

Sabine Lautenschläger: Central bank statistics – moving beyond the aggregates

Speech by Ms Sabine Lautenschläger, Member of the Executive Board of the European Central Bank and Vice-Chair of the Supervisory Board of the Single Supervisory Mechanism, at the Eighth ECB Statistics Conference "Central bank statistics: moving beyond the aggregates", Frankfurt am Main, 5 July 2016.

* * *

Ladies and gentlemen,

I'd like to warmly welcome you all, both on behalf of the Executive Board of the European Central Bank (ECB) and personally, to our Eighth biennial ECB Statistics Conference. I am very pleased to see that so many distinguished participants from all over the world have gathered here in Frankfurt today.

The title of this year's conference is "Central bank statistics: moving beyond the aggregates".

Beyond the aggregates lies a whole world of facts and observations that feed into granular statistics – namely, statistical data that need to be collected, checked, disseminated and, eventually, analysed by policymakers.

In fact we would like, this year, to discuss with you the paradigm shift that central banking statistics are currently undergoing: the move from macro to micro statistics, or from aggregate to granular statistics, if you like.

That move is nothing short of a big bang for central bank statistics. In a sense, it resembles what happened in the 1930s, when the Great Depression moved the focus of economic analysis from equilibrium to fluctuations and downturns, eventually fostering the development of national income accounts.

Similarly, in the aftermath of the recent financial crisis, we saw a huge and sudden increase in heterogeneity, or "fragmentation", in several dimensions: across and within economic sectors, geographic areas and market segments. We have learnt that aggregate statistics, although of high quality and internationally standardised, do not suffice any more as a basis for good decision-making. What we need are high-quality and timely granular datasets and indicators. I will elaborate more on this in a minute.

So, once again, big changes in statistics have been triggered by dramatic economic events which challenged the status quo and added to policymakers' demands for information.

Moving beyond the aggregates – impetus and initiatives

For the European System of Central Banks (ESCB), there are at least two reasons for moving "beyond the aggregates".

First, we had to change the way in which we fulfil our mandate, that is, maintaining price stability through monetary policy. Following the dramatic change in the economic environment since the onset of the financial crisis, we have, for instance, adopted a number of unconventional monetary policy measures, such as targeted liquidity provision and an asset purchase programme. These measures are quite different from the traditional approach of setting an interest rate. As you can imagine, conducting this kind of *unconventional* monetary policy is rather difficult when decisions have to be taken on the basis of conventional data, i.e. traditional aggregate statistics. Mitigating systemic risk in very turbulent times on that same basis is equally difficult.

Second, the ECB has been assigned additional functions as a result of the crisis. We became responsible for banking supervision, of course, but we also received a mandate

on macroprudential policy. In addition, the ECB now has to support the statistical needs of the European Systemic Risk Board.

Not surprisingly, these new functions and new data users have led to new challenges and a remarkable development in the field of statistics: in the last few years, the ESCB has enriched its datasets in several dimensions. Let me recall some of the most recent ones in the area of granular data, which is the theme of tomorrow's conference.

As a result of the money market statistical reporting, we started collecting in April this year up to 35,000 transaction-by-transaction data on a daily basis from 52 large banks in four different segments of the euro money market.

We have gathered security-by-security data on issuances and on corresponding holdings by euro area residents. Worldwide holdings of securities by 26 individual banking groups based in the euro area are also available, and are planned to be extended to all banking groups under the ECB's direct supervision by 2018.

Using the AnaCredit dataset, we will provide policymakers, as of the second half of 2018, with a considerable amount of harmonised loan-by-loan information collected from all euro area banks by the respective national central banks.

Credit data have of course a special role for the ECB. The euro area is a bank-based economy: loans are the main source of financing for companies and almost the only one for small and medium-sized enterprises. Symmetrically, loans are an important asset in banks' balance sheets. Therefore detailed and high-quality information on credit is essential both for monetary policy and for financial stability.

This is why I am convinced that AnaCredit will play a central role in supporting our key central bank functions.

So, as you can see, within the ESCB we are quickly moving "beyond the aggregates", and I am sure that this move will bring huge benefits. Let me explain why.

Lost in aggregation – the benefits of granular data

By asking for aggregate data, which is pre-organised and aggregated by the reporting agents or by the national central banks, we miss lots of valuable information. After all, it is not only the average that matters, but also the underlying distribution. And in order to analyse the distribution we need the "basic" (granular) data.

Let me give you a simple example. Assume that we see credit to businesses accelerating in a given country. We can think of several different developments underlying this "aggregate" fact.

It might be that solid companies, with low debt and good economic perspectives, are taking out more or larger loans. Or it could be that fragile and highly indebted companies are borrowing more or are restructuring their debt just to survive. Credit might also be growing because more companies have access to bank financing. In turn, this might reflect better economic prospects and a greater appetite for investment – which is good – or just a deterioration of credit standards – which is not so good.

All these possibilities (and I didn't mention the possible factors on the credit supply side!) are consistent with growth in aggregate credit. But, clearly, they might have very different implications in terms of the monetary policy stance and risks to financial stability.

Granular data will help us look beyond aggregates and reveal the underlying developments. By having granular loan-by-loan data we will know, abstracting from the respective identities, the characteristics of specific groups of counterparties involved in each transaction. Then we can assess the "driving forces" behind any aggregate development and distinguish genuine and "healthy" growth from potential bubbles. This is very important for policymakers.

And the benefits of going granular also extend to those who have to provide the data in first place: the reporting agents. After the initial high set-up costs resulting from the large volumes, reporting information at a granular level brings significant savings. This is because statistical requirements are going to be more stable over time and *ad hoc* data requests for special or urgent policy needs are minimised.

Also, the more we ask banks to provide information which resembles what they already have in their internal systems, the easier it will be for them to comply with the reporting requirements.

Last but not least, in those countries where feedback loops are established by the respective national central bank, reporting agents will benefit from much more complete and harmonised information on the creditworthiness of their counterparts, in particular when they are foreign residents.

On our side, even when policy decisions are taken on the basis of aggregate statistics, as is usually the case, moving towards granular data offers the big advantage of timeliness and flexibility: raw information can be organised and aggregated in different ways depending on the specific policy question at hand.

This is key, given that it takes three to five years to develop new aggregate statistics, while policymakers usually want the data “yesterday”. Granular data allow us to provide the necessary statistics in a flexible and timely manner – so probably not “yesterday” but not too far from “today”.

Allow me, for a moment, to compare the compilation of statistics with the making of a cake, say, a Sachertorte. Going granular is like asking banks for the basic ingredients such as chocolate, cream, flour, eggs, butter and sugar instead of asking them for a slice of Sachertorte.

Our clients – policymakers, analysts and researchers – are always hungry, and they need different cakes depending on the policies they have to decide on.

Once we have all the basic ingredients, we can bake a variety of cakes just by combining these ingredients in different ways. So we don't have to go shopping every time, which makes it easier for us, and which makes it easier for those who have to provide the ingredients, that is, the banks.

New challenges ahead

Ladies and gentlemen, I hope that I have convinced you all of the benefits of moving “beyond the aggregates” – including those of you who do not fancy baking cakes. I should add though that granular statistical information also poses new challenges.

First and foremost, granular datasets need to be standardised and well integrated. We have to prevent any build-up of disconnected “silo” data collections. We should no longer request the same data twice, and whenever we cannot avoid requesting similar information twice we have to explain why. Once collected, we have to make sure we can combine the data from different granular datasets as necessary.

Second, granular datasets must be multi-purpose: they need to serve different users and support different analyses, all at the same time. AnaCredit is a great example of this; it is designed to support different key central bank functions.

Third, keeping in mind the need for integration and versatility, we must provide financial institutions with a unified reporting framework – a framework that is based on a consistent and stable set of rules and definitions. This is precisely what we are doing with the current work to define an ECB Single Data Dictionary and an integrated European Reporting Framework.

Fourth, we need to adopt and manage state-of-the-art IT solutions that can process huge data volumes and ensure the necessary data confidentiality.

Last, but not least, I have already said that granular data allow us to satisfy our customers' craving for "cakes" of all different kinds. So the remaining question is: can our customers digest all the cakes they request? Can they handle large sets of granular data? I am pretty sure they can!

So, as you can see, moving "beyond the aggregates" entails quite a few challenges, but I am confident that we will master all of them.

Conclusion

Ladies and gentlemen, we are currently in the middle of a paradigm shift in which granular data is rapidly gaining relevance.

And one thing is obvious: good statistics – both micro and macro – are essential tools for policymakers, helping them to take decisions and to assess the impact of those decisions on the economy.

But let's be realistic. Granular information will not save us from any future financial crisis. It will, though, put policymakers in a better position to mitigate the risks *ex ante* and to limit their potential impact *ex post* by taking the appropriate corrective measures and by monitoring their effectiveness in a timely manner.

Besides offering extraordinary opportunities, this move beyond the aggregates brings big challenges, as I have already said.

New questions arise. For instance, will granular data ever replace aggregate statistics? Let me put it another way: is it possible to reduce and simplify the collection and production of aggregate statistics once sufficient individual information is available? Or do we need to collect and compile both granular and aggregate data for the purpose of cross-checking?

These two kinds of statistics nicely complement each other, in my view. Micro data can be used to fill important data gaps. At the same time, aggregate statistics provide a useful benchmark for granular data, especially while these are still at an early stage of development.

Nevertheless, keeping the reporting burden to a minimum is an important priority for us, therefore micro and macro data collections as we know them now cannot continue in parallel forever. As statistical reporting inevitable creates costs for third parties, we will regularly reassess our requirements, taking account of the information needs of policymakers, and at the same time constantly striving for the best possible balance between merits and costs.

I am very confident that the conference tomorrow will help us to find the right answers to these and many other questions relevant for the future of central bank statistics.

Thank you very much for your attention.