

# FAQ on the Pilot project for research access to confidential statistical data

## Contents

<b>Eligibility</b>	<b>3</b>
Do all researchers on a project need to be affiliated with a scientific research body accredited by Eurostat for the project to be eligible?	3
Is it possible for an affiliated student to write the paper in cooperation with a consulting firm, assuming the data will not be passed on to the consulting firm?	3
Are master's students eligible?	3
Are PhD students eligible?	4
Is it possible to submit an application if the Eurostat listing process is still ongoing at the time of application?	4
The website mentions that five euro area central banks are participating in the project. Are researchers and universities located in other countries still eligible?	4
<b>Availability and anonymisation of datasets</b>	<b>5</b>
What data will be available?	5
What is the level of granularity of the data provided? Will the data granularity extend to the assets held by banks?	5
What is the time span of the data provided?	5
Do the anonymised identifiers allow researchers to track the same bank consistently across time periods within a given dataset and across the different datasets?	5
Do the identifiers preserve information on institutional relationships, for example whether a bank is part of a group structure or affiliated with a parent organisation, as is available in the Register of Institutions and Affiliates Data?	6
<b>Access to data</b>	<b>6</b>
What are the different ways to access the data?	6
Will access to data via remote access and remote execution allow access to data without being physically present at a research data centre? If yes, will researchers be able to run code on the data?	7
<b>Application procedure and deadlines</b>	<b>7</b>

In the case of a team of researchers, will the application form have to be filled in only by the lead researcher or by all researchers in the team? Is it possible to add a member at a later stage?	7
When will the final selection of research projects be communicated to applicants?	7
Once a project is selected, when will the researchers be able to start working with the data?	8

**European Central Bank**

Directorate General Communications

Sonnemannstrasse 20, 60314 Frankfurt am Main, Germany

Tel.: +49 69 1344 7455, email: [media@ecb.europa.eu](mailto:media@ecb.europa.eu), website: [www.ecb.europa.eu](http://www.ecb.europa.eu)

Reproduction is permitted provided that the source is acknowledged.

## Eligibility

Do all researchers on a project need to be affiliated with a scientific research body accredited by Eurostat for the project to be eligible?

Yes, for the project to be eligible, all researchers who are part of the project team have to be affiliated with a scientific research body listed by Eurostat.

Is it possible for an affiliated student to write the paper in cooperation with a consulting firm, assuming the data will not be passed on to the consulting firm?

No, this is not possible. As stated in [Pilot project on research access to CSI collected under Council Regulation 2533/98 – Criteria for designating eligible researchers](#), section 1 on Affiliated researchers and eligible scientific research bodies:

*“As a general rule, persons who are currently performing a commercial activity for banks or financial service providers (in particular insurers or investment companies), a commercial activity pertaining to financial topics on behalf of other companies, a commercial self-employed financial activity, or are due to start such an activity during the term of the research project, cannot be eligible researchers. This also includes consulting services and the preparation of expert opinions. If such an activity is being performed, the ECB has the right, after weighing up existing or potential conflicts of interest, to withdraw access to confidential data, thus ending the involvement of the researcher in the research project.*

*The performance of a function in an official body (e.g. supervisory board) or employment with a bank or a financial service provider, or the appointment to start such a function/employment during the term of the research project, leads to the ECB immediately terminating access to its confidential data and the involvement of the researcher in the research project for good cause.”*

Are master’s students eligible?

Yes, master’s students are eligible, providing they meet all other eligibility criteria, in particular the Eurostat listing of their university.

As stated in [Pilot project for research access to confidential statistical data – Research Dataset Request Form](#), “Bachelor and Master students shall fill out the forms with their professor as lead researcher”.

## Are PhD students eligible?

Yes, PhD students are eligible, providing they meet all other eligibility criteria, in particular the Eurostat listing of their university.

## Is it possible to submit an application if the Eurostat listing process is still ongoing at the time of application?

Yes, it is possible to submit an application if the Eurostat listing process is still ongoing at the time of application.

A Eurostat listing is a mandatory eligibility criterion for access to the data. Applicants who intend to submit an application while the Eurostat listing process for their university is still ongoing are required to:

- (i) provide a clear and detailed account of the status of the listing process on the application form;
- (ii) keep the ECB informed of the progress throughout the process.

Access to data will be provided only when the listing process with Eurostat has been finalised.

## The website mentions that five euro area central banks are participating in the project. Are researchers and universities located in other countries still eligible?

Yes, researchers affiliated with universities in other euro area countries can apply to the pilot project and get access, provided their application is assessed as eligible and their project is selected. Access to data is not linked to the nationality of the researcher or to the location of the eligible scientific research body.

With regard to the process for getting access to the data, please see the “Data access” section of [Pilot project for research access to confidential statistical data](#).

## Availability and anonymisation of datasets

### What data will be available?

The data in scope of the project are:

- (i) balance sheet information for all euro area banks, broken down by assets and liabilities; loans, deposits and other financial instruments; stocks and transactions; counterpart area and counterpart sector (the dataset covers around 3,000 institutions);
- (ii) bank interest rates and business volumes.

### What is the level of granularity of the data provided? Will the data granularity extend to the assets held by banks?

The data provided are balance sheet items, so the level of granularity is that of a balance sheet item, not that of an asset.

### What is the time span of the data provided?

The time span is from July 2007 to December 2023 for both individual balance sheet items and interest rate statistics.

For balance sheet items, the sample covers around 3,000 credit institutions with backdata to (at least) December 2019.

### Do the anonymised identifiers allow researchers to track the same bank consistently across time periods within a given dataset and across the different datasets?

Yes, we confirm that the anonymised identifiers are designed to allow researchers to track the banks consistently across the years and to link the two datasets (bank balance sheets and bank interest rates).

Do the identifiers preserve information on institutional relationships, for example whether a bank is part of a group structure or affiliated with a parent organisation, as is available in the Register of Institutions and Affiliates Data?

Variables allowing direct identification will be removed from the datasets. Direct identifiers will be replaced with the same dummy code in both datasets to enable data linking between the two datasets.

## Access to data

What are the different ways to access the data?

There are three types of data access:

- (i) **On-site access:** An access mode allowing an external researcher to consult anonymised confidential statistical information (CSI), while not allowing direct identification, in a secure environment, usually at the premises of the data access point, with measures in place to ensure that confidential information cannot be copied, removed from the secured environment or compared with data from other sources. In this setting, direct identifiers are removed from the data, but the statistical information may still allow indirect identification of statistical units.
- (ii) **Remote access:** An access mode allowing an external researcher to consult anonymised CSI, while not allowing direct identification, on a secured device or application at a research facility, with measures in place to ensure that confidential information cannot be copied or downloaded from the secured device/application or compared with data from other sources. In this setting, direct identifiers are removed from the data, but the data may still allow indirect identification of statistical units.
- (iii) **Remote execution:** An access mode where an external researcher sends scripts to a data access point to be executed on CSI datasets. In this setting, no confidential information is ever disclosed to the external researcher.

For further information on data access and available software, please see [Pilot project for research access to confidential statistical data – Types of access and software provided](#).

Will access to data via remote access and remote execution allow access to data without being physically present at a research data centre? If yes, will researchers be able to run code on the data?

For both access types – remote access and remote execution – physical presence is not required.

With remote access, researchers can log onto a server provided by the central bank to view the data and run code directly on it. The results of the analysis can only be downloaded once the central bank has reviewed them for potential confidentiality risks.

In remote execution, the central bank provides researchers with dummy data to help them write their code. Researchers can then send their code to the central bank, which will apply the code to the original data. Following a review of the coding results for potential confidentiality risks, the central bank will provide the results to the researchers.

## Application procedure and deadlines

In the case of a team of researchers, will the application form have to be filled in only by the lead researcher or by all researchers in the team? Is it possible to add a member at a later stage?

Please see [Pilot project for research access to confidential statistical data – Research Dataset Request Form](#):

*“In case of a joint research proposal, only one application form, filled in by the lead researcher in the project, needs to be completed in full. All other researchers need to submit an application form with sections 1, 2.2, and 8 filled in, and referring in section 2.1 to the title of the joint proposal. Upon acceptance of the project, all researchers will have to submit the confidentiality declaration, the signed privacy statement and the documents requested under section 7.”*

Access will be granted to the researchers who have passed the eligibility assessment, provided the project is selected. No additional access will be provided thereafter.

When will the final selection of research projects be communicated to applicants?

The timeline for finalising the selection process depends largely on the number of applicants. We are therefore unable to commit to a specific deadline at this point in time.

Once a project is selected, when will the researchers be able to start working with the data?

The target access date for researchers is of the last quarter of 2025.