Banking regulation and the interconnectedness of the banking industry

A Data Management Plan created using DMPonline.be

Creator: Cédric Huylebroek

Affiliation: KU Leuven (KUL)

Funder: Fonds voor Wetenschappelijk Onderzoek - Research Foundation Flanders (FWO)

Template: FWO DMP (Flemish Standard DMP)

Grant number / **URL:** 11C7923N

ID: 198974

Start date: 01-11-2022

End date: 01-11-2026

Project abstract:

This research project studies the effects of banking regulation and supervision on the banking sector and the real economy. In particular, the first project uses administrative data from Italy to study whether regulatory reforms have affected banks' systemic interconnectedness; the second project uses administrative data from Norway to study whether regulatory reforms have affected banks' domestic interconnectedness; the third project uses administrative data from Brazil to study the effect of supervisory interventions on banks' balance sheet and lending decisions. In this way, this research project provides a better understanding of the effects of banking regulation on the stability of the financial system and the real economy.

Last modified: 22-04-2023

Banking regulation and the interconnectedness of the banking industry
Application DMP

Questionnaire

Describe the datatypes (surveys, sequences, manuscripts, objects …) the research will collect and/or generate and /or (re)use. (use up to 700 characters)

The project will use supervisory datasets managed by Bance d'Italia (the central bank of Italy), Norges Bank (the central bank of Norway), and Banco Central do Brasil (the central bank of Brazil).

Specify in which way the following provisions are in place in order to preserve the data during and at least 5 years after the end of the research? Motivate your answer. (use up to 700 characters)

The datasets are part of the central banks' main operations and will be preserved on the premises of the central banks for the foreseeable future. The central banks have sufficient storage capacity and can guarantee safe and secure storage. The data cannot be shared, but anyone can request access from the central banks.

What's the reason why you wish to deviate from the principle of preservation of data and of the minimum preservation term of 5 years? (max. 700 characters)

Not applicable.

Are there issues concerning research data indicated in the ethics questionnaire of this application form? Which specific security measures do those data require? (use up to 700 characters)

Not applicable.

Which other issues related to the data management are relevant to mention? (use up to 700 characters)

The data are owned by central banks and can therefore not be shared. In general, the data are still accessible through cooperation with the central banks or its employees.

DPIA

Have you performed a DPIA for the personal data processing activities for this project?

- Not applicable

GDPR

Have you registered personal data processing activities for this project?

- Not applicable

Banking regulation and the interconnectedness of the banking industry
FWO DMP (Flemish Standard DMP)

1. Research Data Summary

List and describe all datasets or research materials that you plan to generate/collect or reuse during your research project. For each dataset or data type (observational, experimental etc.), provide a short name & description (sufficient for yourself to know what data it is about), indicate whether the data are newly generated/collected or reused, digital or physical, also indicate the type of the data (the kind of content), its technical format (file extension), and an estimate of the upper limit of the volume of the data.

Dataset Name	Description	New or reused	Digital or Physical	Only for digital data Digital Data Type	Only for digital data Digital Data format	Only for digital data Digital data volume (MB/GB/TB)	Only for physical data Physical volume
Bank board member data for Italian banks	The register of bank board members comes from the Bank of Italy and allows to identify individuals serving in boards of directors of different banks.	Reused	Digital	Compiled/aggregated data	.csv	<1GB	/
Administrative credit register data for Italy	Data on banks' loan market share at the provincial-quarter level can be retrieved from the Italian credit register. This dataset contains the individual loans granted by banks to firms active in Italy.	Reused	Digital	Compiled/aggregated data	.csv	<100GB	/
Administrative security holdings data for Italian banks	The Security Register dataset is managed by the Bank of Italy and includes microdata on all securities investments—at the security level (ISIN code)—for each bank in Italy (bonds, ABS, equities, derivatives, and shares of mutual funds). Data are available at monthly frequency from 1999. For each security, banks must report the notional amount they hold at the end of each period.	Reused	Digital	Compiled/aggregated data	.csv	<10GB	/

Data on interbank loans between Italian banks	Data on interbank loans (month-end gross unconsolidated bilateral interbank deposits) administered by the Bank of Italy. The database allows for a distinction between different claims (CDs, current accounts, repos, other loans, subordinated and unsubordinated bonds) classified according to their maturity (up to 18 months, over 18 months), currency of denomination and counterpart nationality (Italy vs. rest of the world)	Reused	Digital	Compiled/aggregated data	*.csv*	<10GB	/
Norwegian bank financial statement data	Detailed data on Norwegian banks' financial statements. This dataset reports components of balance sheets and income statements for all Norwegian banks – including commercial banks, savings banks, subsidiaries of foreign banks, branches of foreign banks and bank-affiliated mortgage companies	Reused	Digital	Compiled data	.cvs	<10GB	/

Administrative data from the Norwegian Tax Authority	This data is a high-quality matched firm-bank administrative register from Norges Bank. The unit of observation in this database is an individual loan and the frequency is annual. For every loan, we observe the firm-bank identifiers as well as the flow of interest paid during the year and the end-of-year stock of debt. Because the data is collected and maintained by the tax authority as a basis for corporate taxation, the variables are essentially measurement error-free.The data set covers all limited liability companies for the time period of 2000-2019, which accounts for roughly 90% of private sector employment for most years.	Reused	Digital	Compiled data	.cvs	<10GB	
Administrative firm data from Bisnode for Norwegian firms	Firm data comes from the credit rating agency Bisnode, which provides annual balance sheet and income statement items on all Norwegian firms for 1999-2019. The data also includes information on firms' credit rating scores and other firm characteristics such as age, location and industry.	Reused	Digital	Compiled data	.cvs	<10GB	
Bank financial statement data for banks active in Brazil	This data contains the balance sheet and income statement data for the individual institutions as well as the financial consolidates active in Brazil at a monthly frequency. Variables include, among others, total assets, total deposits, total equity, total liquid assets, gross loans, non-performing loans, and loan loss provisions.	Reused	Digital	Compiled/aggregated data	.csv	<1GB	/

Administrative credit register data from the central bank of Brazil	This data is a high-quality matched firm-bank administrative register managed by the central bank of Brazil. The unit of observation in this database is an individual loan and the frequency is monthly. For every loan, we observe the firm-bank identifiers as the end-of-month outstanding loan, interest rate, loan maturity, collateral, and other loan performance characteristics.	Reused	Digital	Compiled/aggregated data	.csv	<100GB	/
Administrative supervisory interventions data from the central bank of Brazil	We obtain detailed information about the supervisory interventions of the central bank of Brazil from a proprietary dataset administered by the Central Bank of Brazil. This supervisory dataset comprises information on the date of each supervisory intervention, the underlying issue (or motive) of the intervention, and the time needed by the affected bank to resolve the issue.	Reused	Digital	Compiled/aggregated data	.csv	<1GB	/
Firm financial statement data for Brazilian firms	We obtain quarterly firm employment and profitability data from RAIS (Relação Anual de Informações), which essentially comprises all tax-registered firms operating in Brazil. The dataset is administered by the Brazilian Ministry of Labor and Employment, but available via the central bank of Brazil.	Reused	Digital	Compiled/aggregated data	.csv	<100GB	/

If you reuse existing data, please specify the source, preferably by using a persistent identifier (e.g. DOI, Handle, URL etc.) per dataset or data type:

The datasets are all managed by central banks.

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? Describe these issues in the comment section. Please refer to specific datasets or data types when appropriate.

- No

Will you process personal data? If so, briefly describe the kind of personal data you will use in the comment section. Please refer to specific datasets or data types when appropriate.

- No

Does your work have potential for commercial valorization (e.g. tech transfer, for example spin-offs, commercial exploitation, …)? If so, please comment per dataset or data type where appropriate.

- No

Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material/Data transfer agreements/ research collaboration agreements)? If so, please explain in the comment section to what data they relate and what restrictions are in place.

- Yes

As the data is managed by the central bank, a co-author from the central bank is needed to be able to work with the data.

Are there any other legal issues, such as intellectual property rights and ownership, to be managed related to the data you (re)use? If so, please explain in the comment section to what data they relate and which restrictions will be asserted.

- Yes

As the data is managed by the central bank, a co-author from the central bank is needed to be able to work with the data.

2. Documentation and Metadata

Clearly describe what approach will be followed to capture the accompanying information necessary to keep data understandable and usable, for yourself and others, now and in the future (e.g., in terms of documentation levels and types required, procedures used, Electronic Lab Notebooks, README.txt files, Codebook.tsv etc. where this information is recorded).

We will provide:
- documentation on data definitions
- documentation on how to merge the datasets
- a "read me" file (also called "do files") describing the precise handling of the datasets and empirical analysis
These are requested by journals during the publication process and are made available online to allow replication of results by other researchers (or for personal future projects).

Will a metadata standard be used to make it easier to find and reuse the data? If so, please specify (where appropriate per dataset or data type) which metadata standard will be used. If not, please specify (where appropriate per dataset or data type) which metadata will be created to make the data easier to find and reuse.

- No

The "read me" files ("do files") that are made available together with the data sources should make sure that researchers can easily find and reuse the employed data.

3. Data storage & back-up during the research project

Where will the data be stored?

The data will be stored on the central banks' premises and is heavily protected by their IT systems.

How will the data be backed up?

The data will be backed up by the central banks.

Is there currently sufficient storage & backup capacity during the project? If yes, specify concisely.
If no or insufficient storage or backup capacities are available, then explain how this will be taken care of.

- Yes

The central banks have more than sufficient storage given that these datasets are part of their main operations.

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

The data is highly protected by the central banks and cannot leave the central banks' premises.

What are the expected costs for data storage and backup during the research project? How will these costs be covered?

The central banks do not charge additional costs for using the data.

4. Data preservation after the end of the research project

Which data will be retained for at least five years (or longer, in agreement with other retention policies that are applicable) after the end of the project? In case some data cannot be preserved, clearly state the reasons for this (e.g. legal or contractual restrictions, storage/budget issues, institutional policies...).

The data will be preserved (for at least 5 years) by the central banks by default.

Where will these data be archived (stored and curated for the long-term)?

The data will be archived for the foreseeable future at the central banks' premises.

What are the expected costs for data preservation during the expected retention period? How will these costs be covered?

These costs will be born by the central banks.

5. Data sharing and reuse

Will the data (or part of the data) be made available for reuse after/during the project? In the comment section please explain per dataset or data type which data will be made available.

- Yes, in a restricted access repository (after approval, institutional access only, …)

The data are owned by the central banks, and can therefore not be shared. In general, the data are still accessible through cooperation with the central banks or their employees.

If access is restricted, please specify who will be able to access the data and under what conditions.

All coding and programs are made available when papers are published at journals. Data from the central banks is accessible through cooperation with the central banks or their employees.

Are there any factors that restrict or prevent the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)? Please explain in the comment section per dataset or data type where appropriate.

- Yes, Other

The data are owned by the central banks, and can therefore not be shared. In general, the data are still accessible through cooperation with the central banks or their employees.

Where will the data be made available? If already known, please provide a repository per dataset or data type.

Upon publication of the research results.

When will the data be made available?

Upon publication of the research results.

Which data usage licenses are you going to provide? If none, please explain why.

The data cannot be shared (due to central bank restrictions) so a data usage license is not applicable.

Do you intend to add a PID/DOI/accession number to your dataset(s)? If already available, you have the option to provide it in the comment section.

- No

What are the expected costs for data sharing? How will these costs be covered?

There are not expected costs related to this.

6. Responsibilities

Who will manage data documentation and metadata during the research project?

My coauthors from the central banks and me (Cédric Huylebroek).

Who will manage data storage and backup during the research project?

My coauthors from the central banks and me (Cédric Huylebroek).

Who will manage data preservation and sharing?

My coauthors from the central banks and me (Cédric Huylebroek).

Who will update and implement this DMP?

My coauthors from the central banks and me (Cédric Huylebroek).

Created using DMPonline.be. Last modified 22 April 2023

12 of 12