How can big data support financial stability work?

Bruno TISSOT
Head of Statistics and Research Support, BIS
Head of Secretariat, Irving Fisher Committee on Central Bank Statistics (IFC)

2018 CIRET Conference - Workshop on Big Data for Economic Statistics: Challenges and Opportunities
Rio de Janeiro, 11 September 2018

The views expressed are those of the author and do not necessarily reflect those of the BIS or the IFC.
Overview

- Introduction
  - Big Data for Central Banks
  - Three Key Developments
  - Nature of Financial Big Data
  - Financial Stability Work with Big Data

- Looking Forward

- Annexes: Selected References/ Projects
Introduction: Big Data / Big Noise?

Sources:
I – Big Data for Central Banks: increased interest...

• **Private sector** use

• New opportunities **also for Central Banks (CBs)?** – as well as macro-prudential authorities and financial supervisors?

• **Focus on information** supporting
  - Monetary policy eg economic forecasts & analysis
  - Financial stability eg macro/micro prudential, payment systems etc.
  - Other types of data – eg geospatial information – of lower interest
I – BD for Central Banks: ... with significant opportunities...

• Big data provide **new “business opportunities”** for CBs, such as:
  ➢ Qualitative statements to decipher communication
  ➢ Large number of big data pools generated by financial regulations
  ➢ In turn, big data can strengthen the monitoring capacity of public authorities

• **Feedback loop** inherent to policy-makers
  ➢ Big data sources can affect policy
  ➢ Policies implemented can generate new data-sets
I – BD for Central Banks: ... and the need to be proactive

• **IFC survey** of central banks (2015, 2017)
  ➢ Big data work still on an exploratory mode, yet increased interest

• Key objective for central banks is to **better understand**
  ➢ The new data-sets and related methodologies
  ➢ The value added in comparison with “traditional” statistics

• Focus on **pilot projects**
II – Nature of Financial Big Data: A broad approach...

- **Big Data**: by-product of commercial or social activities, providing a huge amount of very granular information

- **Coverage**:
  1. **Unstructured data-set** (often quite large):
     - The “internet of things”, produced organically
       - *Strong interest, but not really the core*
  2. **Large records**, relatively well-structured
     - By-products of 3 **types of activities**: financial, commercial & administrative
       - “Simpler”, *but can benefit from big data techniques*
II – Nature of Financial Big Data: ... 4 main types of datasets

- Commercial data
- Web-based indicators
- Administrative records
- Financial market data
III – Three key developments

• Big Data as a result of the combination of three key developments in the financial area
  ➢ The internet of things
  ➢ Digitalisation
  ➢ Expansion of micro financial data-sets in the aftermath of the Great Financial Crisis (GFC) of 2007-09
III – 3 Developments: Internet of things (new data)

• Information generated by **web and electronic devices**

• **Complement “standard” statistical processes**
  - More rapid information and improved timeliness
  - New ways of analysing / estimating economic patterns

• Usage relatively **limited**, often targeted at methodological improvements (eg quality), reducing reporting lags & revisions
  - Especially for economies with **less developed** statistics?
III – 3 Developments: Internet of things (new insights)

- **Estimates in advance** of actual publication dates (nowcasting)

- **Unsuspected data patterns**
  - BD algorithms to capture various effects without ex ante assumptions

- **Qualitative information**
  - Sentiment indicators
  - Important but difficult factors to model (non-linearities, network effects)
III – 3 Developments: Digitalisation (new data)

• **Expanded access** to digitalised information

  ➢ Rise in textual information moving to the web

  ➢ Not just produced by internet activities strictly speaking

  ➢ Reference documents can be digitalised, accessed and analysed like “web-based” indicators
III – 3 Developments: Digitalisation (new insights)

• Text can be more easily and automatically **exploited through ad hoc BD techniques**: eg text semantic analysis

• Measuring economic agents’ **sentiment & expectations**

• Assessing **policy**
  - Perceived stance of policy
  - Impact of policy communication / action
III – 3 Developments: Revolution in financial statistics (new data)

- **Impact of the GFC**
  - “To see the forest as well as the trees within it” (Borio, 2013)
  - Distribution matters: “fat tails”

- Unprecedented **efforts to use more micro** information
  - High demand for large, granular data-sets
  - Well structured, often more complex compared to “typical” internet data
  - Often derived from confidential registers
  - Example: **loan-by-loan / security-by-security** datasets
III – 3 Developments: New financial statistics (new insights)

• **Richer view** of the population of interest
  - Data collected regularly, over a long period of time
  - Need for anonymization / confidentiality protection
  - Information often already available but not exploited (administrative data)

• **CBs learning from private sector**
  - Increased experience in dealing with large data-sets (eg “stress tests”)
  - Supervisors of financial firms to develop their expertise in these areas too
IV – Financial stability work with big data: variety...

- In practice various & heterogeneous “big data”
  - Usually not designed for a direct statistical purpose
  - Indirectly, data exploited for addressing statistical needs

- Public authorities at the beginning of making sense of these data
  - Use of specific sources depends on policy questions
  - Eg payment systems: of interest to supervisors and tourism analysis
  - “Smart data”: treatment of the raw, “organic” data is key
IV – Financial stability work: ... complexity...

• **Micro-level BD universe** is complex and evolves over time
  ➢ Interaction between data available, specific needs and policy actions

• **Transforming data into information that is relevant for policy**

→ "*Connecting the dots is as important as collecting the dots, meaning the right data*" (Caruana, 2017)
IV – Financial stability work: ... and challenges...

• Specific **challenges** reflecting central banks’ nature
  - **Public status** of financial authorities and public trust
  - Central banks concerned about **ethical & reputational** consequences
  - Risk of **misusing** big data for policy actions?

• **Security concerns linked** to internet / big data, such as:
  - Risk that large private records of individual information could be accessed and potentially misused by unauthorized third-parties
  - Peculiar position of central banks if private information is reported to them but not protected adequately
  - Resilience of financial market infrastructures
IV – Financial stability work: ... esp. in handling big data...

- **Resources** (IT, staff) and proper arrangements for managing BD

- The statistical *production process* itself has to be adapted
  - Need to set up a clear and comprehensive information management process

- **Reputation risk when handling** big data
IV – Financial stability work: ... and using big data

- Does “big data” provide a more **accurate economic** picture?
  - Coverage bias unknown, can be significant (e.g., social media users)
    - Extremely **large big data samples may compare unfavourably** with (smaller) traditional probabilistic samples

- **Reputation risk** when using big data
  - Concerns about lack of transparency, poor quality of some sources
    - Social costs of misguided policy decisions

- Can big data **alter decision-making**?
  - Bias towards responding to news, encouraging shorter horizons?
  - Risk of fine-tuning policy communication based on expectations?
  - How to communicate the results of “black boxes”?
V – Looking forward: CBs to be alert...

- Information needs evolve over times:
  - The **financial system changes**... not least due to policy actions
  - Assessment of how **fragilities are building up** typically rely on **aggregated** statistics to spot “abnormal patterns”
  - In contrast, **resolution work** in the aftermath of a financial crisis will request much more timely and **granular** information

→"**Generally, rough aggregates suffice to indicate that imbalances are building. Once a crisis breaks out, however, more granular data are needed for taking decisions** (Carstens, 2018)."
V – Looking forward: ... focus on information strategy...

- **Decisions on data** have become of strategic importance for central banks

- **Big Data reinforces the need** to:
  - Balance costs and resources implications
  - Consider the various financial, legal and reputational risks

- Proper **information governance frameworks** needed to adequately manage BD-sets collected / used by central banks
V – Looking forward: ... and evolving business models

- What is still unknown is whether and how far Big Data will trigger a change in CBs’ “business models”
  - CBs are relatively new in exploiting big data, in contrast to the greater experience gained by NSOs
  - CBs have traditionally been data users rather than data producers, though the situation has clearly changed since the GFC
  - CBs are thus in a key position to ensure that BD can be transformed in useful information supporting policy
Annex (1): Selected references


Hammer, C., Kostroch, D., Quiros, G., & Staff of the IMF Statistics Department (STA) Internal Group (2017). *Big data: potential, challenges, and statistical implications*, IMF Staff Discussion Note, Staff Discussion Notes (SDN)/17/06, September.


## Annex (2): Selected BD projects by central banks

<table>
<thead>
<tr>
<th>Big data areas</th>
<th>Types of data-sets</th>
<th>Examples of projects</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Administrative records</strong></td>
<td>Foreign trade operations / investment transactions</td>
<td>Balance of payments statistics eg tourism, exports</td>
</tr>
<tr>
<td></td>
<td>Taxation / payroll / unemployment insurance</td>
<td>Employment, wages, business formation (SMEs)</td>
</tr>
<tr>
<td></td>
<td>Central balance sheet offices</td>
<td>Performance vulnerabilities assessment</td>
</tr>
<tr>
<td></td>
<td>Loans registers</td>
<td>Measurement of credit risk, FX exposures</td>
</tr>
<tr>
<td></td>
<td>Financial market supervisors</td>
<td>Network analysis, exposures</td>
</tr>
<tr>
<td></td>
<td>Public financial statements</td>
<td>Corporate balance sheet, group-level supervision</td>
</tr>
<tr>
<td></td>
<td>Financial market activity indicators</td>
<td>Payments systems, Trade repositories</td>
</tr>
<tr>
<td><strong>Web-based indicators</strong></td>
<td>Internet clicks</td>
<td>Google searches</td>
</tr>
<tr>
<td></td>
<td>social networks</td>
<td>confidence &amp; economic sentiment</td>
</tr>
<tr>
<td></td>
<td>Digitalised content / text</td>
<td>policy communication, analysis of expectations</td>
</tr>
<tr>
<td></td>
<td>Websites’ scraping</td>
<td>Various uses</td>
</tr>
<tr>
<td></td>
<td>Job portals</td>
<td>Employment / activity</td>
</tr>
<tr>
<td></td>
<td>Prices posted directly on websites</td>
<td>Measure specific components of the CPI, PPIs, Inflation nowcasting / forecasting</td>
</tr>
<tr>
<td></td>
<td>Real estate agencies</td>
<td>House price indices</td>
</tr>
<tr>
<td><strong>Commercial data-sets</strong></td>
<td>Credit card operations</td>
<td>Payments patterns, Tourism</td>
</tr>
<tr>
<td></td>
<td>Mobile operators</td>
<td>Mobile positioning data (eg travelers’), Financial inclusion</td>
</tr>
<tr>
<td></td>
<td>Geo spatial information</td>
<td>National statistical system Tasks</td>
</tr>
<tr>
<td></td>
<td>Credit institutions</td>
<td>Balance sheet exposures, Investor behaviour/expectations</td>
</tr>
<tr>
<td><strong>Financial data-sets</strong></td>
<td>Settlement operations</td>
<td>Operational risks, Market functioning</td>
</tr>
<tr>
<td></td>
<td>Securities issuance</td>
<td>Security-by-security databases</td>
</tr>
<tr>
<td></td>
<td>Market liquidity</td>
<td>Bid/ask spreads</td>
</tr>
<tr>
<td></td>
<td>Custodians records</td>
<td>Securities holding statistics</td>
</tr>
<tr>
<td></td>
<td>Tick-by-tick data</td>
<td>Real-time analysis of financial patterns</td>
</tr>
</tbody>
</table>
Thank you!!

Questions?

bruno.tissot@bis.org
IFC.secretariat@bis.org