
IFC Satellite Seminar on “Granular data: new horizons and challenges for central banks”

Granular statistical data –
Are we charmed with a blessing or a curse?¹

Jasminka Dodeva Kovaceva,
National Bank of the Republic of North Macedonia

¹ This contribution was prepared for the IFC Satellite Seminar held at the ISI 64th World Statistics Congress, co-organised with the Bank of Canada in Ottawa, Canada, on 15 July 2023. The views expressed are those of the authors and do not necessarily reflect the views of the Bank of Canada, the BIS, the IFC or the other central banks and institutions represented at the event.

Granular statistical data – Are we charmed with a blessing or a curse?

Jasminka Dodeva Kovacheva¹

Abstract

The past has shown that granular statistical data very often comes at the forefront for policy makers. This is particular the case, during or after crisis episodes, when new sources of risks are detected. Individual level data can provide punctuality, clarity and timeliness, as surveys data or aggregated administrative data cannot. However, what is important is how to determine the limit: where is the border between the usefulness and the responsibility for data producers and users; where is the border between the benefit and the burden for data reporters. In this presentation, we investigate the importance behind the collection of granular statistical data for Central banks and the obligation it creates, by tackling the specific experience of the National Bank of the Republic of North Macedonia. The presentation will showcase the granular statistical data we use, how we try to converge from history to new ways of data collection without losing the quality, and how we mitigate the risks associated with managing and using big and detailed databases. There is always a need and a challenge of accepting and introducing new data sources, but there is also a need to preserve what is established, useful and broadly accepted in the country. And the trick is to balance it all and find what the best option is. In the presentation, we show what we think the best option is for our case.

¹ Email: DodevaJ@nbrm.mk, corresponding author. National Bank of Republic of North Macedonia.

Contents

1. Introduction.....	3
2. What is granular data and how detailed data should be?.....	4
2.1 What works for us?	6
3. Cost and benefits of data granularity.....	14
4. Data management.....	15
5. Final remarks/conclusion.....	16
6. References.....	17
Annex 1: Detailed data included in the reporting form for loans	19

1. Introduction

The barriers of statistical research have a heavy duty of providing timely, quality and compatible data so that end-users, such as analysts, policy makers and private sector can make most out of it and can embroider the data into sound decisions for better future. This is an obvious task. However, what is not obvious and is more complex is how to put the theory into practice.

The producers of statistical data are facing many challenges and pressures of responding to the ever-increasing demand of high volume data. Reasons behind this are many, globalization and digitalization being the leaders of the changes in the last few decades. New financial instruments are continuously emerging that could pose higher threat to the financial stability and are putting pressure on defining new regulations. Multinational companies that have trading agreements all over the world that are difficult to capture. Special Purpose entities with activities that are scarce but complex. Economic Unions that provide free, completely liberalized, movements. There are many opportunities offered for economic transactions and all of them have to be statistically captured in order to gather the facts and to provide a valuable insights of the economic (and social) reality. Granularity of data can provide these insights and give answers to many questions; however, there is always a challenge to determine what is sufficient and where is the limit of the granularity. When the granular data becomes a burden instead of an advantage.

In addition, new economic reality provokes introducing new data sources and limiting the “old” ones that are in many cases connected with or originated from certain economic and financial regimes in the countries (trade regimes, restricted financial transactions, pre-approvals and conditions). However, the peculiarity is for the institutions responsible for gathering statistical data to investigate and implement or adopt the combination of both. The need of statistical data should not pose a limit or a delay in the financial transactions, or in the free movement of goods and services, but in the same time, we should not lose the quality and the details of some of the already established statistical systems and cause additional burden to the reporters.

The question of the level of detail of data is something that we will try to cover thought-out the paper and we will try to answer, in certain degree, what works best.

This paper aims to contribute to the ongoing national and international discussions of the relevance of granular statistical data. In addition, to shed light on the risks involved with overburdening the reporters, overflow of information, as well as the responsibility of the data management especially with preserving the individual data and their usage only for statistical purposes.

The paper is organized as follows. The first part covers the definition of granular data and the dimensions it comprises, with a separate part capturing the details and practical aspects in the case of the National Bank of the Republic of North Macedonia. The second part provides insights of the cost and benefit of having detailed data. The third one gives an overview of the need for a complex data management and the ways to mitigate the risks. The last section concludes.

2. What is granular data and how detailed data should be?

In the last decade, the phrase “less is more” has become very popular in many different areas. We have minimalistic style in architecture; design and art in general, used to create larger effect. In addition, it is used very often to allow people to focus on what matters most and not to be distracted by the “irrelevant”. An oxymoron that it is based on the value of simplicity and became a way of creating a life of more.

However, this phrase does not apply in the field of statistics. In statistics, less data means less information for analysts, policy makers and other data users to make solid, data driven decisions. It means less information for quality crosschecks, data comparability and resolving asymmetries. Therefore, it means less quality data and “blind” decision-making. We can freely say nowadays, in statistics “more is more”, and lately we are witnessing a trend of “everybody going granular”.

We see international and country initiatives of enchanting the databases with more details in very different areas, just to name a few: OECD initiative on Municipal Migration Database; or AnaCredit (Analytical Credit Dataset) project set up by the ECB. In addition, on a global level many initiatives are arising for creating unified database for different identifying elements, for example the introduction of the Legal Entity Identifier (LEI) that enables clear and unique identification of legal entities participating in financial transactions or the IMFs initiative for Annual Exchange of Sector Information on Securities Issuers (ESI).

With the rapidly changing nature of how the world works with the globalization and digitalization the need for more granular data is essential in order to understand its complexity and the new phenomenon, new instruments arising.

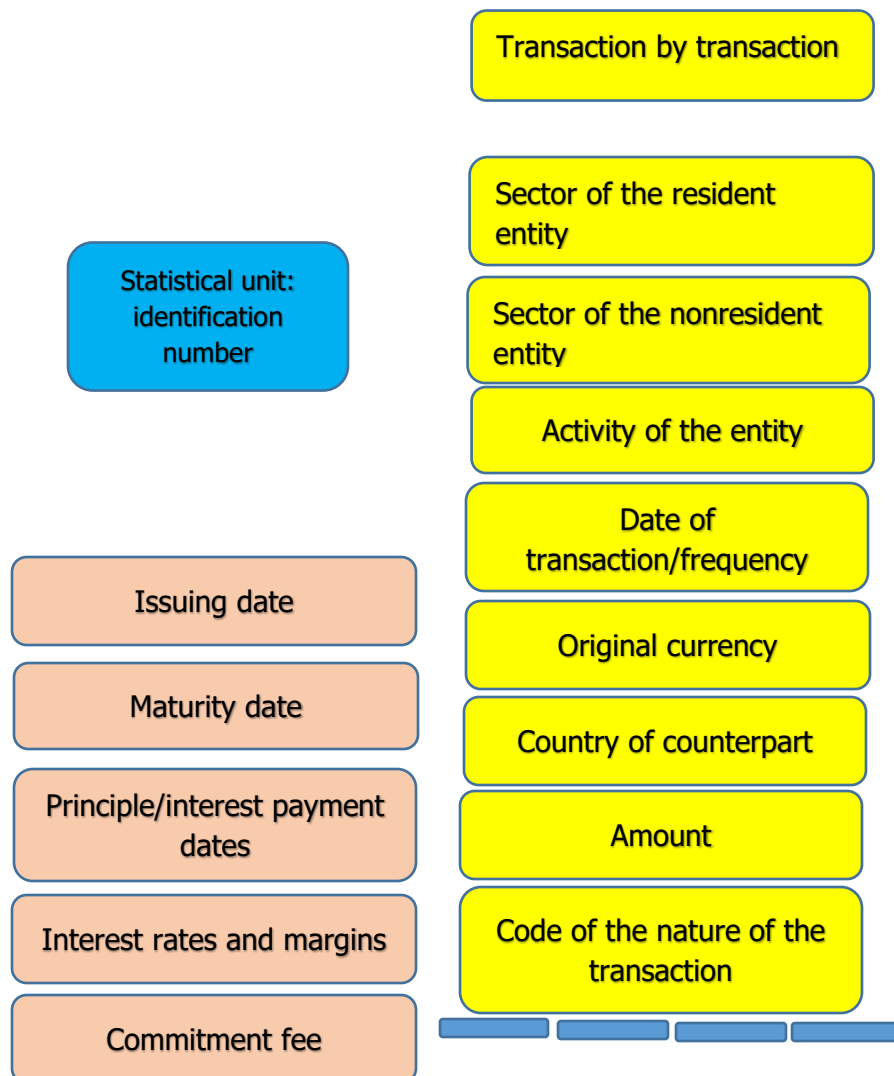
With globalization and multi-international companies gaining pace, the need for data increased even more. The need to understand the international transactions became dependent from the detailed data and the ability to understand their complexity through data linkages. Without having the proper information of how these transactions are conducted, which are the involved parties, what is the type of the transaction and what is their ultimate goal it is difficult to estimate their effect on the economy and how those transactions would influence other entities and instruments. Only granular data can reveal the nature of the integration of the companies in the global economy.

The importance of data is unquestionable, but what is more important in statistical terms is what is considered granular data and how detailed granular data should be.

Granularity is the amount of information contained within the data structure. Granular data usually involves micro-data or data broken down across multiple dimensions of information. In statistical terms, granular data usually comprise data on a level of transaction, instrument and/or account collected on an individual statistical unit.

The level of detailed data is determined primarily by the statistical requests of data, according to the internationally accepted methodology, and secondly, by the demand for additional information for a sound decision-making. Just for an illustration, Picture 1 shows the most common information contained in the granular *external* statistics data set.

Picture 1: Detailed external statistics data



As presented on the picture, the first step is gathering data on transaction-by-transaction level and the identification of the statistical unit through its identification number. Second is to attach all information available from all data sources. In general, this information covers the characteristics of the underlining financial instrument or economic transaction. Most common information of each instrument are: the value of the transaction, sector of resident and non-resident, activity, date of transactions (frequency), original currency, country of the counterpart and code capturing the nature of the transaction. Depending of the data source these dimensions can be broaden with a lot more detailed data, for example data for loans can include information of date of agreement, dates for disbursement, principle and interest payments, interest rate, margins, commitment or other type of fee and charges and so on. Not all instruments have the same information structure, however it is important to determine the ones that are essential and every data source should provide in order to have unified data structure and form integrated database.

Collecting granular data does not come without a challenge. Data producers should always be careful, because there is a significant risk of information overflow. On the other side, for data analysts there is a trade-off between the advantage of performing a micro level analysis and the risk of losing the focus of the macro perspective. The border between producing, analyzing too many granular data versus aggregated data have to be set very carefully.

2.1 What works for us?

In regards to granularity of statistical data, every country must pinpoint what is important for its specific economic circumstances. The National Bank of the Republic of North Macedonia (NBRNM), as an institution responsible for external statistics data, monetary statistics data and financial accounts data, have long established tradition in collecting data through various data sources. Data are collected through administrative data sources; direct reporting or estimates are used. All data have different structure, which pose a challenge to create one integrated database. The intention is with each newly introduced data source the necessary data structure to be ensured.

The standards for granular data, which are used, are based on the standards agreed and implemented on a global level. These standards result in a large efficiency gains for both data providers and users. Standardization leads to easing the reporting of detailed data, especially because in that way many data reporters could include the reporting obligation within their IT system.

Data sources that the National Bank use are designed to obtain specific information to fulfill the methodological requirements. Many of the direct reporting are relatively newly introduced in order to capture the financial instruments that are gaining importance in recent decades and to provide information according to the up-to-date statistical methodology. On the other side, some data sources have been historically established and only certain changes were implemented. With every introduction of new data source, it is very important that timely and effective interaction between reporters and drafters of statistical reporting regulation exists. Collaboration between public and private sector, as well as public-public collaboration in the case of administrative data, remains the best approach for creating integrated system for statistical reporting.

If we analyze more specifically the data sources used for external statistics data, recently two trends have emerged and have taken pace. With the process of cross-border liberalizations, many data sources have become insufficient to provide full quality data and to capture the economic reality, and even some data sources were completely abolished. In the same time, the process of liberalization created an environment where the statistical reporting have to be limited in order not to be a burden to the realization of the financial transactions. In some cases, this led to decreasing the quality of the reported data and in addition decreasing the possibility of conducting detailed quality crosschecks.

The data sources that the NBRNM use have different structure, containing different level of details. When introducing new direct reporting, the process of defining the information that it would contain is twofold. First, it is essential to start from the international statistical requirements for the specific type of statistical data. Second, how detailed the data source should be also depends on the current relevance (in value terms) of that instrument for the country and its

future prospects. Very often those instruments that are most relevant usually are covered more extensively (in a sense of frequency, level of details, periodicity). What is also very important is enabling an automated process and the possibility of the reporting to be integrated in the IT systems of the reporters, and if possible to be used for other purposes as well (e.g. connected with their accounting, or data that are used for supervisory purposes as well). What is also very important is determining the reporting unit and whether there is a possibility of indirect reporting, or reporting through different institution for e.g. regulators, in order not to burden the units with double reporting.

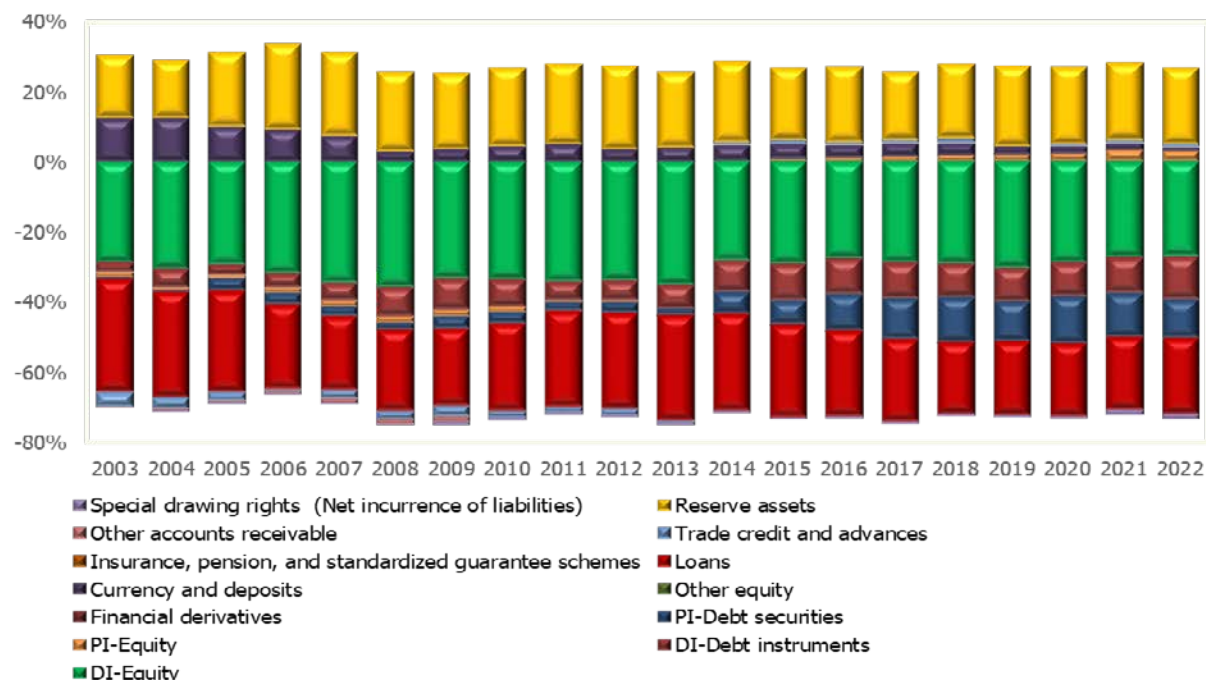
Having this in mind, we have established, or enhanced the historically inherited, several granular data sources. In the following paragraphs, few of the data sources used in external statistics will be elaborated in more details.

Loan by loan

When analyzing the International Investment Position (IIP) of the Republic of North Macedonia it is very notable that “traditional” financial instruments, such as *Loans*, are still one of the most important instruments for foreign financing compared with portfolio investments or financial derivatives (Graph 1). If we add the part of the loans, which are included in the direct investment-debt instrument component, then the importance is even higher. This is to be expected having in mind the status of the country as a developing country.

Graph 1: Structure of International Investment Position of the RNM, by instruments

(in %)

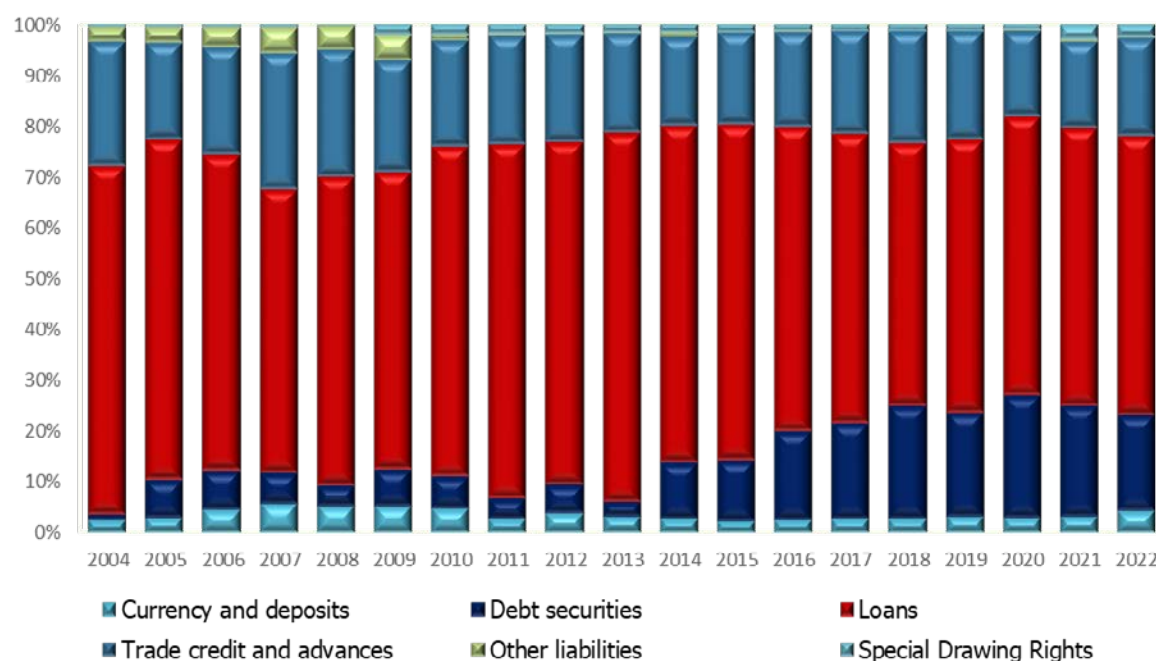


Source: NBRNM

If we analyze the debt instruments within the Gross External Debt Position of the country, the percentage share of the Loans is ranging from 68.6% in 2004, to 72.6% in 2013 and 52% in 2018, which shows relatively constant level over the years.

Graph 2: Structure of Gross External Debt, by instruments

(in %)



Source: NBRNM.

In this environment, the focus is very much on analyzing the effect that this instrument have on the economy and future financial transactions in regards to new disbursements or repayments of debt. The detailed data on loans are used mostly for assessing the ability to service outstanding debt and estimate the future external borrowing, as well as for making economic projections and analysis of a country's overall economic position.

In addition to this need, detailed data on loans are required by international institutions, e.g. by the World Bank for fulfilling the obligations of Debtor Reporting System (DRS). Even though the requirements for loan-by-loan information within this system is only for the public sector, we have applied the same concept for all reporting units. Therefore, the loan component, in our reporting system, is covered at the most detailed level.

According to the regulations (Law on Foreign Exchange, Decision on the manner and the terms of recording and reporting on the concluded credit operations and Manuals for the manner and the conditions for individual reporting of loans taken by residents), each entity that borrows or extend a loan from/to a non-resident have to report that loan to the National Bank within 10 working days after signing the loan agreement. All changes in the agreement and all realized transactions have to be reported within 5 working days. The level of reporting is quite in details,

each reporting set contains several forms: one form is for the general terms of the agreement, second form is for disbursements, third type of form is for schedule of principle payments and the forth is specification in case of multiple creditors in the loan agreement. The most detailed document is the first form. This form contain several information about the loan agreement, as presented on Annex 1.

As can be seen from the picture, all information are on a loan-by-loan basis for each resident, containing information that further provide the basis for additional calculations and deriving information on future schedule of interest payments, stock and flows of accrual interest, arrears, residual maturity. In addition to this, several codes for the type of transactions are applied when recording the disbursement or payments of principal and interest. With this information appropriate classification of the flow within the integrated International Investment Position is ensured, as well as for other statistics for e.g. Financial accounts. All basic elements collected provide a solid ground to create a whole set of information that can be utilized differently, according to the needs.

Historically analyzed, this data source was first introduced because of certain regulatory bans on the movement of the capital to and from abroad, which are usually covered by the term “foreign exchange regime”, defined in the Law for foreign exchange operations, but also in order to fulfill the international statistical requirements. Currently, most of the capital transactions, such as credit operations, direct investments, investments of non-residents in securities and in real estate are completely free. Therefore, the recording of loan transactions are for statistical purposes only and it does not require submitting any supporting documents. Having this detailed data provide rich possibilities for different analyses, comparison with other data sources in order to increase the quality, fulfilling the international methodological standards and quickly adopting to the changes in the recommendations.

The ownership of such database very often creates a challenge. This is especially in cases such as classifying personal information of the reporters; requests from other institutions to obtain detailed data for not statistical purposes; occasions where as a convention in order to ensure symmetry among countries certain classifications should be re-located (e.g. “all interbank positions other than securities and accounts receivable/payable are classified under deposits”² even though the loan agreement clearly testify otherwise). In addition, as mentioned earlier, maintaining big systems that process and calculates detailed level data on a daily basis, require sophisticated technology and application, and advanced IT human resources.

The decision to collect detailed type of data reports for the Loan component might be seen as too cautious or prudent, but the importance of the instrument on the overall perspective of the economy overwhelms the challenges and drawbacks.

International Transaction Reporting System (ITRS)

Another data source that requires closer attention (not only because of its importance but also for its controversy) is the “historically inherited” International Transaction Reporting System (ITRS). This data source have been originally inherited from the former system in which very rigid foreign exchange regime existed, with clearly defined regulation of all inflows and outflows

² Balance of Payments and International Investment Position Manual (BPM6), page 87, 2009, IMF.

from/to abroad. One of the main reasons behind this was the need to sustain the stability of the domestic currency, through controlling and approving the cross-border transactions. Throughout the years, this database full of information became less important for transaction regulation, but became more important for statistical purposes, as one of the main data sources for Balance of Payments compilation (BoP). The system was additionally enriched with separate codes for each type of transaction in order for clear classification within the BoP.

However, with the process of globalization and liberalization, free movement of capital, growth of multinational corporations, and development of foreign financial markets and the surfacing of new financial instruments have led to questioning the sustainability and justification of the ITRS as a statistical data source with adequate coverage. In addition, the difference between the methodological statistical guidance versus the coverage of the data source (e.g. accrual vs payments) have become even broader.

Within the EU, the liberalization of the transactions and the pressure for decreasing the expenses and reporting burden of deposit-taking corporations, as well as speeding the process of realization of each cross-border transaction, led to adopting the EU Regulation 924/2009 on cross-border payments, by the European Parliament and the Council of the European Union. With this regulation, it was foreseen that from January 1st, 2016, the deposit-taking corporations within the EU Member states would not conduct classification of the cross-border transactions for the need of the statistical systems. However, the regulation allows the "Member States to continue to collect aggregated data or other relevant readily available information, provided that such collection has no impact on the straight through processing of the payments and can be fully automated by payment service providers"³.

The ITRS that the National Bank uses is a so-called "closed system" in which full integration is achieved between stock at the beginning of the period, transactions and stock at the end of the period. It provides data mostly on a daily basis, in original currency. In the last two decades, the National Bank has made significant improvements in the coding list describing the nature of the transaction, in order to be aligned with the international classifications of the transactions and in order for the residents to have easier selection of the code when classifying the transaction. In addition, the National Bank has invested largely into the IT infrastructure for providing fast, easy and secure electronic transfer of data by the banks.

With all being said, how useful is this data source and what kind of data it provides? As mentioned, the reporting provides data on a transaction-by-transaction level, mostly on a daily basis, in original currency of the transaction, without applying any threshold for foreign currency payments⁴. The information it contains mostly capture the ID of the resident legal entity or the code for the resident natural person, the code for the nature of the transaction, the country of the non-resident, an information about the non-resident ("free text"), the ID of the bank through which the transaction is realized.

The biggest advantages of the system are the short deadline in which data are obtained from the reporters (5 working days after making the payment); granular data obtained through census (all

³ REGULATION (EC) No 924/2009 on cross-border payments in the Community.

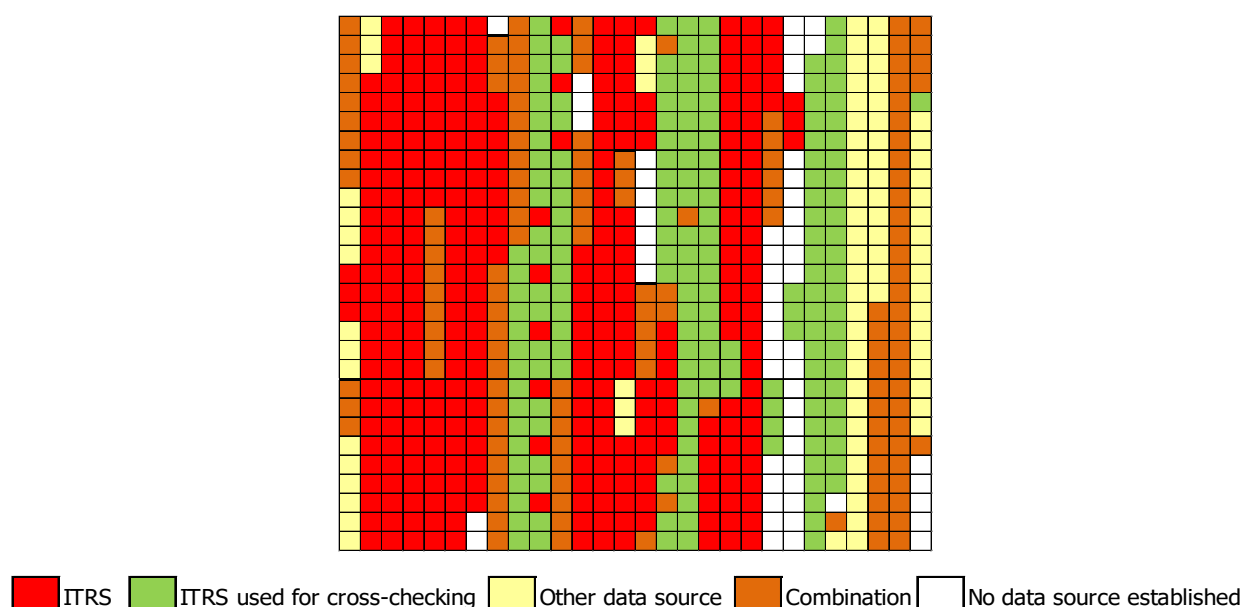
⁴ For payments in Macedonian denars, the transactions under 100.000 denars (ca. EUR 1500) are reported on an aggregated basis.

data are reported, without threshold, through total of 14 banks); relatively good data quality, classified according to the statistical requirements; the whole process of reporting is completely automatized; the responsibility of providing correct code of the nature of the transaction is on the resident receiving/sending the money; the companies and the banking sector have already established internal work procedures which is easing the whole process.

The ITRS as a data source is mainly used in the BoP compilation. In addition to being a data source, the ITRS has its application as a source of information for crosschecking of the quality of data received through other data sources. This characteristic is very much appreciated because it provides granular information like no other data source can: who is the reporting unit, when the transaction happened, and why it occurred. What is also very important benefit is providing information of the ID of the legal entities that are involved in certain type of transaction. This information is used for defining the sample of data reporters for a direct reporting (surveying through a sample of companies). These are two of the most important side-benefits from the ITRS that have a direct effect on increasing the data quality.

Picture 2: Data sources - ITRS usage in the Balance of Payments ("The Rubik`s cube")

*Each square present an item from the BoP standard presentation



Source: Authors presentation based on NBRNM data.

Some disadvantages of the ITRS that should be mentioned is the lack of compatibility with the internationally accepted methodological standards. Transactions that are based on payment ("cash") basis are not aligned with the accrual requirements stipulated in the statistical manuals. In the same time, the moment of payment does not capture the moment of change of the economic ownership, which is required for the statistical reporting. In addition, this data source does not capture transactions that do not require payments (e.g. compensation, debt forgiveness) or stock data. One cannot overlook the fact that coding of the transaction could pose delays or

could “slow” the realization of the international payments, especially in the case of inflows where the money are not available to the receiver until the appropriate code is stated. The risk of having this disadvantage should be very carefully managed and if the system is properly set, with an advance IT infrastructure within the Central bank and the commercial banks, it can be mitigated or at least significantly decreased.

Even though there are some shortcomings in this data source, one must agree that it has values that other data sources does not. The trick is to adjust it in order to incorporate the new requirements, but to keep all the advantages it provides. In cases where, as in RNM, the historical habits are so enstamp in the whole work culture of the companies and the system is already incorporated in their accounting systems, being that the ITRS also relies on accounts, is very important to seize the values and utilize the benefits it provides.

Security by security

One granular data source that was relatively newly introduced was the security by security direct reporting for portfolio investments abroad made by residents. Having in mind that portfolio investments became more significant for the country during the recent years and in an absence of a possibility for having a direct access to any global security-by-security database, in July 2014, the National Bank made changes in legislation and introduced new manual for reporting on security-by-security basis.

The reporting was made obligatory for all residents who invest in securities in foreign financial markets, and who at the time of the introduction of the new report, had the possibility to invest according to the Law of foreign exchange operations. Those were deposit-taking corporations, insurance companies, investment funds and pension funds. The reporting to the National Bank is done directly or indirectly through the supervisory agencies. Reporting is ISIN by ISIN, on a monthly basis, and includes all securities (foreign and domestic) in the investor’ portfolio acquired on foreign financial markets.

The reports include mainly detail data for:

- Investor (ID, Name, Institutional Sector);
- Security (ISIN, Currency of issue, Currency of trade, Market and Country of trading, Type of security, Issuing date, Maturity date, Type of interest rate etc.);
- Issuer of the security (Name, Country, Institutional Sector);
- Stock at the beginning of the period (Number/Market price per security at the beginning of the period);
- Transactions (Number/Market price per security of purchased/sold securities);
- Other changes (Price changes, Exchange rate changes, Other changes);
- Stock at the end of the period (Number/Market price per security at the end of the period).

With this reporting, the National Bank established an integrated system of data collection on the stocks and transactions in securities, electronically transferred to the National Bank through web application, by imposing only low burden on the reporters due to the fact that many supervisory agencies collect these data for their own purposes and then report this data to the National Bank. The system provides full reconciliation of data, better quality of data (market value, accrual basis) as well as bigger details of data (unified Data Structure Definition-DSD).

The introduction of this data source enabled the compilation and dissemination of Coordinated Portfolio Investment Survey Data (Assets), with complete breakdown of data by Currency of Denomination of the Securities Held, by Sector of Resident Holder, by Sector of Nonresident Issuer, and cross tabulation by Sector of Resident Holder, and Economy and Sector of Nonresident Issuer for Specified Economies.

With these, the RN Macedonia fulfilled the last SDDS Plus requirement for external statistics data and starting from 28.1.2019 it became a SDDS Plus adherent, together with 18 other countries at that time.

However, new challenges lay ahead. Starting from 2019, with the beginning of the second phase of the Stabilization and Association Agreement with the EU, free movement of capital related to portfolio investment was ensured. This means investors in foreign securities can also be individuals and non-financial corporations, which are not captured with the security-by security data source. Currently, these investments are monitored through other data sources (mainly ITRS) and the amounts are not that significant. However, potential data sources are being analyzed.

New Integrated Reporting System for monetary data – Ongoing Project ISIDORA

The need for granular data, mainly for statistical purposes (monetary, financial and external statistics), but also for supervisory purposes, have prompted starting a project for integrated reporting system for monetary statistics data. With the main goal of the project to align the National legislation with the European legislation, the project ISIDORA started in 2016. The intention was set at collecting balance sheet data from deposit-taking corporations (including the Central Bank) and saving houses.

Key elements that were identified as essential for collection, that are not captured with the current system, were:

- currency structure;
- geographical distribution of non-residents' instruments;
- sector classification of non-residents;
- sub-sectoral classification of residents;
- flows data, as opposed to stocks (in order to separate the influence of price and exchange rate changes on the movement of monetary aggregates, as well as other changes in volume).

Many of these dimensions represent a requirement for statistical reporting, however, in order to minimize the reporting burden, the ISIDORA project was broadened in scope by capturing also the requirements for the Credit Registry, as part of the supervisory purposes.

With the implementation of this project, the National bank would achieve a complete reorganization of the current data reporting system (from accounts to instruments described through predefined dimensions and attributes) and an integration of various statistical reports, but also integration part of the supervisory reports. So far, a lot has been done, but many more activities still await, such as development of by-laws (decisions and reporting manuals), technical performance of the systems in the National Bank and with reporters, continued collaboration and

holding meetings and workshops with reporters, testing of the system, parallel reporting and finally accepting the system as an official data source.

3. Cost and benefits of data granularity

The cost and the benefits of data granularity can be analyzed from different perspective. The decision on what prevails, the cost or the benefit, largely depends whether that perspective is from the angle of the data collector, data reporter or data user. On average most of the common benefits can be summarized in several different points.

When we consider statistical data, the most important benefit from granular data is the possibility of fulfilling the statistical methodological requirements. Even more so, the availability of microdata provide an opportunity to revise the time-series data in cases of updates in the methodology. In addition, microdata provide a good ground for increasing the quality of data by crosschecking of data between different data sources and the quality checks can be done even on single statistical unit. In addition, the analyses of data, when individual data are available, can be done on a single Balance of Payment report produced for one statistical unit, providing a very direct view of the effects that entity have on the cross border transactions, the possible pressure on the foreign exchange market or foreign reserves assets.

Lately, models for estimation have become very widely used among statisticians, especially when data are scarce or less frequent. They provide information that can easily be applied once the initial parameters are determined. However, what became very much in focus on a global level is also the large detected data asymmetries between countries. Not to neglect the importance of estimation models, but in situations when different countries have different input parameters for the estimates, it is very possible that the asymmetries would increase between those two countries. Having micro level data instead of estimates increase the quality of the analyses and ease the process of detecting the potential differences in data.

In addition, when stating the benefits we cannot overlook the fact that granular data provide an opportunity more detailed data to be published and be available for different kind of users (academia, researchers, students, officials, private sectors, media). It provide possibility of conducting multidimensional analyses and it stipulates larger creativity and determining the causality of the effects.

If we focus on the needs of the data collectors and analysts, at first glance, it is obvious that more granular data means larger possibility of it use. However, knowing the obvious does not mean we should lose sight and neglect the possible traps that the possession and reporting of detailed data can have.

When analyzing the costs, it is important to make a distinction whether we discuss about the costs of the reporters or of the collectors. One of the most common mentioned cost of microdata reporting is the burden of data reporters that it creates. Reporting microdata usually means reporting not only detailed level data but also high frequency data that can be very time consuming process. However, when stating this burden one have to have in mind that it is not

always straightforward to determine what creates more pressure microdata or aggregated data. In cases where the IT structure is enhanced and all reporting is automated and from already available databases then it is much easier to transfer the database from the reporting unit to the data collector, than to recalculate, aggregate according to certain rules. However, even in these cases certain aspects have to remain in focus of the data collectors: like making sure not to collect the same information twice or to have a clear vision of the data that are useable (or would become in use in the future) and not to overburden the reporters with excess data.

In order to decrease the reporting burden, while keeping the details of data, it is very important to ensure harmonization and standardization of data. When harmonizing the data on global level, it is very important to keep in mind the interest of all involved parties, as much as possible.

From the aspects of data collectors, the possession of a detailed database does not come without a responsibility. Maintaining the confidentiality of the individual data and having in place profound procedure of how, when and for what purpose these data can be used, is at most importance. Managing these micro databases usually implies that the institutions responsible have to be able to provide sufficient human capital, devoted to managing and preserving these data, as well as developed IT resources that can enable secure process of statistical production (acquisition, production, exploration and dissemination of data).

4. Data management

At the National Bank, the focus today is on producing high quality and granular data and in managing this high-volume of data. Granular data should be properly collected, organized, and managed to maintain its integrity. Data quality and validation processes are crucial when dealing with granular data, because errors on individual level can have a significant impact on the overall analysis.

When working with granular data it is important to adhere to privacy regulations and ensure appropriate data anonymization or aggregation when publishing in order to protect individuals' identities.

In the Republic of North Macedonia, the National Statistical System is defined in the Law of state statistics and in the five-year Statistical Programme. The State Statistical Office and the National Bank (NBRNM) are the main institutions barriers of statistical research, together with several ministries and agencies who are responsible for fewer statistical surveys noted in the five-year statistical programme. The regulatory framework consist of Law on State Statistics and the Law of the National Bank, as well as agreements for cooperation and data sharing between institutions. The legislation, among other things, covers the statistical confidentiality as well. Confidential data are data that allow direct or indirect identification of the statistical unit. All confidential data collected for statistical purposes should be protected under the law.

Special procedures are in place at the Bank that regulate the derogations in which confidential data can be disclosed. The exceptions cover requests by external auditors or regulatory authorities in order to perform their duties, court orders, or when a consent from the reporter is

obtained. Within the institutions, access to the confidential data is allowed only to the employees responsible for collecting, analyzing and managing the data. Measures are in place to ensure data protection (password protection, restricted access, encryption of data or data transfer through special application to ensure privacy). Internal procedures are also established for regular data exchange between departments within the institution or in cases of special requests.

5. Final remarks/conclusion

The conclusions can be summarized in several key points. In our complex world that we live in, having available granular statistical data is a necessity, in order to understand the economic reality. It provides insights of how our economic, financial, social dynamics works. Adding micro data one to another and piling, building more details on top of that can create the picture of the movements and structures in the country and how they are interconnected with the rest of the world, like pieces of a puzzle. Nowadays is not the question on do we need granular data, but the question is more on how to obtain them without creating additional burden, to obtain them with less effort and how to best utilize them and protect them.

What is evenly important to granular data is to have high frequency and quality data. In order to ensure that, having appropriate data sources is essential. The adaptation of different data sources depends entirely of the type of the economy-level of development and the reporting culture in the country. Even though harmonization of data is a necessity for providing similar, aggregated data structure on a global level, "cut-off models" and harmonization of data sources in different countries is not always a good choice. Each data source have some particularities that each country should assess and determine whether it would be a good fit for its case. Even though some data sources are considered "historical", "not appropriate" and should be neglected, one have to keep in mind the advantages those data sources have and the possibility that little change can ensure keeping their value and continue to use them. The focus should not always be on "one side of the coin", but to re-think and re-invent the options presented and maximize the already established granular data sources while respecting the needs of all involved parties and having in mind the necessity for ever evolving modernization in statistics.

6. References

Marques C. (2014): "How granular data should be to compile BOP statistics", IFC Bulletin No 37, January.

OECD/Statistics Finland (2021), "Measuring Trade in Value Added for Finland", OECD Publishing, Paris.

Balance of Payments and International Investment Position Manual (BPM6), 2009, IMF.

International Reserves and Foreign Currency Liquidity: Guidelines for a Template, IMF, 2013.

External Debt Statistics: Guide for Compilers and Users, IMF, 2013.

OECD Benchmark Definition of Foreign Direct Investment 4th edition BD4.

Regulation (EC) No 184/2005 of the European Parliament and of the Council of 12 January 2005 on Community statistics concerning balance of payments, international trade in services and foreign direct investment. OJ L 35, 8.2.2005, p. 23, as amended by Commission Regulation No 555/2012 (OJ L 166, 27.06.2012).

Commission Regulation (EC) No 601/2006 of 18 April 2006 implementing Regulation (EC) No 184/2005 of the European Parliament and of the Council as regards the format and the procedure for the transmission of data. OJ L 106, 19.4.2006.

Commission Regulation (EC) No 1055/2008 of 27 October 2008 implementing Regulation (EC) No 184/2005 of the European Parliament and of the Council, as regards quality criteria and quality reporting for balance of payments statistics. OJ L 283, 28.10.2008, as amended by Commission Regulation (EU) No 1227/2010 (OJ L 336, 21.12.2010, p. 15-16).

Commission Implementing Regulation (EU) No 228/2014 of 10 March 2014 amending Regulation (EC) No 601/2006 implementing Regulation (EC) No 184/2005 of the European Parliament and of the Council on statistics concerning balance of payments, international trade in services and foreign direct investment, as regards the format and the procedure for the transmission of data. OJ L 70, 11.3.2014, p. 16–17.

Regulation (EU) 2016/1013 of the European Parliament and of the Council of 8 June 2016 amending Regulation (EC) No 184/2005 on Community statistics concerning balance of payments, international trade in services and foreign direct investment. OJ L 171, 29.6.2016, p. 144–152.

Law on the National Bank of the Republic of North Macedonia (Official Gazette of the Republic of Macedonia No. 158/10, 123/12, 43/14, 153/15, 6/16 and 83/18 and Official Gazette of the Republic of North Macedonia No. 110/21).

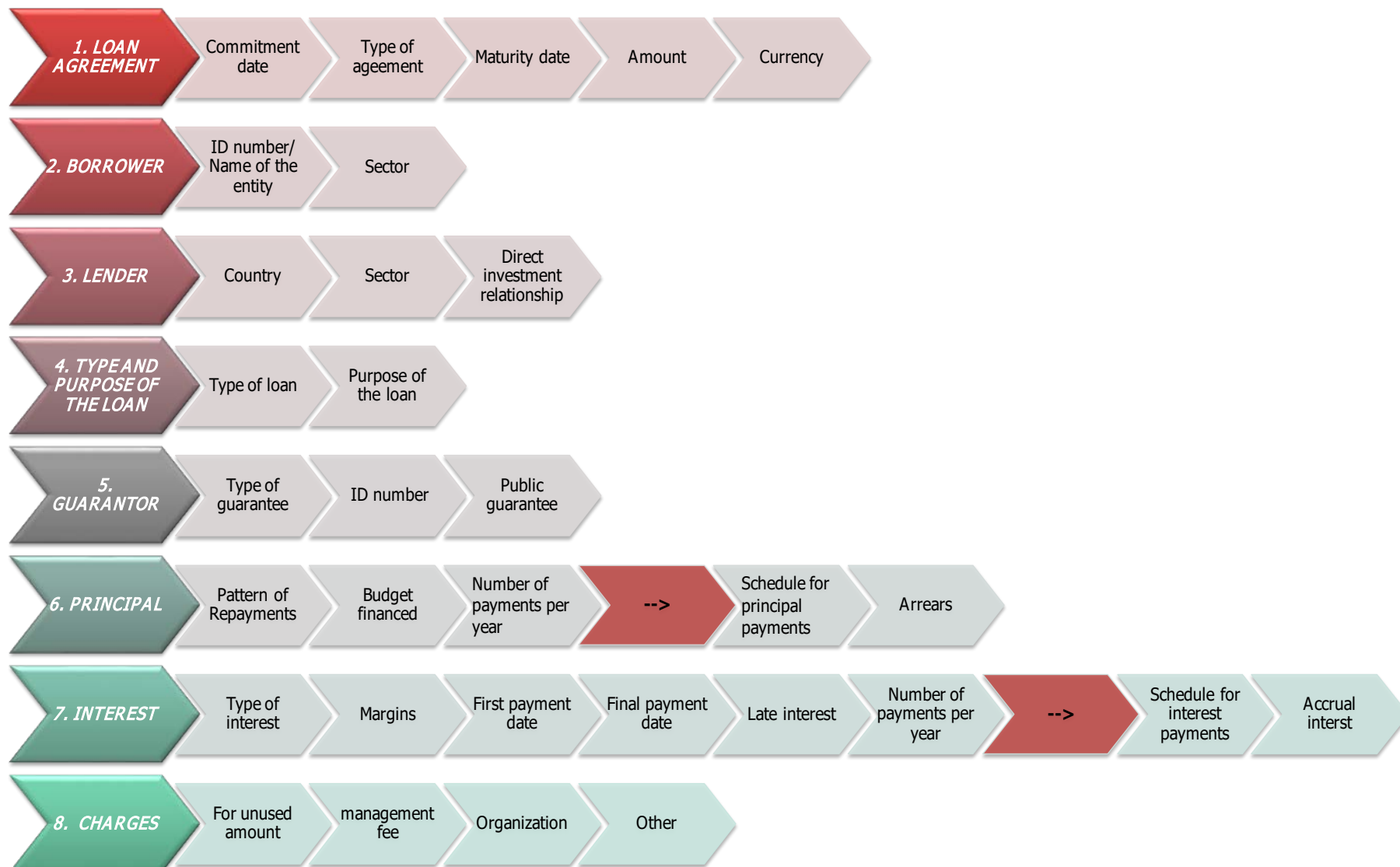
Law on Foreign Exchange Operations (Official Gazette of the Republic of Macedonia No. 34/01, 49/01, 103/01, 51/03, 81/08, 24/11, 135/11, 188/13, 97/15, 153/15 and 23/16 and Official Gazette of the Republic of North Macedonia No. 110/21).

Decision on the manner and the procedure for reporting on transactions with non-residents (Official Gazette of the Republic of Macedonia No. 87/14).

Decision on the manner and the terms of recording and reporting on the concluded credit operations (Official Gazette of the Republic of Macedonia No. 173/15 and 105/18).

Decision on the manner of performing international payment operations (Official Gazette of the Republic of Macedonia No. 42/11).

Annex 1: Detailed data included in the reporting form for loans





NATIONAL BANK OF THE REPUBLIC OF NORTH MACEDONIA

Granular statistical data – Are we charmed with a blessing or a curse?

Jasminka Dodeva Kovacheva

BoC-IFC Satellite Seminar,

Bank of Canada

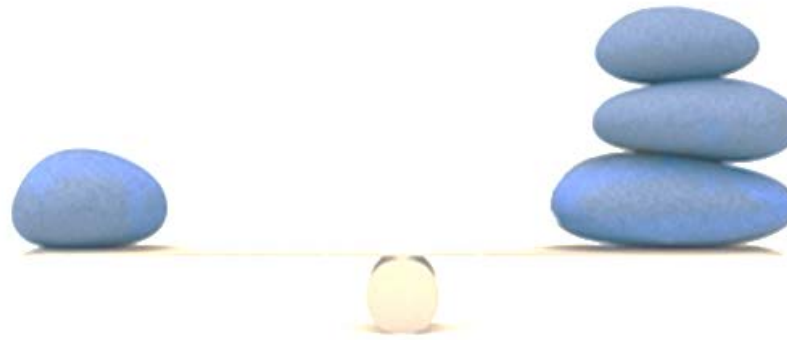
July 15, 2023

***Granular statistical data – Are we
charmed with a blessing or a curse?***

Introduction

- “New reality” - globalization and digitalization, new financial instruments, multinational companies, special purpose entities-bring new need for more data;
- How to make a choice which data sources to be used?...
- ...and what level of details to be used;
- Is there a prototype?;
- The obligation and responsibility that obtaining granular data carries.

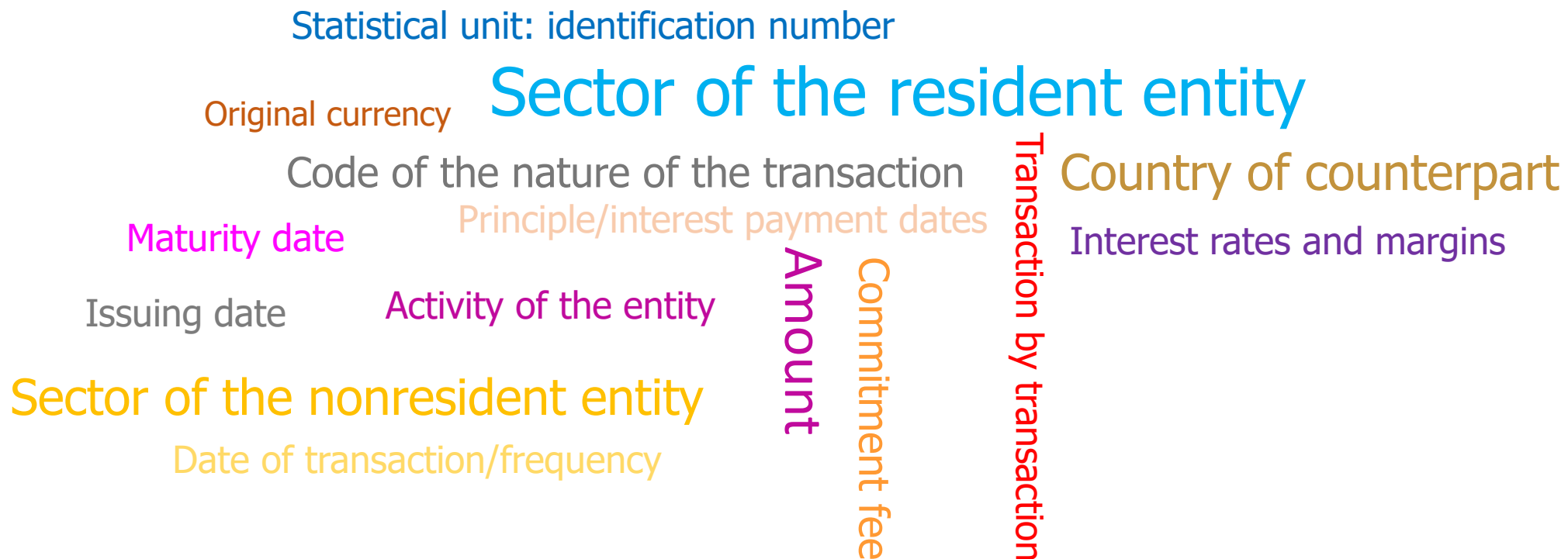
What is granular data and how detailed data should be?



Less is More?

What is granular data and how detailed data should be?

- Global initiatives for granular databases – “everybody is going granular”



What works for us?



What works for us?

- Accepting and implementing internationally accepted standards;
- Electronically reporting and integrated IT systems;
- Maximizing the level of details with minimizing the reporting burden;
- Collaboration between public and private sector, as well as public-public sector;
- Custom made data sources;
- History matters, re-think how to utilize it for the future.

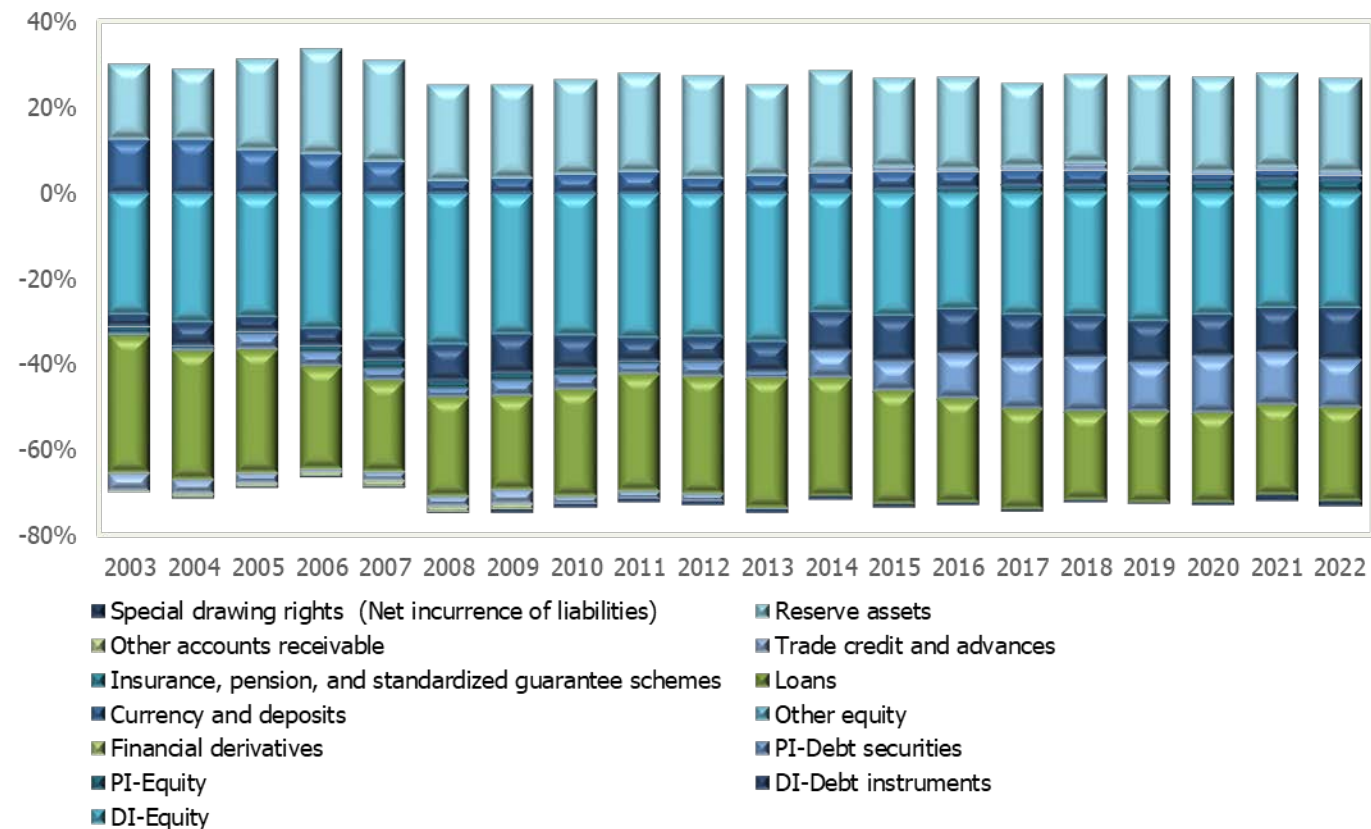
What works for us?

Granular data sources used:

- Loan - by - Loan;
- International transaction reporting system (ITRS);
- Security - by - Security;
- Integrated Reporting System for monetary data – Ongoing Project ISIDORA.

Loan - by - Loan

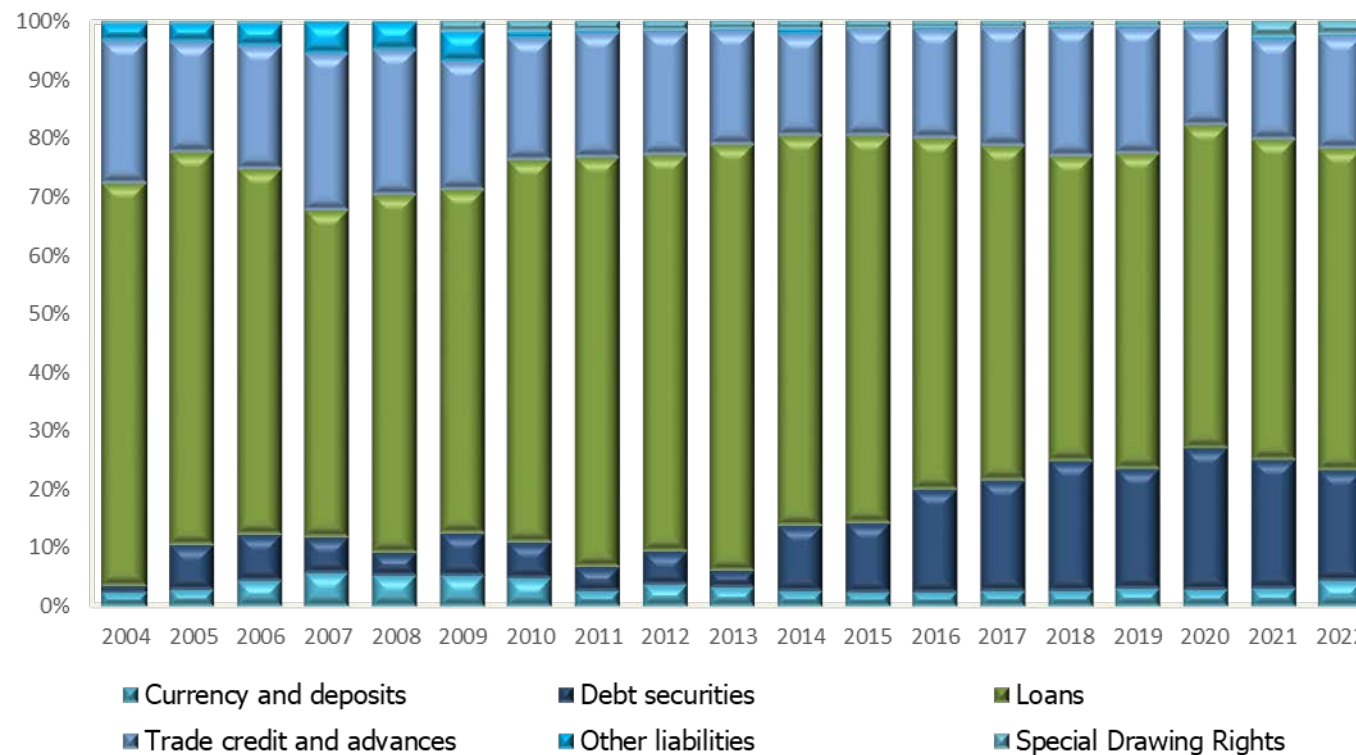
Graph 1: Structure of International Investment Position of the RNM, by instruments



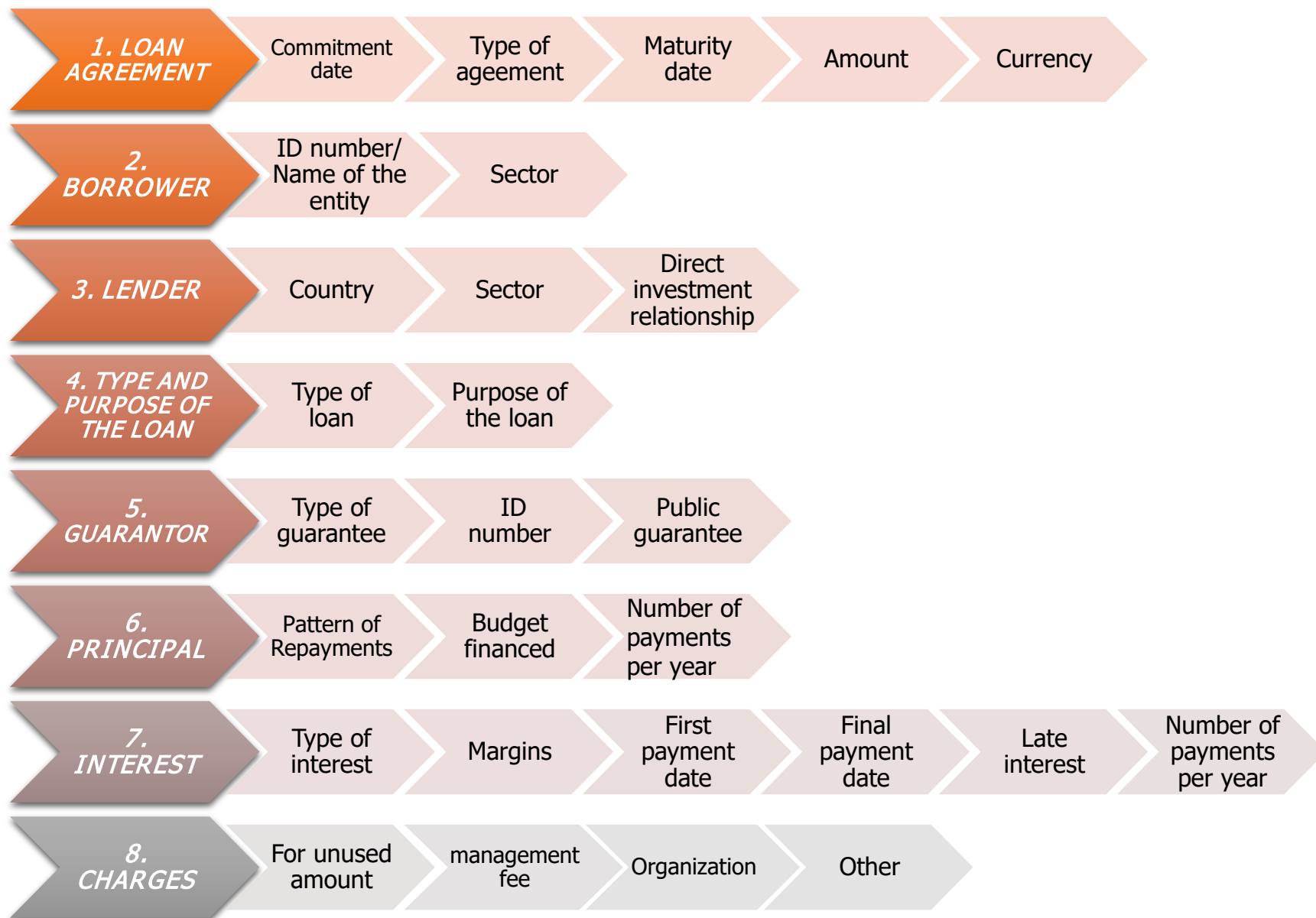
Source: NBRNM

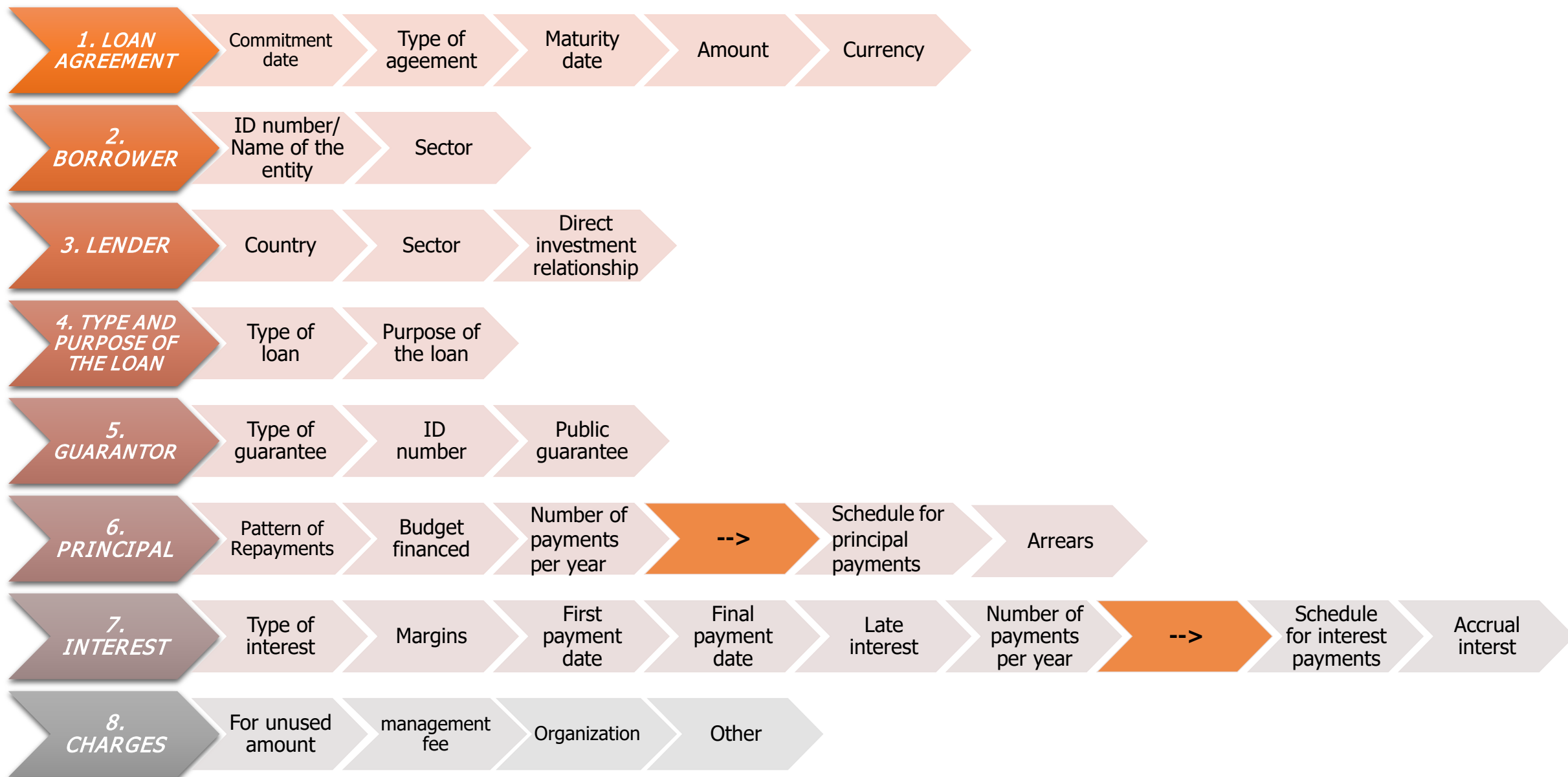
Loan - by - Loan

Graph 2: Structure of Gross External Debt, by instruments



Source: NBRNM



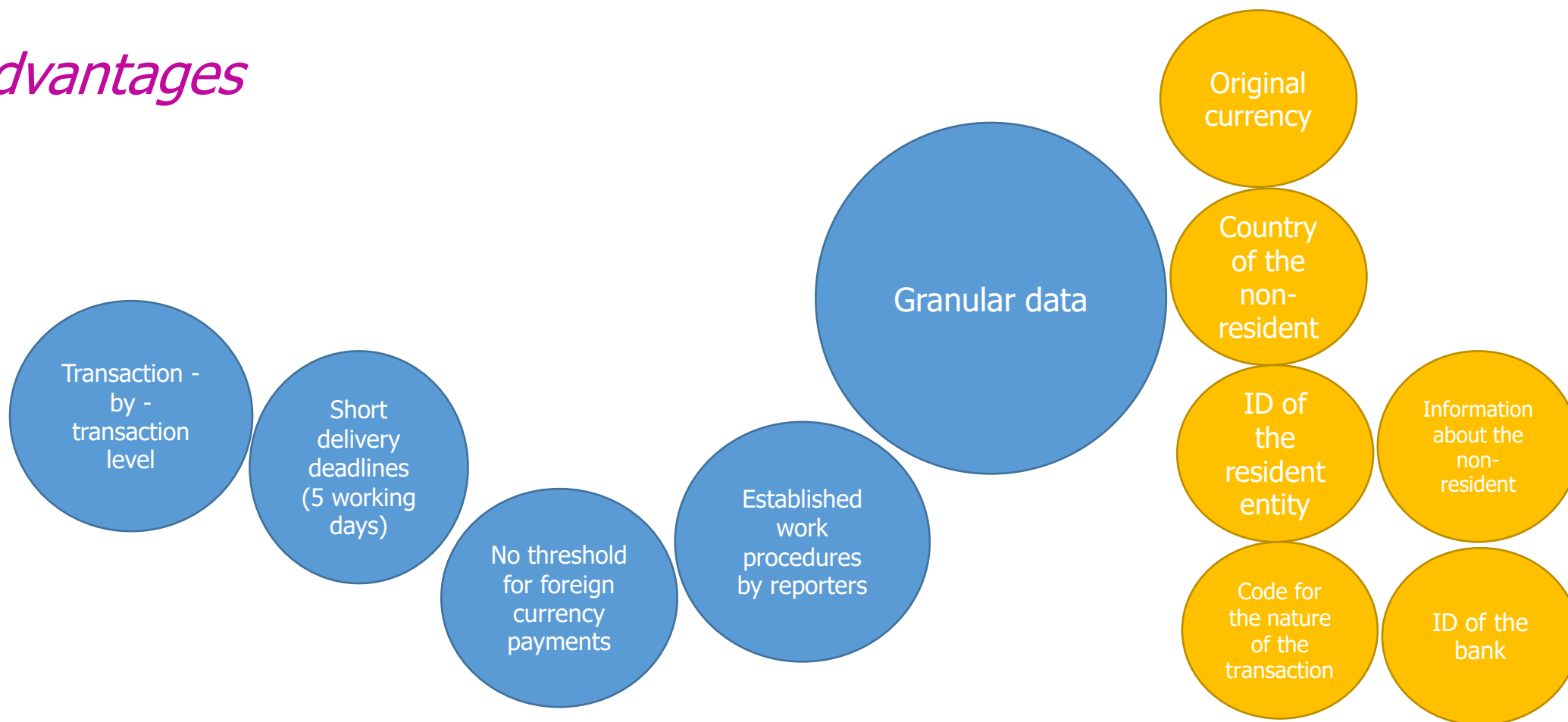


International Transaction Reporting System (ITRS)

- “Historically inherited system”;
- “Closed system” with full integration between stocks at the beginning of the period, transactions and stocks at the end of the period;
- Received mostly on a daily basis;
- Separate codes for each type of transaction (for BoP purposes);
- Electronically transmitted by the deposit-taking corporations.

International Transaction Reporting System (ITRS)

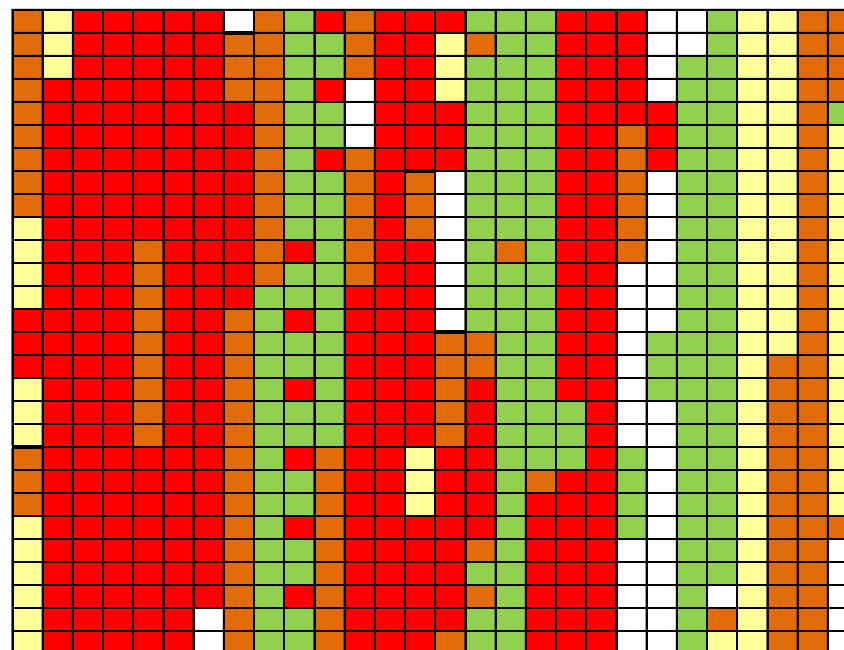
Advantages



International Transaction Reporting System (ITRS)

Picture 2: Data sources - ITRS usage in the Balance of Payments (“The Rubik’s cube”)

*Each square present an item from the BoP standard presentation



■ ITRS ■ ITRS used for cross-checking ■ Other data source ■ Combination ■ No data source established

Source: Author`s presentation based on NBRNM data.

International Transaction Reporting System (ITRS)

Disadvantages



Lack of compatibility with the internationally accepted methodological standards



Does not capture transactions that do not require payments or stock data



Delays or “slow” the realization of the international payments

Security - by - Security

- Newly introduced in 2014;
- Portfolio investments aboard – assets;
- ISIN – by – ISIN;
- Monthly reporting (t+20 days), made electronically through web application;
- Direct or indirect reporting made by supervisory agencies of insurance companies, investment funds and pension funds;
- Full reconciliation of data;
- Ensures compilation of Coordinated Portfolio Investment Survey ( SDDS Plus adherent, 2019).

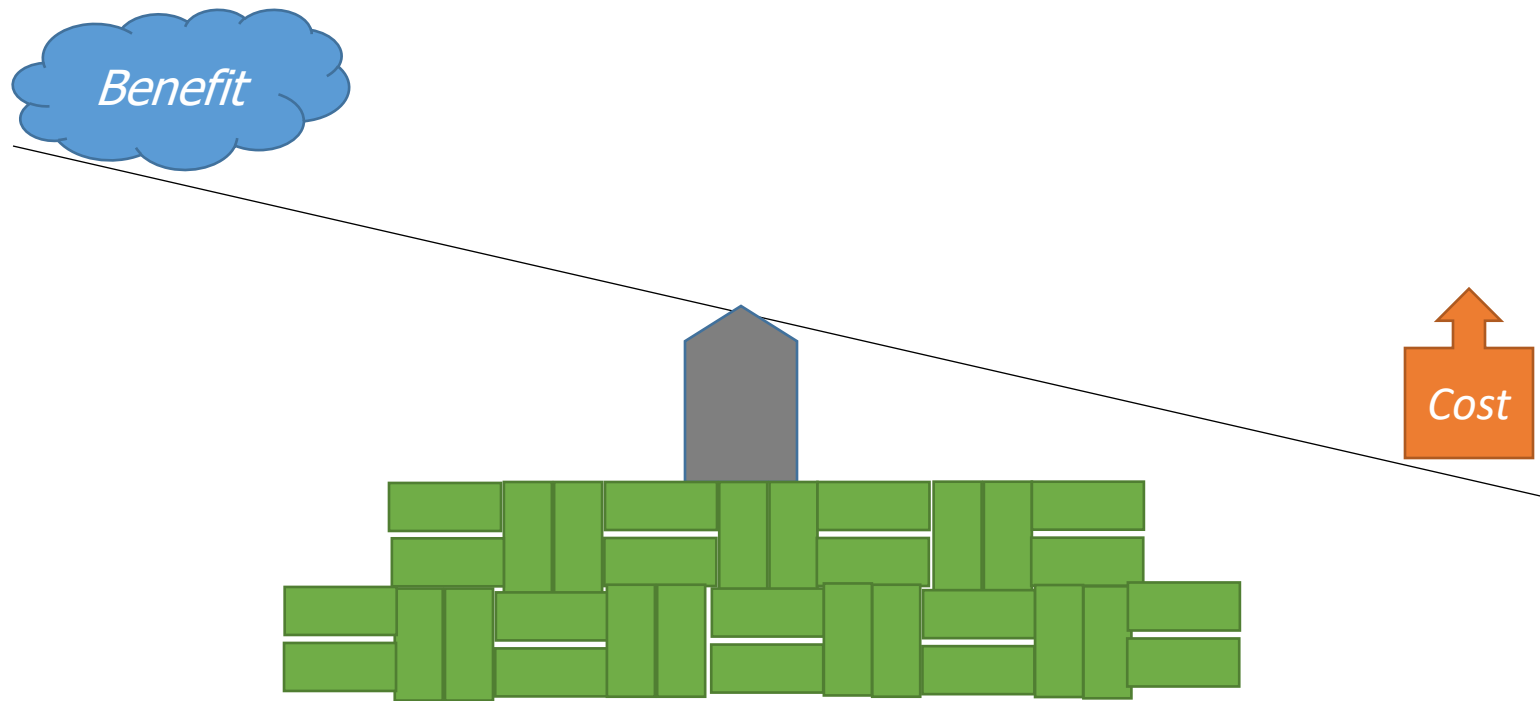
Security - by - Security

- Data are fully reconciled. Forms include data for:
 - Investor (ID, Name, Institutional Sector)
 - Security (ISIN, Currency of issue, Currency of trade, Market and Country of trading,..)
 - Issuer of the security (Name, Country, Institutional Sector,..)
 - Stock at the beginning of the period (Number/Market price per security at the beginning of the period)
 - Transactions (Number/Market price per security of purchased/sold securities)
 - Other changes (Price changes, Exchange rate changes, Other changes)
 - Stock at the end of the period (Number/Market price per security at the end of the period).

A glimpse into the future - ISIDORA

- Ongoing project;
- Collecting monetary data (balance sheet data from deposit-taking corporations, including the Central Bank, and saving houses) for statistical and supervisory purposes (e.g. Credit registry);
- Complete reorganization of the current data reporting system (from accounts to instruments described through predefined dimensions and attributes);
- Detail data structure (currency, country for non-residents, sector classification of non-residents, sub-sectoral classification of residents, flows data, separate the influence of price and exchange rate changes, as well as other changes in volume).

Cost and benefits of data granularity



Cost and benefits of data granularity

Benefits/**costs**

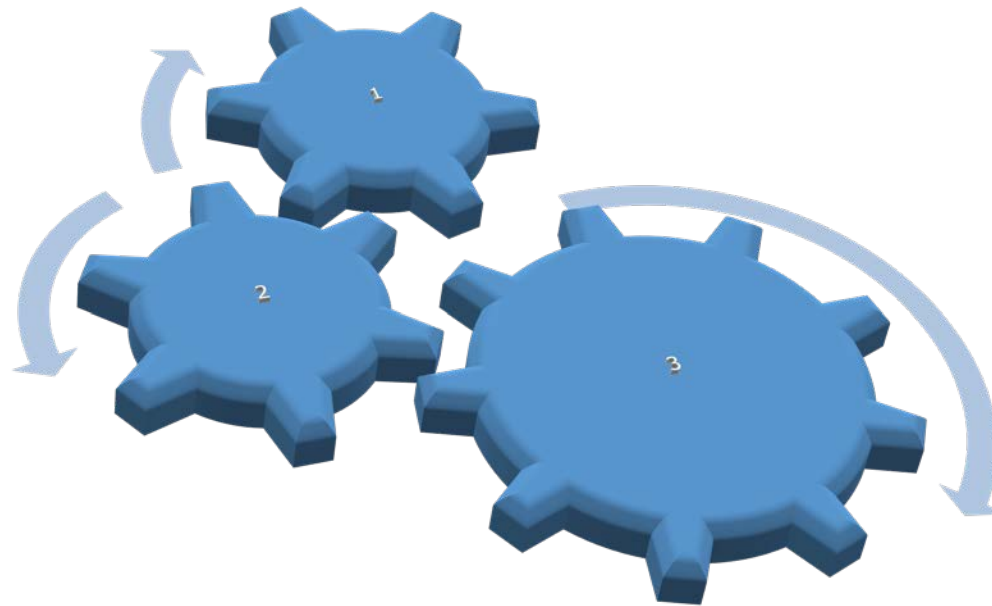
- Fulfilling the statistical methodological requirements;
- Easy to revise the time-series;
- Improved data quality by **crosschecking**;
- Better data analyses of the effects on the economy on a micro level;
- Decreasing the asymmetries between countries;
- **More detailed data** available for data users;
- Sometimes, **easier transfer** of the database from the reporting unit to the data collector;
- **High frequency reporting.**

Cost and benefits of data granularity

Stating the obvious

- Make sure not to collect the same information twice;
- Have a clear vision of the data that are useable or will become;
- Harmonization and standardization of data;
- Provide possibility for electronic transfer of data - have a developed IT resources;
- Ensure good communication and collaboration with the reporters.

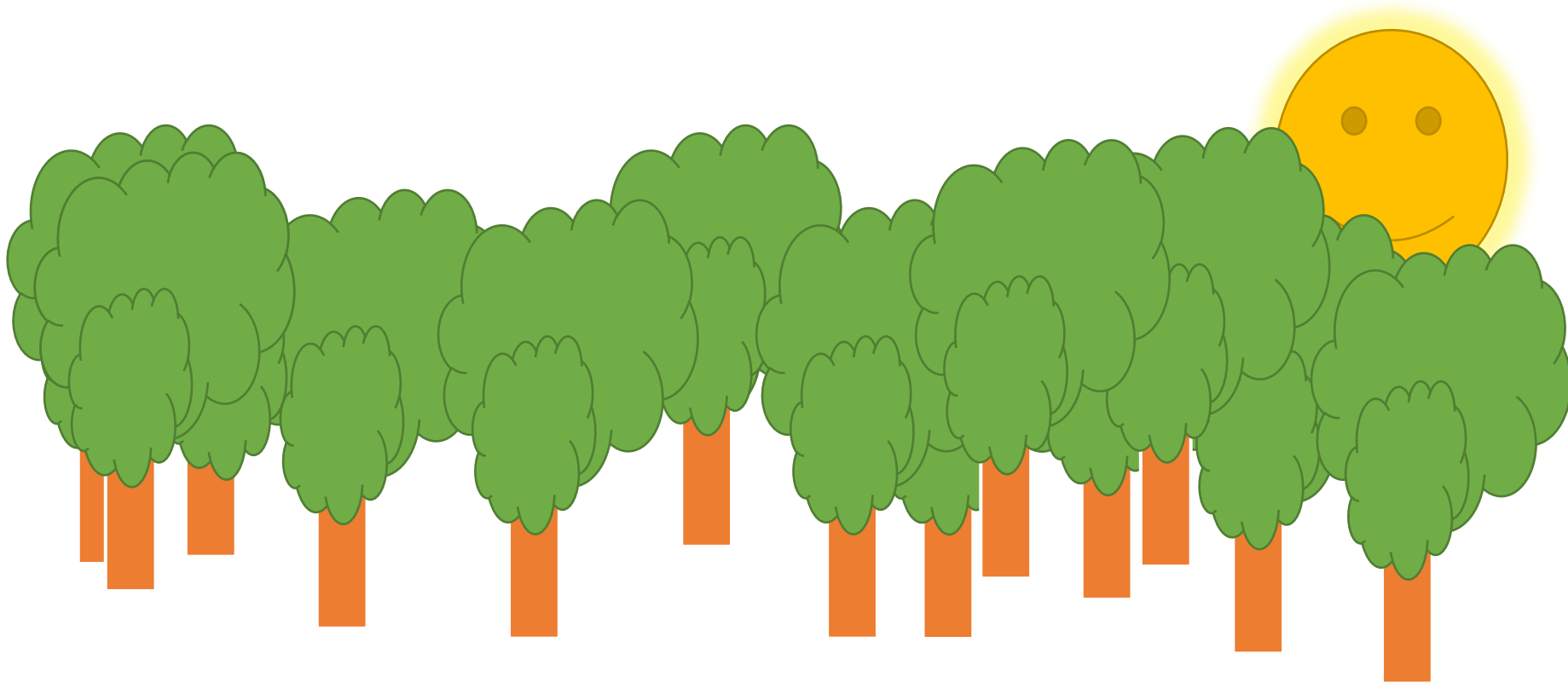
Data Management



Data Management

- The National Statistical System is defined in the Law of state statistics and in the five-year Statistical Programme;
- The regulatory framework consist of Law on State Statistics and the Law of the National Bank, as well as agreements for cooperation and data-sharing between institutions;
- Under the law, confidential data are data that allow direct or indirect identification of the statistical unit, and should be protected;
- Internal procedures and measures are in place to ensure data protection;
- Special procedures regulate the derogations in which confidential data can be disclosed.

Conclusions



Conclusions

- Granular, high frequent and quality data are essential;
- Introduce appropriate data sources;
- Adjust the data sources to the country;
- Apply international harmonization of the data structure, less to the data sources;
- Re-think and re-invent the “historically inherited” data sources;
- Consider the data reporters needs;
- Forward looking approach, modernization, by keeping the old charm.

Thank you