



EUROPEAN CENTRAL BANK

EUROSYSTEM

Developments around AnaCredit

Frank Mayerlen

21 June 2022

Item 3.2
Dialogue with the banking industry on ESCB statistics



Overview

- 1 AnaCredit relevance and usage
- 2 Developments in the data quality approach
- 3 What have we achieved in terms of data quality improvements?
- 4 Data quality next steps

1. AnaCredit relevance and usage

- **A dataset of very high relevance in recent crisis situations**
 - AnaCredit has proven its value as it allows **flexible aggregation and analysis to address urgent questions without going back to the reporting agent**
 - **Covid:** analysis of **exposures to certain industries**
 - **Climate:** rich source in developing a very **initial set of statistical indicators**
 - Further increased value through **horizontal integration with other** granular data, e.g. securities issues and holdings

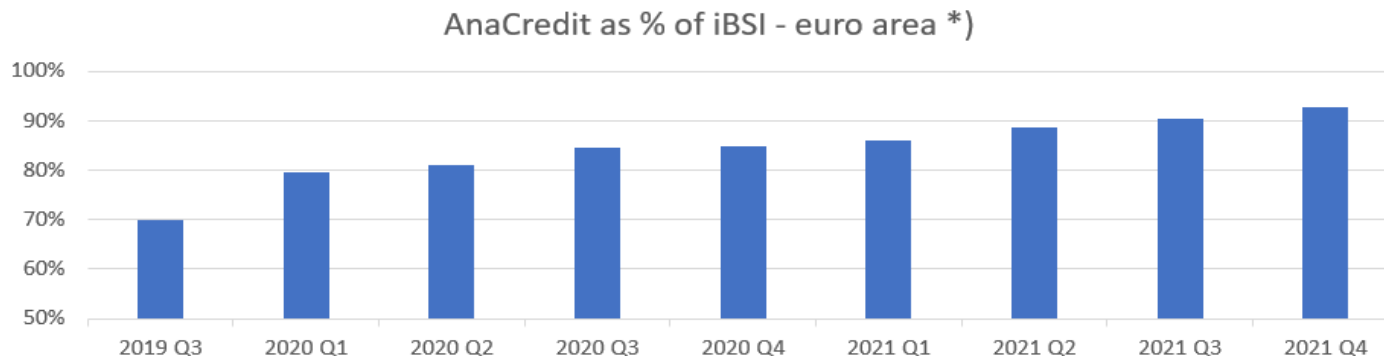
2. Developments in the data quality approach

- **Adding an external plausibility check – iBSI benchmark comparison**
 - Addressing the problem that **AnaCredit would still not sufficiently match the respective aggregated balance sheet data** as reported by individual banks (iBSI)
 - **Detailed check published** in March 2022^{*)} to ensure transparency also with reporting agents
 - Benchmark comparison facilitates the check whether AnaCredit coverage and its match with existing datasets is sufficiently good; **implicitly this checks many data quality dimensions (volume driven)**
 - Once coverage problems are well identified and (on the way towards) being resolved, we put focus on **more granular attributes**

^{*)} https://www.ecb.europa.eu/pub/pdf/other/ecb.220311.AnaCredit_external_plausibility_checks~e622cae8aa.en.pdf

3. What have we achieved in terms of data quality improvements?

- The iBSI check over time as an example



- This check is verifying the **grand-total, no breakdowns** but shows clear **evidence for improvements in covering the relevant loans**. Remaining problems of this high level iBSI check are largely identified and being worked on;
- However, **data are not yet considered complete and stable enough to build publishable aggregates or timeseries**; this will require more work

*) iBSI sample of the 250 largest banks
(covering about 75% of BSI), end-quarter data

4. Data quality next steps

- **2 broad dimensions well prioritized and concentration on more recent data**
 - **More detailed AnaCredit attributes most relevant to our users** such as e.g. interest rates, probabilities of default, defaults
 - Development of a **further high-level benchmark check vs FINREP**
 - Back data revisions should **focus on the last 12 months** (from the moment the erroneous data were identified)