Revision in the Sample Design of TANKAN using the Economic Census of Japan

Kyosuke Shiotani
Economic Statistics Division
Research and Statistics Department
Bank of Japan

The views expressed are those of the speaker and should not be attributed to the Bank of Japan.
Index

I. Overview of Tankan survey
II. Sample Design of Tankan
III. Improving the Sampling Methodology
IV. Conclusion
**What is TANKAN?**

Tankan = Short-Term Economic Survey of Enterprises in Japan

<table>
<thead>
<tr>
<th>Purpose</th>
<th>Providing an accurate picture of business conditions ⇒ Contributing to the appropriate implementation of monetary policy</th>
</tr>
</thead>
</table>
| Survey Items | • Judgement survey (e.g., Business conditions index)  
• Quantitative data (e.g., Sales, Fixed investment)  
• Inflation Outlook (Output Prices and General Prices)  
• Number of New Graduates Hired |
| Coverage | • Sample enterprises : 11,066 (Jun. 2015)  
• Population enterprises : 212,277 |
| Frequency | Quarterly (March, June, September, December) |
Business Conditions DI

Manufacturing

(Diffusion index, % points)

Large Enterprises
Medium-sized Enterprises
Small Enterprises

"Favorable"
"Unfavorable"

Forecast

Note: Shaded areas indicate periods of recession (according to the Cabinet Office).
Strength of the Tankan Survey

1. High response rate
   - 99.5% (Jun. 2015)
     : keeping a high response rate (95.6%) even immediately after the Great East Japan Earthquake (March 2011)

2. Long history
   - The predecessor of the present Tankan started in 1957.
     : on the model of “Economic test” by the IFO Institute
   - The present Tankan started in 1974.
     : registering the 162th survey in Sept. 2014

3. Statistical accuracy
   - Stratified sampling to reduce sampling errors
   - Regular revision of the sample enterprises

4. Quick release
   - Response collected in one month
   - Released on the day after the end of the survey period

5. Various survey items
   - Judgement survey (13 items)
   - Quantitative data (9 items)
   - Inflation Outlook

The value of the Tankan survey shared among respondents
Credible data
Useful data
Heavily used by policy makers and analysts
Virtuous cycle
Index

I. Overview of Tankan survey
II. Sample Design of Tankan
III. Improving the Sampling Methodology
IV. Conclusion
Calculation Methods

- **Judgment Survey**
  Aggregated into Diffusion Index (DI)

- **Inflation Outlook**
  The percentage share of the number of respondents choosing each alternative

- **Quantitative Data**
  The amount of "population estimates" is calculated by industry and size classifications
## Error Ratio of the Sales

<table>
<thead>
<tr>
<th></th>
<th>Large</th>
<th>Medium-sized</th>
<th>Small</th>
<th>Sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010</td>
<td>1.5%</td>
<td>3.1%</td>
<td>2.6%</td>
<td>4.3%</td>
</tr>
<tr>
<td>2015</td>
<td>1.0%</td>
<td>3.0%</td>
<td>2.7%</td>
<td>4.6%</td>
</tr>
</tbody>
</table>

Tankan has set the targets on error ratio of these 6 categories by industry and scale. 3 % for Manufacturing and 5% for Nonmanufacturing.
Economic Census

- In March 2015 survey, we made the first revision using the “Economic Census” as the population base for the Tankan survey.

- The "Economic Census“ aims to provide a comprehensive overview of the actual business activities of establishments and enterprises in Japan.

- It surveys the economic activities of establishments and enterprises, with its survey items including accounting information. We can access detailed information for the target population of Tankan.
To achieve the accuracy target, we divide strata into smaller segments by “industry,” by “capital” and by “the number of employees” (stratified sampling).

The number of employees was the only effective information to divide strata in the previous population base statistics for the Tankan.
Stratified Sampling

If “sales” and “the number of employees” are highly correlated, we can decrease the population variance of sales for each stratum.

The variance of is smaller than one of.
Why did error ratios increase?

- The *Economic Census* uses administrative record information to collect information on establishments and enterprises, and has resulted in greater coverage of enterprises.

- Using the sales data from the *Economic Census*, the population variance increases particularly sharply for the real estate sector and some other sections.

- Their number of employees are not highly correlated with sales. Thus, it is not an effective characteristic to decrease the error ratio of “sales.”
Sample Bias of the Tankan

Using accounting information for all enterprises in the population, we can check the sample bias of Tankan results compared with the true value.

<table>
<thead>
<tr>
<th></th>
<th>Large</th>
<th>Medium-sized</th>
<th>Small</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sales</strong></td>
<td>2.3%</td>
<td>11.3%</td>
<td>3.3%</td>
</tr>
<tr>
<td><strong>Fixed Investment</strong></td>
<td>3.5%</td>
<td>9.8%</td>
<td>6.2%</td>
</tr>
</tbody>
</table>
Index

I. Overview of Tankan survey
II. Sample Design of Tankan
III. Improving the Sampling Methodology
IV. Conclusion
Classified by Sales

- Applying the “Economic Census” as the population base for the Tankan, sales data for all enterprises in the population is available now. We can use it to divide the population into strata.

- By dividing the population into strata by sales, we can make a perfect ordering of groups by sales and thereby decrease the population variance of sales.

- The error ratios and the sample biases based on the sales classification are lower than those based on the employee classification.
### Error Ratio (Strata Classified by Sales)

<table>
<thead>
<tr>
<th>Items</th>
<th>Classified by</th>
<th>Large</th>
<th>Medium-sized</th>
<th>Small</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales</td>
<td>Employees</td>
<td>1.0%</td>
<td>3.0%</td>
<td>2.7%</td>
</tr>
<tr>
<td></td>
<td>Sales</td>
<td>0.5%</td>
<td>0.8%</td>
<td>1.8%</td>
</tr>
<tr>
<td>Fixed investment</td>
<td>Employees</td>
<td>3.0%</td>
<td>3.6%</td>
<td>6.5%</td>
</tr>
<tr>
<td></td>
<td>Sales</td>
<td>2.4%</td>
<td>3.1%</td>
<td>6.3%</td>
</tr>
<tr>
<td>Item</td>
<td>Classified by</td>
<td>Large</td>
<td>Medium-sized</td>
<td>Small</td>
</tr>
<tr>
<td>-----------------------</td>
<td>---------------</td>
<td>----------------</td>
<td>-----------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>Sales</td>
<td>Employees</td>
<td>2.3%</td>
<td>11.3%</td>
<td>3.3%</td>
</tr>
<tr>
<td></td>
<td>Sales</td>
<td>1.8%</td>
<td>1.6%</td>
<td>1.2%</td>
</tr>
<tr>
<td>Fixed Investment</td>
<td>Employees</td>
<td>3.5%</td>
<td>9.8%</td>
<td>6.2%</td>
</tr>
<tr>
<td></td>
<td>Sales</td>
<td>3.4%</td>
<td>8.4%</td>
<td>10.9%</td>
</tr>
</tbody>
</table>
Overestimation

- There is still a problem of overestimation.
- The Tankan samples in each stratum are biased towards enterprises with large number of employees.

Tankan Estimation (e.g. Sales) > Population Value
To reduce Statistical Bias

- The main reason is a difference in the acceptance rate for survey across enterprises with different numbers of employees.
- The acceptance rate for enterprises with very few employees is lower than for enterprises with dozens of employees owing to their insufficient resources for accounting and responding the Tankan survey.
- However, adding and eliminating a significant number of enterprises are controversial in as much as the post-revision results would end up differing significantly from the pre-revision results.

Acceptance rate: the ratio of the enterprises that agree to be in the Tankan sample to all enterprises that are asked to be in the sample.
## Decreasing the Sample Bias

<table>
<thead>
<tr>
<th>Item</th>
<th>Sample Size</th>
<th>Large</th>
<th>Medium-sized</th>
<th>Small</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sales</strong></td>
<td>11,126</td>
<td>1.8%</td>
<td>1.6%</td>
<td>1.2%</td>
</tr>
<tr>
<td></td>
<td>10,826</td>
<td>0.4%</td>
<td>0.6%</td>
<td>0.3%</td>
</tr>
<tr>
<td><strong>Fixed Investment</strong></td>
<td>11,126</td>
<td>3.4%</td>
<td>8.4%</td>
<td>10.9%</td>
</tr>
<tr>
<td></td>
<td>10,826</td>
<td>3.2%</td>
<td>5.1%</td>
<td>9.1%</td>
</tr>
</tbody>
</table>

Eliminate just a few percent of enterprises, focusing on those that contribute the most to overestimation of the number of employees.
I. Overview of Tankan survey
II. Sample Design of Tankan
III. Improving the Sampling Methodology
IV. Conclusion
Conclusion

1. Using data from the “Economic Census,” we find that the true error ratio is higher than previous estimation, and that the Tankan results overestimate the population values.

2. By stratifying based on sales, we can decrease both the error ratio and the sample biases reflected in the Tankan results.

3. To eliminate just a few percent of enterprises that contribute the most to overestimation of the number of employees, we can decrease the biases for sales and fixed investment.