co-operation in carrying out the joint project

"Fraunhofer Impact Study"

The type and scope of the work to be performed by each Partner is defined in detail in the work plan attached as Annex A. Every Partner shall be responsible for the implementation of its assigned tasks at its own costs.

2 Designated Representatives

The designated representatives of each Partner for the performance of the joint project "Fraunhofer Impact Study" are listed in Annex B.

3 Duration

The joint project started on January 01, 2015, has an expected period of performance of 18 months and ends when FhG has approved a final version of the Fraunhofer Impact Study. Details are specified in the work plan attached as Annex A.

4 Rights and Obligations

- 4.1 The Partners shall inform each other about the attained research results or the work progress, and shall exchange informations, data and interim and final reports as given in the work plan attached as Annex A.
- 4.2 For the duration and implementation of the project only, the Partners shall grant each other a non-exclusive, non-transferable, non-sublicensable, royalty free right of use to their information, data and any other industrial property rights of the Partners already existing prior to the signature of this Project Agreement (hereinafter referred to as "BACKGROUND"), whereas this right of use is limited insofar as this right of use is necessary for the Partners to perform their work packages according to Annex A. No further rights of use are granted to the other Partners to the BACKGROUND.

Notwithstanding the foregoing the Partners shall keep the BACKGROUND of the other PARTNERS strictly confidential according to section 5 of this Cooperation Agreement. , Regarding any information and/or data of FhG section 6 shall apply in addition.

- 4.3 The Partners agree the result of the joint project "Fraunhofer Impact Study", including but not limited to the final study, (hereinafter referred to as the "RESULTS") shall vest in FhG. ZEW, Prof. Comin, CIRCLE shall transfer right and title to the RESULTS, insofar as legally possible, to FhG and shall keep the RESULTS strictly confidential unless FhG agrees in writing prior to any intended disclosure of any of the other Partners.

The other Partners shall grant to FhG an exclusive, unrestricted in time, place and content, right to use the RESULTS in any form or matter; this right includes the right to copy, distribute, adapt, or modify the RESULTS. If the RESULTS are protected by copyright, FhG is especially entitled to develop and modify the RESULTS in any way and to offer, publish, copy, exhibit, or distribute – also via networks or satellite transmission – the results in their initial or modified form. FhG is entitled to transfer its rights of use and to grant exclusive or non-exclusive licenses – including sublicenses – to third parties to the RESULTS.

5 Confidentiality

The Partners shall keep in confidence for the confidential information of the other Partners according to the Non Disclosure Agreement (version January 20, 2015) at January 21, 30, 2015 which term shall be extended until the end of this project according to section 3 and which is attached as Annex C.

6 Data of FhG

- 6.1 The data of FhG may be used by the other Partners only for the purpose of the joint project "Fraunhofer Impact Study". These data may not be processed or used for any other purpose; in particular commercial or any other business purposes for private or public clients is not permitted.
- 6.2 The Recipient of data of FhG is obliged to delete the data and any backup copies, selected files and auxiliary files at the latest at the end of the project agreement and a representative of the Recipient of data of FhG shall confirm promptly in writing that the data have been deleted.

- 6.3 In the event that the provided data of FhG not used in accordance with this Project Agreement the data Recipient shall be liable for any damage incurred to FhG.

7 Liability

- 7.1 The Partners shall not be liable for the correctness of the research results exchanged or the information communicated during the project. Likewise, the Partners do not warrant that the rights of use granted by them can be executed without infringement of any third party's rights. Such limitation of liability shall not apply in cases of intent.
- 7.2 Unless otherwise stipulated in this Agreement, the Partners shall, including liability for their senior executives, legal representatives and vicarious agents, not be liable for breach of duty or tort except in case of intent. The foregoing limitation of liability shall not apply in cases of violation of any of the obligations stated in section 6.

8 Publications, Advertising

- 8.1 Solely FhG shall be entitled to publish the RESULTS according to the sole discretion of FhG, whereas FhG shall name the author or the other Partners in cases of publications of the final version of the Fraunhofer Impact Study.
- 8.2 ZEW, Prof. Comin or CIRCLE shall ensure that any results of the research that they intend to publish or disseminate otherwise (further publication) do not contain information which may permit the identification of Fraunhofer individual records (such as customer etc.) of the data. To avoid any disclosure of confidential data (customers etc.), Fraunhofer reserves the right to check all preliminary, mid-term and final results (such as estimations, publication iterations etc.) that might be components in planned publications. If ZEW, Prof. Comin or CIRCLE are interested to publish the RESULTS or any parts thereof they are therefore only allowed to do so after the prior written consent of FhG, which shall not be unreasonably withheld.
- 8.3 The Partners may only mention the name of any other Partner for purposes of advertising with the prior express consent of such Partner.

9 Subcontracts

The Partners may conclude subcontracts only with the prior written consent of the other Partners. If - in case of such consent - a Partner assigns some of his tasks under this Project Agreement to a subcontractor this does not affect its own obligations resulting from this Project Agreement. The respective Partner shall secure that the subcontractor will comply with all obligations – especially with regard to confidentiality – resulting from this Agreement and that the RESULTS attained by the subcontractor will be made available to FhG according to Section 4.

10 Termination

Each Partner may terminate its participation in the project for good cause only. There is no right of termination for convenience.


11 Miscellaneous

- 11.1 Ancillary agreements, amendments, additions hereto shall be made in writing. This applies also if the requirement of the written form shall be waived.
- 11.2 This Agreement shall be governed by the laws of the Federal Republic of Germany. Any dispute or claims arising out of or in connection with this Agreement shall be subject to the jurisdiction of the District Court of Munich.
- 11.3 If any provision of this Agreement is determined to be illegal or in conflict with the applicable law, the validity of the remaining provisions shall not be affected. The ineffective provision shall be replaced by an effective provision which is economically equivalent. The same shall apply in case of a gap.
- 11.4 This Agreement will come into force upon the signature of all Partners with retroactive effect on January 01, 2015.

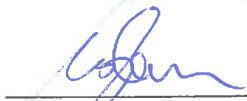


Mannheim, June, 1st, 2015

ZEW, Zentrum für Europäische Wirtschaftsforschung GmbH

A handwritten signature in black ink, appearing to be 'T. Kohl', written over a horizontal line.

Thomas Kohl
(Director)

A handwritten signature in blue ink, appearing to be 'C. Rammer', written over a horizontal line.

Dr. Christian Rammer
(Deputy)

Hanover,

Prof. Diego Comin

A handwritten signature in blue ink, consisting of several overlapping loops and a long horizontal stroke extending to the right. The signature is written over a thin horizontal line.

signature

26-P-315466-15/B6-Fis

Lund,

CIRCLE, Lund University

A handwritten signature in blue ink, appearing to be 'C. Sjö', written over a horizontal line.

signature

Munich, 08.06.2015

Fraunhofer-Gesellschaft zur Förderung
der angewandten Forschung e. V.



Dr. Birgit Gebler



Thomas Fischer