



Irving Fisher Committee on
Central Bank Statistics

BANK FOR INTERNATIONAL SETTLEMENTS

IFC – Bank Indonesia International Workshop and Seminar on *“Big Data for Central Bank Policies / Building Pathways for Policy Making with Big Data”*

Bali, Indonesia, 23-26 July 2018

A personal view on big data and policymaking¹

Naruki Mori,
Bank of Japan

¹ This presentation was prepared for the meeting. The views expressed are those of the author and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the meeting.

”Building Pathways for Policy Making with Big Data”
IFC–Bank Indonesia Satellite Seminar on Bali, 26 July 2018

日本銀行
BANK OF JAPAN



A Personal View on Big Data and Policymaking

Naruki Mori,
Associate Director-General,
Research and Statistics Department
Bank of Japan
naruki.mori@boj.or.jp

The views in this presentation are my own and do not necessarily reflect those of the Bank of Japan.

Today' s topics

1. Analytical use of big data
2. Possible use of big data for statistics
 - A case study: big data for the price index
3. Useful for policymaking?

1. Analytical use of big data at the Bank of Japan

➤ Financial markets analysis

- ✓ Analysis of the Japanese government bond (JGB) markets

- Intraday market liquidity of JGB futures

<https://www.imes.boj.or.jp/research/papers/english/me34-3.pdf>

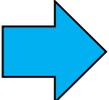
- Indicators measuring liquidity in the JGB markets (cash and futures)

http://www.boj.or.jp/en/research/brp/ron_2018/data/ron180329a.pdf

➤ Payment and settlement system analysis

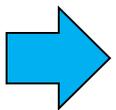
- Effects of the improvement in payment and settlement system

<http://www.boj.or.jp/research/brp/psr/psrb160629.pdf> (Japanese only)

 High-frequency data such as tick data are increasingly used in the analysis of financial markets.

2-1. Possible use of big data for statistics

- We are exploring possibilities of using big data in compiling the existing statistics: for example,
 - ✓ Securities statistics in the flow of funds accounts
 - Use of micro (security-by-security) data sets with description on individual issues (e.g. interest rate, currency, maturity) available on the website of Japan Securities Depository Center (JASDEC)
 - ✓ Producer Price Index
 - Use of web scraped data for splitting price differences between old and new products into those due to quality changes and those due to pure price changes



New data sources are complementary to traditional data sources such as those compiled by national statistical agencies.

2-2. A case study: big data for the price index – Traditional and Non-traditional approaches –

Based on our staff's draft paper available at:

<https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.22/2018/Japan.pdf>

- We combine features of both traditional approach and non-traditional approach by applying machine learning methods in order to pair old and new products accurately. We use big data obtained from Japan's leading price comparison website *Kakaku.com* and **machine learning methods**.
- We can call our new method “Webscraped Prices Comparison Method (WSM).”

Traditional approach: Price index which is created by carrying out changes of sample prices reflecting the product life cycles and quality adjustments between old and new products.

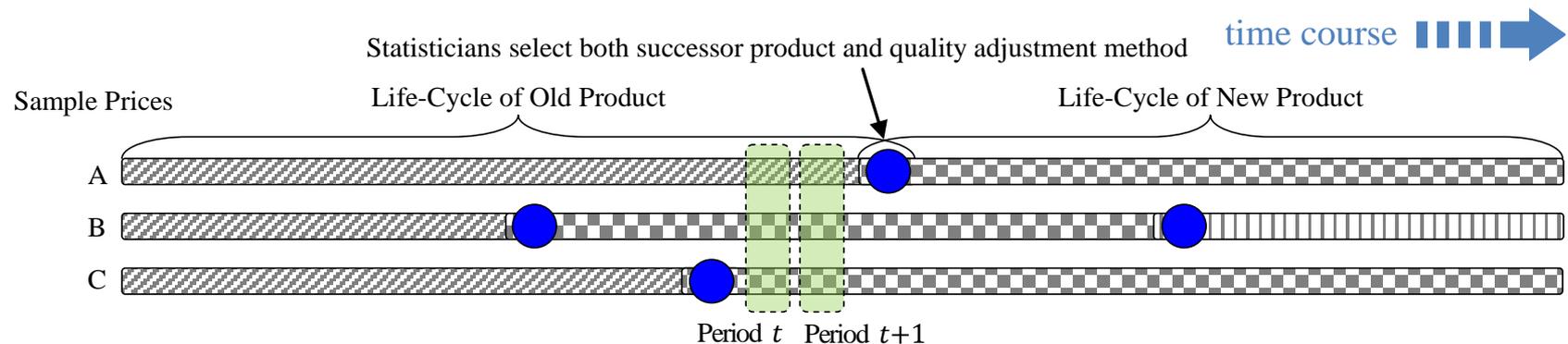
and

Non-traditional approach: Price index which is compiled by making use of big data and computing capabilities.

2-2. How to make old and new product pairs

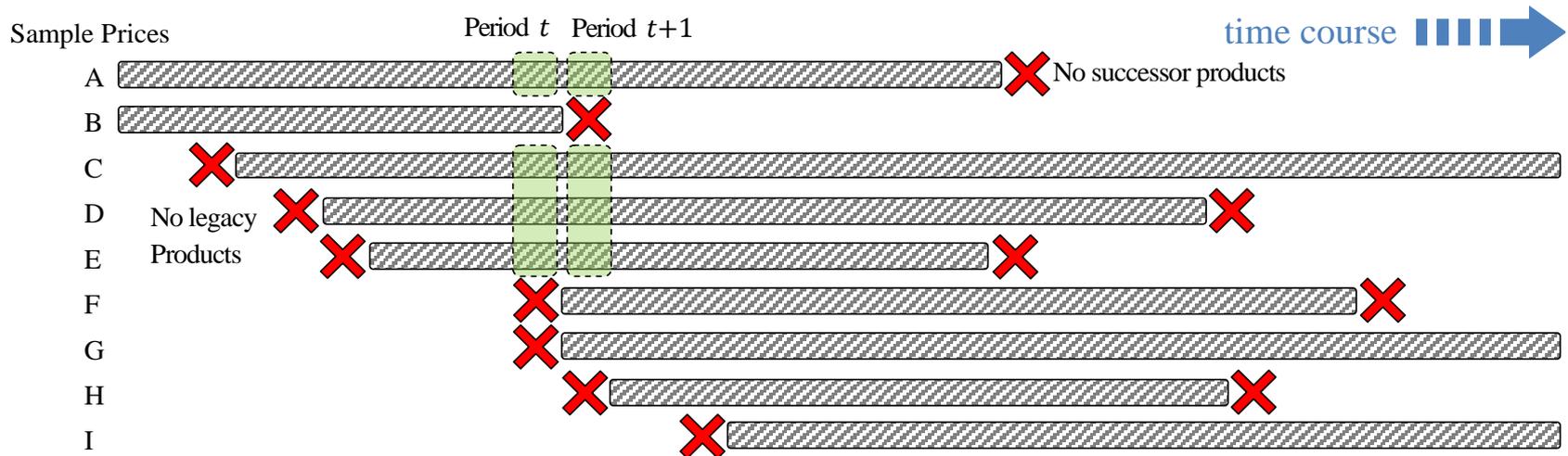
– Traditional approach adopted by statistics agencies–

- Price statisticians select representative products to be surveyed, considering product specifications and data availability. At time of changing sample prices, they apply optimal quality adjustment to remove price change arising from changes of quality.
- Resources constraints at statistics agencies and reporting burden normally lead to a small number of sample prices.



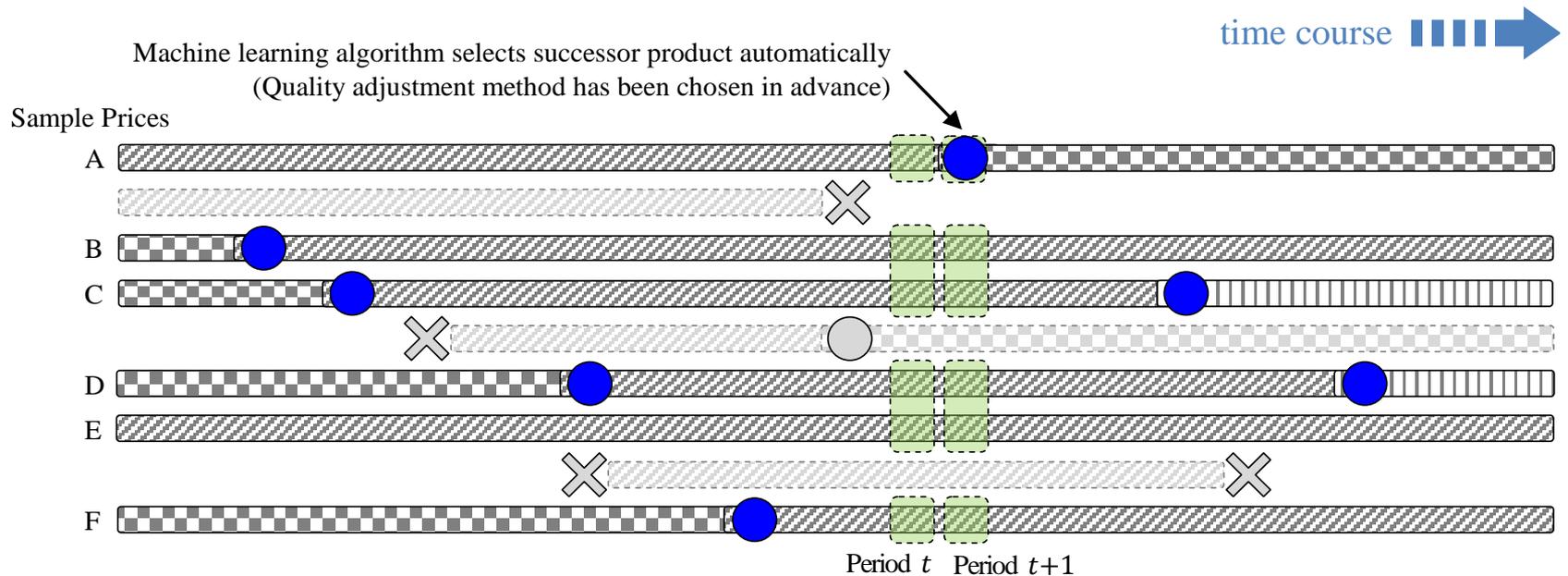
2-2. How to make old and new product pairs – Non-traditional approach using big data –

- Compiling price index by using big data and **Matched-Model Method** (MMM) which calculates the percentage change of price for products which exist in both *survey period* and *following period*.
- If price pushbacks are constantly conducted when launching new products, the index cannot properly reflect the impact of such price pushbacks, and may cause a downward bias.



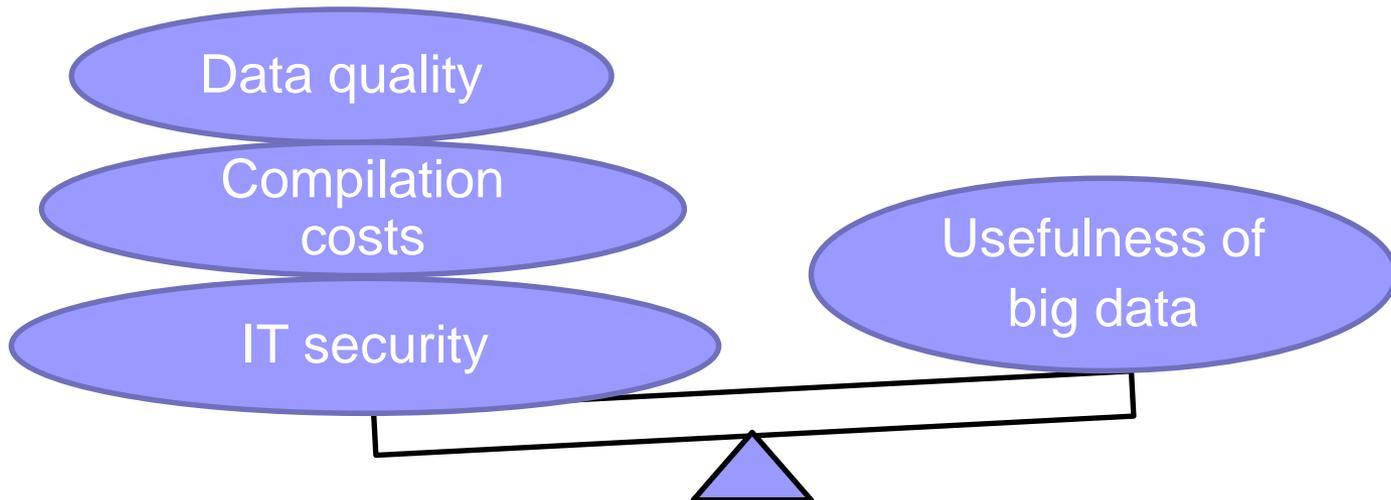
2-2. How to make old and new product pairs – A new method which adopts machine learning –

- We developed a **supervised machine learning** algorithm which pairs legacy and successor products with high precision to conduct appropriate quality adjustment for given big data.
- This index properly reflects the impact of price pushbacks.



2-3. Challenges for BOJ

- Constraints on human and financial resources
- ✓ In order to utilize big data for economic research and statistics, we need to have staff with necessary expertise (e.g. data scientists who can make good use of machine learning methods).
- ✓ Handling big data requires not only significant resources but also data quality (reliability in data sources), methodological soundness, IT security, and proper arrangements for the protection of privacy and confidentiality. We need to strike a balance between costs and benefits.

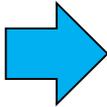


3. Big data for policymaking?

➤ Inputs for Policy Board members

- ✓ High-frequency data of prices (e.g. weekly, daily)
 - Daily price index compiled by scanner data (T-point price index)
 - Weekly purchasing price index of households
- ✓ Text mining to gauge business sentiment
 - Text mining of Economy Watchers Survey compiled by Cabinet Office
 - Google Trends

➤ Readily available and high-frequency economic data may help policymakers to assess current economic developments timelier (e.g. nowcasting).



However, does the policymaking framework quickly adopt them? Otherwise, the use of big data is limited to providing background information for Policy Board members in making their assessments and decisions.

3. General remark: useful for policymaking?

- How can we fit big data into the policy reaction function? For example, how about the Taylor rule?
 - At major central banks, price stability target (π^*) is headline CPI inflation.
 - Taylor rule:
$$i_t = \pi_t + \alpha(\pi_t - \pi^*) + \beta y_t$$
 - IS curve:
$$y_{t+1} = \theta y_t - \sigma(i_t - E_t \pi_{t+1})$$
 - Phillips curve:
$$\pi_{t+1} = \gamma \pi_t + \delta y_t$$
- where i = policy interest rate, π = CPI inflation, y = GDP gap
- It may suggest that CB's policy reaction function continue to be based on traditional data sources.
 - Appropriate policymaking needs to be based on accurate economic data. Thus, policymakers and statistics agencies have been making efforts to improve accuracy of economic indicators (e.g. preliminary estimates (QE) of GDP, price index). What role will big data play in these efforts?

Thank you!

