ธนาคารแห่งประเทศไทย
BANK OF THAILAND

Stat-Horizon

Uplifting the Usability of Financial Statement Data to Leverage the Quality of the Sectoral Balance Sheet (SBS) Estimates

Jittima Dummee

Natchaya Mahaviriyakul

Pornsawan Rugpenthum

Prossaya Jitapunkul

Bank of Thailand

DECEMBER 2024

Abstract

One of the main challenges in compiling the sectoral balance sheet (SBS) is the sourcing of data for financial asset and liability positions of nonfinancial corporations (NFC), households and nonfinancial institutions serving households (HH & NPISH), particularly claims and liabilities among themselves where mirror data (monetary and financial statistics, external sector statistics, and government finance statistics) are not available as in the case of positions vis-à-vis financial corporations, general government, and the rest of the world. Although administrative records such as financial statements reported to relevant government units, theoretically, could serve as the data source for compiling SBS of these two sectors, the quality, frequency, and timeliness concerns could still hinder their usability in practice.

This paper focuses on developing a framework for improving the quality of the SBS estimates through the use of statistical techniques to

(i) impute missing values for selected elements in the financial statement data which registered companies nationwide reported to the Ministry of Commerce (MOC) annually;

(ii) derive quarterly estimates of the SBS from the annual financial statement data; and

(iii) produce nowcasting estimates for SBS of recent quarters for which annual financial statements are not yet reported to the authorities.

In addressing the three aforementioned issues, a number of statistical techniques were explored to extract 'peer pattern' at different levels of disaggregation; and tested how well each method (or a combination thereof) could be applied to estimate the concerned elements/units of financial statement data. The results and associated statistical errors for each method are evaluated and compared to identify the option that produce the best estimates of selected financial statement elements (e.g., equity, trade credits, etc.).

Given the significance of NFC sector in Thailand's overall SBS, the result of this study should serve as a key steppingstone towards improving the quality of Thailand's SBS compilation and pave way for its dissemination in the near future.

Content

Uplifting the Usability of Financial Statement Data to Leverage the Quality of the Sectoral Balance Sheet (SBS) Estimates

1. Introduction

A number of studies on the root causes and warning signs of the world's recent economic crises (namely, those occurred in Mexico, East Asia, and Russia in the 1990s, the global financial crisis originating in the US financial sector during 2007-2009, and public debt crisis in European countries) reveals that one contributing factor is the lack of data/statistics to support insightful analysis, thus hindering timely detection of associated risks/vulnerabilities.

Quality, timely, and comprehensive data is one prerequisite for well-informed decision-making and effective macroeconomic policy formulation. High-quality data with sufficient level of granularity, frequency, timeliness, and dynamic enough to keep pace with financial innovations are key to early detection of pre-crisis symptoms, thereby allowing reasonably ample time to design preventive and remedial measures to mitigate the impacts.

As the financial system evolves, complexity has been growing in the form of (1) more diversification of players (i.e. investors/net savers, intermediaries, investment recipients /borrowers); (2) more variety of financial instruments and hybrids thereof; and (3) increasing number of investment channels/platforms. These have concertedly deepened financial linkages among different sectors/sub-sectors of players, as well as the transmittal mechanism through which risks could be spilled over to these players.

The statistics which could shed light on these inter-sectoral financial linkages is **"the sectoral balance sheet" (SBS)**[1]. The SBS comprehensively illustrates financial claims that one sector/sub-sector has vis-à-vis one another in the economy, as well as foreign countries (referred to as "rest-of-the world sector" in the SBS jargon). This information enables policy makers to assess and analyze vulnerabilities associated with each economic sector, and the plausibility of the "risk spillovers" transmitted through this inter-sectoral connectedness.

[1] SBS is constructed based on the concept of the balance sheet approach (BSA), developed by the International Monetary Fund (IMF) as a tool for analyzing inter-sectoral financial linkages, as well as associated risks and potential spillovers. (Puntharik Supaarmorakul, et al (2014). "แนวทางการประเมินเสถียรภาพการเงินของประเทศ ผ่านกรอบการวิเคราะห์งบดุลรวม (Guidelines for Financial Stability Assessment through the Balance Sheet Approach Framework)". Stat Horizon (Thai version).

To derive a quality SBS, however, is rather challenging. In Thailand's context, "decentralized statistical system" is adopted, with several government agencies and institutions involved in the data collection and compilation of different economic and financial statistics. Legal ground empowering data collection/reporting for each agency/institution also varies in degree and scope. This results in difficulties in gathering necessary information and data to support SBS compilation, particularly the data for NFC, HH, and NPISH sectors.

As an attempt to overcome these challenges, the author team explored different sources of readily available data (e.g., secondary data from other macroeconomic accounts, administrative records such as "financial statement data" which corporates are required by law to submit to the MOC on an annual basis, etc.) and assessed their usability to support the compilation of the SBS. Details pertaining to the principles and processes for data validation and preparations are elaborated in later sections.

2. Principles for Preparing Thailand's SBS

A balance sheet, also known as "a statement of financial positions", is a statement that illustrates the assets and liabilities of each economic sector at a specific date, revealing the financial linkages and vulnerabilities of each sector which could be illustrated through the following equation:

Net Financial Position[2] = Asset - Liability

Constructed on the ground of this balance sheet concept, the SBS systematically reflects how each sector of the economy relates to one another through financial claims (i.e., asset and liability positions that each sector has vis-à-vis the rest). Financial assets and liabilities are further broken down by financial instrument, namely:

1. Monetary gold and Special Drawing Rights (SDR)
2. Currency and deposits
3. Debt securities
4. Loans
5. Equity and investment fund shares

[2] Positive net financial position means that the respective economic sector is a "net creditor", while negative net financial position means that the respective economic sector is a "net debtor".

6. Insurance, pension and standardized guarantees
7. Financial derivatives and employee stock options
8. Other accounts receivable/payable

Sector-wise, entities in the economy are categorized according to international sectorization standard, as follow:

1. Central bank (CB)
2. Other depository corporations (ODC)
3. Other financial corporations (OFC)
4. Non-financial corporations (NFC), which can be further broken down into:
    - o Public non-financial corporations (PNFC)
    - o Other (i.e., private-owned) non-financial corporations (ONFC)
5. General government (GG), which can be divided into:
    - o Central government (CG)
    - o Local government (LG)
6. Households and nonprofit institutions serving households (HH & NPISH)
7. Rest of the world (ROW)

3. Data sources and limitations

The data sources used by the Bank of Thailand (BOT) for SBS compilation encompass the followings:

1. Monetary and financial statistics (MFS) for positions of CB, ODC, and OFC vis-à-vis the remaining sectors
2. International investment position (IIP) for positions of ROW vis-à-vis the remaining sectors
3. Securities database (i.e., Financial Market Instrument (FMI) system)
4. Financial statements of NFC, which ONFC are required by law to submit to the MOC on an annual basis; while PNFC report to the State Enterprise Policy Office (SEPO).

Data with high quality and comprehensiveness are typically those of CB and ODC sectors (as they are closely regulated and subject to regular auditing), followed by OFC and ROW. Meanwhile, data gaps in terms of the coverage of NFC, HH & NPISH sectors' financial positions remain.

Table 1: Coverage assessment of Thailand's SBS

	CB	ODCs	OFCs	PNFCs	ONFCs	HHs & NPISHs	CG	LG	ROW
Monetary gold and SDRs									
Currency and deposits									
Debt securities									
Loans									
Equity and investment fund shares									
Insurance, pension and standardized guarantees									
Financial derivatives and employee stock options									
Other accounts receivable/payable									

Complete data (reported directly from direct primary sources)
Complete data (some of the data is taken directly from primary sources, while missing information is filled in using reports from relevant economic sectors.)
Incomplete data

In an attempt to narrow such data gaps, this study focused primarily on the fourth data source, which is the most challenging in terms of data processing, yet assessed to have the highest impact in enhancing the coverage of NFC sector (with HH & NPISH positions vis-à-vis NFC as by-products as well).

NFC can be divided into two categories in accordance with the input source:

1. For listed companies: data from securities database (i.e., FMI system, maintained and processed by the BOT)
2. For non-listed companies: financial statements reported annually to the MOC (so called — the Corporate Profile and Financial Statement (CPFS) database), available since 2001

The latter part was particularly of concern, with several limitations encountered as follow:

1. Financial statements of companies in Thailand are reported annually. Therefore, quarterly positions have to be imputed for compiling quarterly SBS.
2. Some companies failed to submit their financial statements in time, hence causing 'missing values' at the time CPFS data were processed for inclusion in SBS compilation.
3. Some companies did not report all items in the reporting template for certain years, resulting in omissions scattering across different financial instruments.
4. CPFS data are made available to SBS compiler with approximately seven-month lag, thus requiring techniques to produce preliminary estimates for the current year.

Subsequent sections illustrate different methods explored with an aim to overcome these limitations in deriving positions of non-listed companies.

4. Data preparation: Imputing missing values for selected financial instruments for private non-listed companies (non-listed ONFC)

Step 1: Filtering CPFS data for non-listed ONFC

- The study was conducted on the most recent 3-year CPFS data.
- Companies in the financial sector (International Standard Industrial Classification: ISIC Rev.4 Section K) were excluded.
- PNFCs were excluded.
- Companies listed in the Stock Exchange of Thailand (SET) and Market for Alternative Investment (MAI) were excluded.
- Companies which were suspended or no longer operated were excluded.
- Only those companies submitting financial statements for at least two consecutive years and have total income in the most recent year of more than 100,000 baht were included (i.e., "cutting off tail" with insignificant value of assets/liabilities).

Step 2: Conducting statistical tests to identify optimal method in imputing missing values

Thorough review of CPFS database revealed that most companies consistently reported data on equity and inter-company lending (if any). On the contrary, approximately 10 percent of companies did not report trade credits receivable; and five percent not reporting trade credits payable for at least one of the three years investigated. Therefore, this study emphasized on exploring different options in imputing missing values for these instruments.

Step 2.1: Creating "mock-up test data":

1. Selecting only those companies submitting financial statements for three consecutive years. Actual reported data were used as "benchmark" for this statistical test.
2. Removing some data points for certain years, so as to create "mock-up missing value" test data:

- Removing data points of some companies for one of the three years
- Removing data points of some companies for two consecutive years
- Removing data points of some companies for two non-consecutive years
- Removing data points of some companies for the three years

Step 2.2: Testing different methods in imputing missing values:

Four different methods tested include:

Method 1: Overall Growth Method

This method involves calculating year-on-year growth (%YoY) of respective financial statement items of all reporting companies combined (i.e., %YoY at aggregate level) and applying this percentage to estimate missing data for non-reporting companies in each of the three years tested. Table 2 illustrates how missing data were imputed under this method (imputed data are in yellow cells).

Table 2: Imputing missing data under Method 1

(assuming 2020 - 2022 %YoY were 15.5%, 12.4%, and 17.9%, respectively)

Company	2019	2020	2021	2022
Company A	1,306	1,306 * (15.5% + 1)	1,500	1,713
Company B	1,306	1,231	1,231 * (12.4% + 1)	1,713
Company C	1,306	1,306 * (15.5% + 1)	(1,306 * (15.5% + 1)) * (12.4% + 1)	1,713
Company D	1,306	1,231	1,231 * (12.4% + 1)	(1,231 * (12.4% + 1)) * (17.9% + 1)

Method 2: Cluster-wise Growth Method

This method involves calculating year-on-year growth (%YoY) of respective financial statement items of all reporting companies in each ISIC section (i.e., %YoY for each ISIC 1-digit level) and applying these percentages to estimate missing data for non-reporting companies in each ISIC section for the three years tested. Table 3 illustrates how missing data were imputed under this method (imputed data are in yellow and blue cells).

Table 3: Imputing missing data under Method 2

(assuming 2020 - 2022 %YoY for ISIC J were 10.8%, 14.2%, and 15.9%, respectively and %YoY for ISIC A & B were 5.3%, 4.7%, and 5.6%, respectively)

Company	ISIC	2019	2020	2021	2022
Company A	J	1,306	1,306 * (10.8% + 1)	1,500	1,713
Company B	B	1,306	1,231	1,231 * (4.7% + 1)	1,713
Company C	J	1,306	1,306 * (10.8% + 1)	(1,306 * (10.8% + 1)) * (14.2% + 1)	1,713
Company D	A	1,306	1,231	1,231 * (4.7% + 1)	(1,231 * (4.7% + 1)) * (5.6% + 1)

Method 3: Hybrid Method

This method applies Method 2 for 'manufacturing and wholesale-retail trade' ISIC sections, and Method 1 for other ISIC sections.

Table 4: Imputing missing data under Method 3

(assuming 2020 - 2022 %YoY were 15.5%, 12.4%, and 17.9%, respectively and for ISIC C were 10.8%, 14.2%, and 15.9%, respectively)

Company	ISIC	2019	2020	2021	2022
Company A	C	1,306	1,306 * (10.8% + 1)	1,500	1,713
Company B	A	1,306	1,231	1,231 * (12.4% + 1)	1,713
Company C	C	1,306	1,306 * (10.8% + 1)	(1,306 * (10.8% + 1)) * (14.2% + 1)	1,713
Company D	A	1,306	1,231	1,231 * (12.4% + 1)	(1,231 * (12.4% + 1)) * (17.9% + 1)

Method 4: Time-Based Method

This method assumes "linear growth" for companies not reporting data for only one year. For companies with missing data for two years or more, Method 2 was applied. Table 5 illustrates how missing data were imputed under this method (imputed data are in purple and pink cells).

Table 5: Imputing missing data under Method 4

(assuming 2020 - 2022 %YoY were 10.8%, 14.2%, and 15.9%, respectively)

Company	2020	2021	2022
Company A	1,500 + (1,500 – 1,713)	1,500	1,713
Company B	1,231	1,231 + ((1,713 – 1,231)/2)	1,713
Company C	1,231	1,500	1,500 + (1,500 – 1,231)
Company D	1,306 * (10.8% + 1)	(1,306 * (10.8% + 1)) * (14.2% + 1)	1,713
Company E	1,500 * (10.8% + 1)	1,500	1,500 * (15.9% + 1)
Company F	1,231	1,231 * (14.2% + 1)	(1,231 * (14.2% + 1)) * (15.9% + 1)

Step 2.3: Assessing efficacy of each method

The statistics used to assess the efficacy of each of the four methods include:

1. Mean square error (MSE)
2. Root mean square error (RMSE)
3. Mean absolute error (MAE)
4. Mean absolute percentage error (MAPE)

The method which yields the lowest value for all (or most of the) four statistics is deemed to have highest efficacy and would therefore be chosen as preferred method, to be applied in actual compilation of the SBS.

Based on statistical tests conducted on the mock-up data, “Hybrid Method” yielded the best results for trade credits receivable. Meanwhile, Method 1 (Overall Growth Method) returned the lowest statistical error for trade credits payable. These two methods were then applied to impute missing data for the entire CPFS dataset for the aforementioned two financial statement items. Completed data were then used for SBS compilation for periods up to 2023.

5. Compiling Quarterly SBS: Constructing quarterly estimates based on annual data

Not all financial statement items and counterpart sectors have quarterly inputs to support quarterly compilation of the SBS. This is particularly the case for selected ONFC's assets and liabilities, for which the annual CPFS data serve as the only source. Thus, estimates of quarterly positions were derived from annual time series, with the assumption that changes to the annual positions occurred at a constant rate and distributed evenly across the four quarters. Table 6 depicts how quarterly series were derived based on this assumption.

Table 6: Example of how to convert annual data to quarterly data

Q4/2021	Q1/2022	Q2/2022	Q3/2022	Q4/2022
1,000	1,000 + ((1,400 – 1,000) / 4)	Q1/2022 + ((1,400 – 1,000) / 4)	Q2/2022 + ((1,400 – 1,000) / 4)	1,400

6. Nowcasting the SBS for the current year

As CPFS data are made available to SBS compilers with relatively long time lag, a method to nowcast the SBS for NFC sector needs to be developed to enable a more timely compilation of the SBS. The following variables were short-listed as conceptually relevant to the targeted financial statement items. These variables are available with significantly less time lag; and hence could serve as indicative information for designing nowcasting models for the targeted items.

- Relevant variables for trade credits receivable: total sales or income (e.g., from financial data of listed companies disclosed by the SET, data on sales reported to the Revenue Department's value-added tax (VAT) database)
- Relevant variables for trade credits payable: production volume, inventory, sales, etc.

CPFS data for 2011 – 2023 were tested for cross-correlation between those short-listed variables and the targeted items (i.e., trade credits receivable/payable). Tests were conducted both on the value of the variables themselves and their year-on-year growth rate. For data on sales, the total sales of all ONFCs as well as sales of those in manufacturing and trading businesses (i.e., ISIC sections C and G, which are deemed to have closer link to trade credits) were tested. The study revealed the following results:

Table 7: Cross-correlation between selected variables and trade credits receivable

For trade credits receivable	Variables' Value	Variables' %YoY
SET	0.66	0.40
VAT	0.90	0.75
VAT C+G	0.88	0.70

Remark : Test period was 2011-2023 (for variables' value), and 2012-2023 (for %YoY)
C denotes manufacturing sector G denotes wholesale and retail sector
Source: SET, The Revenue Department, calculated by the BOT

Table 8: Cross-correlation between selected variables and trade credits payable

For trade credits payable	Variables' Value	Variables' %YoY
SET	0.66	0.39
Production (MPI)	0.87	0.38
VAT_C	0.90	0.70
VAT_C+G	0.88	0.77

Remark : Test period was 2011-2023 (for variables' value), and 2012-2023 (for %YoY)
C denotes manufacturing sector G denotes wholesale and retail sector
MPI denotes manufacturing production index
Source: SET, The Revenue Department, Ministry of Industry, calculated by the BOT

Results in Tables 7 and 8 reflected that "total sales of all business sectors combined" was highly correlated with trade credits receivable; while "sales of manufacturing and wholesale/retail sectors alone" exhibits the highest correlation with trade credits payable. These findings were used as a basis for nowcasting preliminary estimates for ONFC's trade credits receivable/payable. These estimates would be revised thereafter upon availability of the CPFS data of respective year.

7. Conclusion

In recognition of the need for comprehensive tools to monitor inter-sectoral financial linkages and plausible risk spillovers, the BOT has commenced Thailand's compilation of the SBS some years ago, with continuous effort in exploring and integrating supplementary sources to enhance the coverage and quality of the SBS. In an attempt to narrow the data gaps on ONFC's positions, financial statement data reported annually to the MOC were identified as potential data source. Despite its rich information and considerable granularity, limitations exist in regard to the timeliness and completeness of the data reported.

To overcome these limitations, the author team conducted statistical tests to identify appropriate techniques to (i) impute missing values (for companies not reporting trade credits

receivable/payable for certain years); (ii) derive quarterly estimates of the SBS from the annual financial statement data; and (iii) nowcast quarterly SBS for the current year. This, together with the enhanced coverage of other comprehensive data sources (namely, the MFS, the IIP, and securities database) should ensure considerable improvement to the coverage, quality, as well as timeliness of Thailand's SBS.