Linking data for MFIs

Technical Report 2022-03

Metadata version: 3

Deutsche Bundesbank, Research Data and Service Centre

Harald Stahl

Research Data and Service Centre

Abstract

Different departments of the Deutsche Bundesbank use different identifiers for MFIs (banks). This technical documentation describes how different research datasets can be matched on an individual basis.

Contents

1 Banking statistics

Linking data for MFIs from the banking statistics is in general easily done. However, there are some caveats. We will look at the monthly balance sheet statistics (BISTA), the quarterly borrower statistics (VJKRE), the MFI interest rate statistics (ZISTA = MIR), and the selected master data for MFIs (MaMFI). They are all panel datasets and use a common identifier for the MFIs and a common representation for the date. The latter is delivered in separate variables YEAR, MONTH, and DAY so that even different frequencies can easily be linked. The drawback is, that the user has to create his or her own time variable.

The situation is somewhat different for the banks' profit and loss accounts (GuV). Before we can proceed and explain how to link the banks' profit and loss accounts we have to understand what a multi-office bank is.

1.1 Multi-office banks

A multi-office bank (Gesamtinstitut) is a bank with legally dependent branches abroad. Their data is consolidated. We call MFIs multi-office banks in the narrow sense if they actually own legally dependent branches abroad and we call all MFIs multi-office banks in the wide sense. Owning legally dependent branches abroad is something that can change over time. Since our research datasets are panel datasets we need a unique identifier within the dataset for each MFI.

To distinguish multi-office banks from domestic parts they have unique identifiers and different names. By implication, every MFI located in Germany has two distinct identifiers. BAID_DOM is the name for the domestic parts and BAID_MOF for the multi-office banks. If the bank identifier is just BAID the researcher has to care whether the MFI is a multi-office bank or a domestic part.

There is a mapping table map_BAID_DOM_BAID_MOF that relates the domestic parts to the multi-office banks.

1.2 Banks' profit and loss accounts

The results from the profit and loss accounts are based on the published annual reports of the individual institutions in accordance with the provisions set forth in the German Commercial Code (Handelsgesetzbuch or HGB) and the Regulation on the Accounting of Credit Institutions (Verordnung über die Rechnungslegung der Kreditinstitute, or RechKredV). Data for legally dependent branches abroad for example should be included in a bank's profit and loss accounts whereas they are not accounted for in the other statistics. For the external position of banks and for the monthly balance sheet statistics legally dependent branches abroad have their own reporting obligations. Therefore, the identifier of the banks' profits and loss accounts is BAID_MOF.

Next, we have to take a closer look at linking data with different frequencies.

1.3 Linking different frequencies

Data comes in different frequencies. There is the monthly balance sheet statistics (BISTA), the quarterly borrower statistics (VJKRE) and the yearly statistics on the banks' profit and loss accounts (GuV). Master data (MaMFI) is monthly. The easiest way to proceed is to aggregate high frequency data to the lowest frequency available while sacrificing the information contained in the high frequencies. Another approach would be to disaggregate the low frequency data. It is best to start the explanations with linking the master data.

Aggregation to the lowest frequency available can be done by averaging or summing up or by taking the end of period value of the high frequency data. At the beginning of a time-series, low frequency data may cover a time-period that is not available for the high-frequency data. E.g. if the reporting obligations for the monthly balance sheet statistics start in April 2010 but the bank existed before, the bank's profit and loss accounts cover the months January 2010 until March 2010, too. At the end of a time-series, high frequency data may still be available, while there is no low frequency data anymore. E.g. if a bank ceases its business in April 2010, high frequency data is available for the months January 2010 until March 2010 but the bank's profit and loss accounts data ends in 2009. The bank's profit and loss accounts use four variables from the monthly balance sheet statistics. If the reporting obligations for the monthly balance sheet statistics start in April 2010 for example but the bank existed before, the bank's profit and loss accounts imputes the missing data by the first available figure i.e. the value for April 2010.

1.4 Mergers

A merger is a special case of the aforementioned. Assume bank "B" absorbs bank "A" in October 2010. For the absorbed bank, its monthly balance sheet data is available for January 2010 until September 2010 but its profit and loss accounts data ends in 2009. On the other hand, the customers profit and loss accounts data for 2010 includes bank "A"'s data even for the months January 2010 until September 2010. If there is no major interest in the period of the merger one may predate the merger, in the example to December 2009. The high frequency data of the absorbed bank for the suppressed period then has to be assigned to the customer's data. Aggregating the data for the year of the merger can be done by using an RDSC ado-file (mfimerger HV21_330, period(YEAR)).

In the data the date of a merger is when equity shows up in the data of the absorbing institute, not the legal date. These two dates may differ by several months with the drawback that if the legal date is, say November 2002, and the date of the merger in the data is, say February 2003, the institute has no profit and loss account in 2002 and in 2003. Thus, the high frequency data of the absorbed institute has to be assigned to the absorbing institute in 2002 as well as in 2003.

It is possible to take this into account in the ado-file by creating a variable for the year of the exit (for each month), say EXITYEAR, and to create a new variable, say SELYEAR, that is equal to EXITYEAR if YEAR is equal to EXITYEAR or to EXITYEAR-1 and additionally the variable gvK150 from the banks' profit and loss accounts has to be missing. gvK150 is equal to one if a for the corresponding year a profit and loss account is available. Use then mfimerger HV21_330, period(SELYEAR).

1.5 Entries and exits

Banks "enter" the banking statistics at the time they start taking deposits and "exit" at the time when they stop taking deposits. That means that they may exist before they have to report in the banking statistics and they may exist after leaving the banking statistics, e.g. banks that are liquidated after a failure until the liquidation is completed. It may happen that after a MFI got in serious troubles part of the business is transferred into a new "good" bank and the rest stays in the "bad" bank until liquidation is completed. Then, the bad bank exits banking statistics. Actually, the new good bank should get a new ID. However, in a few cases, in banking statistics the good bank keeps the ID of the predecessor since the bad bank does not need it anymore.

2 Banking supervision

Research data from banking supervision is anonymized using BAID_MOF. Banking statistics focuses on calculating monetary aggregates and is therefore interested in the domestic part of MFIs. In contrast, banking supervision looks at the MFI with legally dependent affiliates abroad included.

Supervisory data and the profit and loss accounts of banks can be linked using BAID_MOF. Data of the domestic parts can be merged using the mapping table map_BAID_DOM_BAID_MOF that relates the domestic parts to the multi-office banks.

2.1 Mergers

Banking supervision treats absorbing institutes after a merger as new institutes. mfimerger.ado with the split option creates a suitable identifier.

2.2 Entries and exits

Banks that are liquidated after a failure exist outside the banking statistics until the liquidation is completed. They keep their identifier. E.g. it may happen that after a MFI got in serious troubles part of the business is transferred into a new "good" bank and the rest stays in the "bad" bank until liquidation is completed. Then, the bad bank exits banking statistics but not banking supervision. In the original data, the new good bank gets a new ID and the bad bank keeps the former ID.

3 External sector

External sector data cover enterprises not being MFIs, enterprises being MFIs, the public sector and individuals. Here, the focus is on the balance of payments statistics, in particular the Statistics on international financial and capital transactions (SIFCT). Historically, for the external sector there are two different sets of identifiers for MFIs.

3.1 Statistics on international financial and capital transactions (SIFCT)

In the anonymized version of the SIFCT the eight-digit identifier AWMUSID of a MFI is derived from its corresponding six-digit identifier in the banking statistics BAID_DOM according to 10000000 + BAID_DOM, since the AWMUSID of a non MFI in the SIFCT can have six-digits, too. AWMUSID is for merging SIFCT with banking statistics data or other data using the BAID. For merging SIFCT with MiDi there is the second identifier AWMUS_CD.

3.1.1 Mergers

Merged MFIs are allowed to use the identifiers of the separate institutes before the merger for a transition period of up to two months in their external sector reports. To be in line with the banking statistics the SIFCT uses the identifier corresponding to the MFI in the banking statistics. A more severe issue is the question which identifier to use after the merger. Both banking statistics and external sector almost always continue with one of the previous identifiers. However, they may choose different identifiers. As before, to be in line with the banking statistics the SIFCT uses the identifier corresponding to the MFI in the banking statistics.

3.1.2 Entries and exits

The external sector statistics follow closely the legal definition of an enterprise. Banks may exist before they appear in the banking statistics and they may exist after leaving the banking statistics, e.g. banks that are liquidated after a failure until the liquidation is achieved. E.g. it may happen that after a MFI got in serious troubles part of the business is transferred into a new "good" bank and the rest stays in the "bad" bank until liquidation is achieved. Then, the bad bank exits banking statistics but not necessarily the external sector statistics. The new good bank most likely keeps the former ID.

3.2 Micro database Direct investment (MiDi) - affiliates abroad

MiDi covers MFIs abroad that are either legally dependent or independent affiliates of domestic MFIs so that there is an overlap with the research dataset External positions of banks (AUSTA). MiDi may cover less affiliates because there is a reporting threshold. On the other side AUSTA covers only majority owned affiliates. In the case of legally dependent affiliates abroad, all legally dependent affiliates of a MFI in a country are aggregated so that there is just one report. For

MiDi all affiliates have to be reported separately. Formally, the definition of legally dependent affiliate and legally independent affiliate coincide in both datasets but in practice an affiliate that is classified as legally dependent in MiDi may be an independent affiliate in AUSTA and the other way round.

To be in line with the banking statistics the affiliates are classified as dependent or independent, accordingly. In the case of several legally dependent affiliates in one country of one and the same MFI the mapping is not one-to-one but each legally dependent affiliate is related to the aggregate in the AUSTA. Therefore, the corresponding data normally have to be aggregated.

Since MiDi is a yearly dataset mergers may cause ostensible contradictions during the year of the merger. This is aggravated by the fact that a legally dependent or independent affiliate abroad my get a new majority owner and changes at the same time status from independent to dependent or the other way round.

Citation requirements

For any study or other document which is made available to the public and contains information derived from the provided data, the researcher is obliged to properly cite the data source as:

Stahl, H. and (2022). Linking data for MFIs, Technical Report 2022-03 – Metadata version 3. Deutsche Bundesbank, Research Data and Service Centre.