

Quality issues in the compilation of bank interest rate statistics: A European perspective

Andreas Kuchler

Danmarks Nationalbank (Central Bank of Denmark), Copenhagen, Denmark

E-mail: aku@nationalbanken.dk

Abstract

Statistics on bank interest rates have proved useful for analyses of, e.g., the monetary transmission mechanism and lending conditions for various sectors of the economies before and during the financial crisis. The quality of bank interest rate statistics is mainly the result of an interplay between two broad factors: relevancy and accuracy. By relevancy is meant that the output should be a set of comprehensive aggregate indicators, which are relevant and useful for a diverse portfolio of users. Accuracy refers to the precision and consistency of the published aggregates, which may be achieved by a well-designed reporting system focused on efficient data checking routines and a collaborative dialogue with reporters. Various quality considerations have shaped a new system for banks' reporting to Danmarks Nationalbank, which is being implemented in the summer of 2013. This paper discusses a number of these issues from a conceptual and a practical point of view. The discussion indicates that quality in itself is a dynamic concept, which calls for a continuous balancing of relevancy and accuracy in the process of responding to increased user needs.

Keywords: Bank Interest Rates; Statistical Reporting Systems; Relevancy; Accuracy.

1. Introduction

One of the main goals of statistics on banks' interest rates is to provide users with a set of indicators of the interest rates and flows of interest payments among different sectors of the economy. Traditionally, focus has been on households and non-financial corporations, but there is a growing analytical interest in data for other sectors as well. The overall quality of interest rate statistics is determined in an interplay between two broad factors, namely relevancy and accuracy. First, the statistical output produced should be relevant and add value for the various user groups. This requirement entails that the compiler needs to assess the diverse demands of user groups in terms of e.g. breakdowns by counterparty characteristics, loan sizes, maturity and collateral provided. Also, issues related to publication, such as aggregation levels, frequency and timeliness, play a role for the relevance of bank interest rate statistics.

The decision regarding which data to collect and publish also involves considerations about the level on which data should be reported (e.g. on a loan-by-loan basis or in aggregates). A rather detailed reporting has the advantage of flexibility in the definition of output as the compiler has many options for aggregating and combining breakdowns. To achieve a high quality, data should be accurate and consistent. This requires a well-designed reporting system and an efficient data checking procedure; the latter being more important as the data volume is increasing when more granular data is collected. Also, a good collaboration with reporters is key to a high quality of the individual bank's reporting. To counter the fact that reporting may be seen as a burden by some reporters, provision of incentives for reporters to improve the quality of reporting could be considered.

The remaining part of the paper is structured around the two broad factors just identified. Hence, section 2 focuses on the issue of how interest rate statistics should reflect diverse user needs, and section 3 discusses in greater detail the issues relevant for ensuring that published aggregates are of sufficient accuracy. Section 4 concludes.

2. Ensuring relevancy and usefulness of bank interest rate statistics

Relevancy and usefulness are ultimate goals of any official statistics. The diverse demands of many potential user groups are to be taken into account when considering the societal significance of interest rate statistics. Relevant user groups include the media, financial market participants, analysts, researchers and official authorities. In addition, the general public increasingly uses statistics in various contexts. Maximizing the usefulness of interest rate statistics for all those user groups is therefore a multi-dimensional challenge, consisting of reporting issues, compilation issues as well as dissemination issues.

2.1. *The response to different user needs*

The interests of the media and general public may be broad. Often their focus is on interest rates for a 'typical' bank customer. Official authorities, financial market participants, researchers and analysts are typically interested in more detailed breakdowns of statistics. Topics, for which interest rate statistics are useful input, include analyses of business cycles, credit developments, competition issues, interest burdens of households and non-financial corporations and monetary transmission channels. From the point of view of the compiler this entails emphasis on collecting sufficiently detailed data as well as on employing a flexible dissemination strategy. Based on recent experiences of user demands there seems to be a potential for increased dissemination of e.g. measures of dispersion to complement the more basic aggregates, both generally and for specific topics¹. Analyses of the distribution of interest rates across banks are often found to be limited by difficulty of comparisons due to differences in product compositions, etc. This is further discussed in section 2.2. In addition, there is a potential to add value to users by developing tools which can provide a better understanding of the underlying explanations for interest rate developments over time².

A related consideration is the sectoral coverage of interest rate statistics. Households and non-financial corporations are obvious to include as they are important for many types of analyses, for example analyses of the transmission mechanism of central bank interest rates and economic analyses of debt burden, etc. In addition, interest rates on interbank loans, as well as on positions between banks and insurance corporations, pension funds and other parts of the non-bank financial sector, may be indicative of the transmission of liquidity in the financial system.

Along with increasing detail of published data it could also be considered whether a composite measure, or a number of measures, of cost of borrowing can be constructed. The European System of Central Banks has made an attempt to develop cost of borrowing indicators for households and non-financial corporations for short and long term loans to residents of euro area countries. For an application of the ideas in this work, see ECB (2012, pp. 35-39).

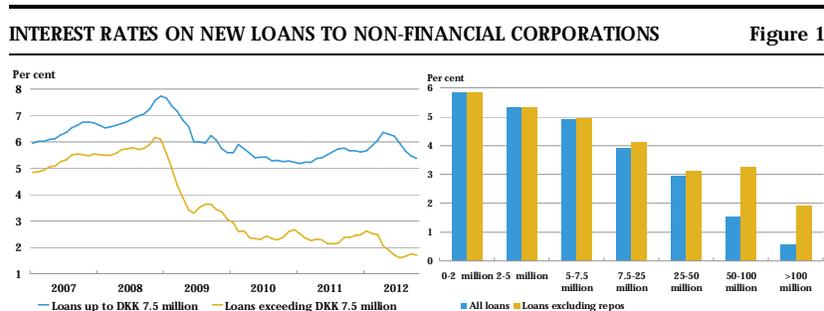
2.2. *Comparability, breakdowns and aggregation levels*

The use of interest rate statistics to address questions such as pricing structures, competition issues and dispersion of interest rates in various categories raises the need

¹ As a very specific example the Danish government imposed a regulation which implies that banks taken over by the state-owned entity for resolving failed banks should not distort competition. Hence, these banks are not allowed to charge interest rates lower than those corresponding to the 9th decile of lending rates of other commercial banks; a level which is determined by the distribution of interest rates reported for the bank interest rate statistics.

² Huerga and Steklacova (2008) discuss a possible decomposition of interest rate developments over time into interest rate effects and weight effects. Their ideas provide a useful starting point for understanding the development in aggregate interest rates over time, and regular publishing of a decomposition along these lines at the national level could be useful.

for sufficiently detailed breakdowns in order to promote meaningful comparisons. For example, in order to identify differences in banks' pricing policies, one should ideally compare similar types of loans provided to similar borrowers (or alternatively, make sufficient adjustments to the aggregates). Hence, detailed breakdowns are needed to distinguish different types and sizes of loans, collateral, maturity and borrower characteristics.



Note: The interest rates in the left panel are shown as 3-month moving averages. The right panel is based on data for the period April-August 2012, and shows the average interest rates for different loan sizes.
 Source: Andreasen et al. (2012).

A recent analysis of Danish bank lending rates to non-financial corporations shows a widening of the spread between small and large loans since 2008, cf. Figure 1 (left). While the regular interest rate statistics data do not allow for a further breakdown of this development, an ad-hoc data collection exercise revealed that the impact of the volume of repo transactions on the aggregate interest rate is substantial. Figure 1 (right) indicates that this is the case, in particular when taking into account that large loans have a very high impact on the aggregate interest rates. This example reiterates the need for sufficiently detailed breakdowns of aggregate interest rates.

An additional example of the difficulty of comparing aggregates in a Danish context is interest rates on loans for house purchase³. In January 2013, the average interest rate for a mortgage bank in Denmark is 2.9 per cent, while the corresponding rate for a commercial bank is 4.3 per cent. However, these differences can only be meaningfully compared by taking into account that loans from mortgage banks are only used to finance up to 80 per cent of the real estate value, while bank loans may cover the remaining part. This means that the risk of default on the average bank loan is higher. Based on this example, compilers should be encouraged to publish interest rate statistics for different bank types, where applicable. In addition to the improvement in comparability, this could also increase awareness of differences in financial structures across countries or regions.

There is, however, a trade-off between tailoring interest rate statistics to the banking system under consideration and at the same time ensuring international comparability. It would be of value to users if the statistical community, after consulting with relevant stakeholders, agreed to a set of indicators which ideally would be compiled in all countries; and which in individual countries could be supplemented by country specific variables. Within the EU, a set of interest rate statistics for member countries are regularly compiled by the ECB based on input from central banks or statistical authorities in member states⁴. The experience from the EU is that international

³ In general, it is difficult to classify loans according to purpose. For example, households may take out loans for consumption purposes using real estate as collateral. In some contexts (e.g. national accounts) the relevant distinction is between loans with different purposes; while in other contexts loans would more naturally be classified according to collateral.

⁴ For a discussion of recent improvements in the interest rate statistics compiled by the ECB, see ECB (2011).

comparison is indeed difficult⁵. This should come as no surprise, however, keeping in mind the limitations to comparability even at the national level, as discussed above.

3. Ensuring accuracy of bank interest rate statistics

In the process of determining which statistical information is relevant and of analytical interest, it should be kept in mind that the usefulness of the statistical output rests crucially on the accuracy. Hence, a well-designed reporting system focused on quality is key. This section focuses on some important issues, backed by experiences acquired by Danmarks Nationalbank during the past years in the process of designing a new reporting system. In particular, the focus in the section is on the level of detail in reporting, checking of reported data and dialogue with reporters. Sampling strategies are not covered by this paper⁶.

3.1. Reporting systems and level of reporting

As discussed in section 2, the demand for detailed statistical information covering many aspects of financial intermediation, including bank interest rates, has increased in the wake of the financial crisis. However, in response to those demands, compilers of statistics need to take into account the burden which new reporting requirements put on reporters. It should also be taken into account that development of reporting systems is mostly done in discrete steps, as it takes time for reporters and authorities to implement new demands in their data management and reporting systems.

As the demand for increased detail in the output is growing, there has, at least in the EU, been a general trend towards asking reporters to provide data at a more granular level. In Denmark, we are currently implementing a new reporting system, scheduled to be fully operational during summer of 2013. Compared to the existing system, data reported in the new system will be at a much more granular level, although not on a loan-by-loan basis⁷. The main driver of the decision regarding the granularity of reporting has been the wishes of the reporting banks, as a granular reporting is closer to their internal data model than a system add-on performing aggregations. Of course, in order for an actual implementation to end up being a reduction of banks' reporting burden, a significant consultation effort is needed to ensure that data to be reported is sufficiently close to the banks' existing data models in order for the banks to be able to reap the economies of scope. An additional advantage of the more granular reporting is that it facilitates greater comparability between the banks' reporting for statistical and supervisory purposes, and also between the reporting to the authorities and banks' own internal and external reporting.

Further, regular collection of granular data has the advantage that a larger part of new data demands may be covered by the existing data collection. This means that

⁵ The ECB has recently analysed the development in bank interest rates on loans to non-financial corporations in selected euro area member states (ECB, 2012; pp. 35-39). Among other things, the analysis concluded that different financial structures are a major explanation for cross-border differences in interest rates.

⁶ It should, however, be noted that in a European context discussions regarding measures of sample coverage are ongoing. In addition to the coverage measured by fraction of outstanding balance covered, representativity may be an issue to the extent that, for instance, larger banks are oversampled. Recent work in the ECB Technical Expert Group for the interest rate statistics aims at assessing the sensitivity of the aggregate interest rates to assumptions about the interest rates for non-reporting banks, see Hueriga et al. (2013).

⁷ The new system is covering bank interest rates as well as balance sheets. Securities will be reported on a security-by-security basis while loans and deposits will be reported at a (slightly) more aggregate level. However, at the same time, Statistics Denmark and Danmarks Nationalbank are currently in the process of setting up a new annual data collection covering large parts of the balance sheet of households. This requires information from mortgage banks at a loan-by-loan basis, including the identification of the counterparty, for the purpose of enriching the data with socioeconomic information from official registers.

compilers of statistics do not necessarily need to collect additional data when new demands from users arise; which is of course a big advantage in terms of keeping the burden for reporting institutions reasonably low. In other words, a granular reporting system increases the probability of being able to answer future questions which are not known at the time of setting up the reporting system – assuming that the dimensions of interest have been subject to sufficient quality control throughout the period for which data would be used.

3.2. Data checking and verification

An integral part of ensuring a high quality of the statistical output is the process of checking and verifying reported data. With large and increasing volumes of data and limited resources for verification, efficient data checking becomes increasingly important. The purpose of data checking is twofold. First, data checking should be aimed at ensuring that published data is accurate. And second, the checking procedure should be a tool for the compiler to ask the right questions in order to gain an understanding of the reasons for the developments in published aggregates. As a response to the increasing data volumes recent trends within optimization of data checking is in line with a macro (or top-down) approach in which the observations contributing the most to the development in the aggregate figures are the primary subjects for checking, bearing in mind the moving target nature of relevant aggregate figures⁸.

An intuitive approach to data checking is to focus on the individual observations which have the largest impact on aggregate developments. Changes in aggregate interest rates compared to previous periods may be the result of either a change in interest rates at the reporting banks (interest rate effect), a change in business volumes at individual banks (weight effect), or both. However, focus on accurate micro data is also increasing. For instance, one reason is the analytical demand for distribution measures utilizing more information than simple aggregates. In addition, where bank level data is used for analytical purposes such as financial stability analysis (if legislation allows such use of data), the accuracy of micro data cannot be disregarded. Hence, the optimal data checking procedure consists of a combination of macro and micro checks where the compiler's attention is drawn to observations which (1) have an impact on the aggregates because of changes in interest rates or weights, or (2) seem to warrant a further explanation because of the development at the bank level, while not necessarily having a large impact on aggregate figures.

Data checking systems are mostly – though not always – focused on explaining changes. For example, outlier identification is often based on the development since the previous period, or sometimes with a longer time horizon. In principle, it should be sufficient to focus solely on changes if a detailed verification of levels has taken place previously. However, since some, possibly erroneous, developments may occur slowly over many periods it is crucial from time to time to check the levels and consistency of the overall reporting by individual banks. Input useful for such a consistency review may come from comparisons of levels among banks, comparisons with data reported by the bank for other purposes, additional breakdowns provided by the bank on an ad-hoc basis, or data from other sources such as annual reports of banks and advertised interest rates.

3.3. Dialogue with reporters

A collaborative dialogue with reporters is important for efficient cooperation. The relation should be focused on quality assurance. Reporters should of course be allowed to correct errors without 'punitive' measures; while at the same time aiming at providing correct information in the first place. Since contact persons at reporting banks are often technical specialists it is also important to ensure that they, to the

⁸ See Drejer (2013) for a discussion of data checking terminology and techniques.

extent needed, consult with other departments in order to adequately answer enquiries which are more focused on business rationales than technical issues.

Reporting is by some banks seen basically as a burden which does not contribute with any payoff *per se*. In addition, compilers' focus may be more aimed at reporters providing data of a clearly poor quality; whereas the incentive to improve the quality of an already reasonable reporting may not be as strong. One way to increase awareness of reported data within the bank is to provide a detailed dataset relating developments at the individual bank to relevant benchmarks for the whole population or subgroups thereof. Such a concept of providing individualized reports to reporting banks has been in operation at Danmarks Nationalbank during the past few years, using data from balance sheet as well as interest rate statistics. Banks' interest rates on loans and deposits are benchmarked with relevant aggregates and distribution measures for the whole population as well as for specified benchmark groups, with the possibility to trace developments over time. The report has been set up in such a way that it enables reporters to see their contribution to the aggregate development. This could further draw the bank's attention to possible reporting errors.

Although substantial start-up costs are associated with the implementation of such a concept of providing individualized reports to reporting banks, the Danish experience is that it is well worth the effort. The set-up can be programmed and automated to a large extent, thus minimizing maintenance costs. One further advantage of the set-up used in Denmark is that the individualized reports have been addressed to the bank management in addition to the department responsible for reporting. This increases awareness of reporting within the bank and provides the bank with an additional incentive to report accurately.

4. Concluding remarks

This paper discusses two necessary conditions for high quality interest rate statistics, namely relevancy and accuracy. In a world in which data demanded for analytical purposes are continually increasing, the balancing of those two conditions is ever-important. It may, at least to some extent, be possible to design a reporting system suited to respond to future needs which are unknown at present. However, the discussion in the paper indicates that quality in itself is a dynamic concept, a fact which implies a continuous balancing of relevancy and accuracy. While a granular reporting system may provide the compiler with a large number of possible combinations of breakdowns, it is more important than ever to ensure that all combinations of aggregates which are published or used for analysis are sufficiently accurate.

References

- Andreasen, B. L., P. L. Kramp and A. Kuchler (2012), The Banks' Interest Rates, *Danmarks Nationalbank Monetary Review*, 4th quarter 2012, pp. 71-84.
- Drejer, P. A. (2013), *Optimizing Checking of Statistical Reports: Conceptualized Experience from a Central Bank*, BIS Irving Fisher Committee on Central Bank Statistics Working Papers, no. 10.
- ECB (2011), Keeping the ECB's Monetary and Financial Statistics Fit for Use, *Monthly Bulletin*, August 2011, pp. 63-75.
- ECB (2012), *Monthly Bulletin*, November 2012.
- Huerga, J. and L. Steklacova (2008), *An Application of Index Numbers Theory to Interest Rates*, ECB Working Paper no. 939, European Central Bank.
- Huerga, J, S. Pérez-Duarte and J.M. Puigvert-Gutiérrez (2013), *Quality measures in non-random sampling: MFI Interest rates statistics (MIR)*, European Central Bank Mimeo.