For Official Use **STD/STESEG(2003)11**

Organisation de Coopération et de Développement Economiques
Organisation for Economic Co-operation and Development

10-Jun-2003

English - Or. English

STATISTICS DIRECTORATE

STD/STESEG(2003)11
For Official Use

OECD Short-Term Economic Statistics Expert Group

STESEG TASK FORCE ON TIMELINESS AND BENCHMARKING

CASE STUDY: TIMELINESS OF RETAIL TRADE STATISTICS IN SWEDEN, ITALY AND THE UNITED STATES

Paris, 26 - 27 June 2003

Paper prepared by:

Richard McKenzie, OECD Statistics Directorate
Anita Ullberg, Statistics Sweden
Pietro Gennari, ISTAT

1. This document is submitted to the Expert Group under item 5 of the Draft Agenda

2. The Expert Group is invited to:

NOTE the analysis in the body of the document

COMMENT on the findings of the case study presented in paragraphs 27 - 29.

Contact: Richard McKenzie (richard.mckenzie@oecd.org), Statistics Directorate, OECD

JT00145788

Document complet disponible sur OLIS dans son format d'origine
Complete document available on OLIS in its original format

English - Or. English

STESEG TASKFORCE ON TIMELINESS AND BENCHMARKING

Case study: Timeliness of retail trade statistics in Sweden, Italy and the United States

a. Introduction

1. The purpose of this paper is to examine the impact on timeliness due to the application of different methods in the statistical production process for retail trade statistics in Italy, Sweden and the United States. Retail trade was chosen for the study as it is an important short-term economic indicator where timeliness is considered a key criterion amongst users. Also, the countries under consideration have large differences in timeliness for this statistic which facilitates a meaningful analysis.

2. It should be clearly stated that this paper is not concerned with comparing the quality of retail trade statistics for the countries involved in the case study. Its entire focus is to identify differences in methods used for aspects of the statistical production process and comment on their impact on timeliness, including any possible trade-offs with accuracy. This is intended to aid the STESEG in identifying methods for improving timeliness across a broader range of short-term economic indicators.

b. Overview of timeliness for the US, Sweden and Italy

3. The approach used to produce retail trade statistics in the US, Sweden and Italy differs significantly, as does the timeliness of release of these statistics. The US publishes *advanced* estimates at aggregate levels 9 working days after the reference month, which are subsequently revised with the publication of preliminary estimates at a detailed level 6 weeks after the reference month. Final estimates at the same level of detail and with very little revision from preliminary estimates are published 10 weeks after the reference month.

4. Sweden has recently altered their processes for producing retail trade statistics and now releases preliminary estimates at aggregate levels 28 days after the reference month. Final data are released with the subsequent issue (i.e. at 58 days) at a more detailed level, but are based on a very similar response rate and thus revisions are small. Italy publishes a retail trade index disaggregated at a number of different levels approximately 8 weeks after the end of the reference month. ISTAT are participating in the European project of a 'country stratified sample', whose main purpose is to produce preliminary estimates at EU level 30 days after the reference month. ISTAT are currently investigating the possibility of producing a preliminary aggregate national index from the Italian businesses in the EU sample, which are a sub-sample of their larger national sample survey. This project is at an initial stage with experiments being carried out on both data collection and index calculation.

c. Comparison of statistical production processes

5. The following section outlines some of the differences in key statistical production processes between the three countries, and comments on the influence these differences may have on the timeliness and accuracy of the statistics produced.

c.1 Sample design and selection

6. Similar methods to obtain the main sample are used by each country. An annual frame is stratified according to industry and business size, from which a sample is drawn. The US make use of historical administrative data on retail sales by performing probability proportional to size sampling within strata, whereas Italy and Sweden both use simple random sampling. The US also supplements their sample each quarter to represent births to the population. Each country includes the largest businesses in the sample with certainty.

7. The main difference between the countries is that the US also selects a sub-sample from which their advanced estimates are compiled. The full sample is re-stratified to reflect the broader level of detail required for publication of the advanced estimates and the sub-sample is then chosen with probability proportional to size. The sub-sample size is approximately 5,000 compared to the full sample size of 12 000.

Impacts on timeliness and accuracy

8. 8 The US approach of supplementing their sample each quarter to represent population births introduces a complexity that could have some impact on response rates (i.e. it is generally more difficult to obtain responses from new respondents) and thus put extra pressure on timeliness. However, the selection of a sub-sample by the US is a methodological technique specifically designed to produce aggregate estimates at an earlier time frame in comparison to the full sample.

c.2 Data items, reference periods and due dates

9. Each country sends a questionnaire by mail to their respondents. The US (total sales) and Sweden (turnover) collect only 1 data item, whereas Italy collects six variables: turnover by type of outlet (e.g. large store, small specialty store etc.) and group of product sold for the reference month of the current year and same month of the previous year; number of persons employed; total selling surface; number of outlets and; number of working days in the month. Italy also collects qualitative information on sales expectations for the following month.

10. In Sweden, there is a legislative regulation that businesses must answer the survey within 15 days of the end of the reference month. This was recently reduced from 25 days after a study showed that the required data should be available from most businesses in accordance with this earlier time frame. In Italy, the survey is compulsory by Italian law, and respondents are asked to answer within 10 days after the reference month. In the US, the survey is voluntary. For the sub-sample, the questionnaire is sent to selected businesses 5 working days prior to the end of the month and the due date is 3 working days after the end of the month. For the remaining sample, the questionnaire is sent at the end of the reference month, and the data is due on the 12th day of the following month.

Impacts on timeliness and accuracy

11. The major difference between the countries is the additional data items collected in Italy. This is currently required to support both their estimation technique and level of detail required for published estimates (see paragraphs 21 and 24). However this is likely to have a major impact on timeliness, due to the extra burden placed on respondents to complete the survey, and the time required for businesses to compile the required data (e.g. for turnover by product, which mainly affects larger businesses). The US approach for their sub-sample relies on early notification through the despatch of their questionnaire 5 days before the end of the reference month, and the availability of data in businesses records very soon after the end of the reference month (i.e. 3 days). The fact that around 65% response is obtained (see paragraph 17)

indicates that many business can provide retail trade turnover very soon after the reference month, at least in the US.

c.3 Methods of data collection, non respondent follow up and data validation

12. Sweden has implemented touchtone data entry and 60% of their respondents respond in this way. Other responses are received by telefax and mail. Italy receives approximately one third of their responses by telefax and the rest by mail, whereas the US predominantly rely on response by mail.

13. Sweden sends mail reminders to non-respondents one week prior to the due date, and then telephones non respondents over a 3-4 day period 1 day after the due date. The largest businesses receive priority in telephone reminders, to try and ensure they contribute to estimates. The US also performs intensive telephone reminders for a 3-4 day period 1 day after the due date for their sub-sample. For the remaining sample, telephone reminders commence after the due date (i.e. the 12th of the month) and continue until about 1 week before the preliminary estimates are published. Italy use a similar approach, beginning telephone reminders approximately 1 week after the due date and continuing to collect data until 1 week before the publishing date.

14. Each country applies validation rules to data at the micro level, and will query respondents if anomalies are identified. Sweden has the added advantage of automating data validation in conjunction with data capture where touchtone data entry is used by the respondent.

Impacts on timeliness and accuracy

15. The use of touchtone data capture with automated data validation by Sweden has a major influence on improving timeliness, through avoiding the delay in mail receipt and minimizing separate follow up with respondents to validate data. This method is particularly suited to surveys with a small number of data items. However, given the amount of information collected at present for the retail trade survey in Italy, the use of touchtone data capture may not be possible.

16. The use of telephone reminders where resources are concentrated on this process for a short period of time immediately after the due date also benefits timeliness. In addition, the approach by Sweden of prioritizing telephone reminders to the largest businesses assists in minimising any loss in accuracy resulting from attempting to produce the timeliest estimates.

c.4 Response rates, estimation methods, level of detail published and revisions

First estimates

17. The US advanced estimate, published 9 working days after the reference month, is based on a weighted response rate of approximately 65%. Estimates for each level to be published (i.e. based on the industry stratification of the sub-sample) are developed by applying a ratio of current month to previous month sales (derived from the sub-sample) to the preliminary estimate of sales for the previous month (derived from the full monthly sample). Industry estimates are then summed to derive the advanced estimate of total retail sales. Data are released at aggregate, 14 sub-aggregate and 3 detailed levels consistent with the stratification of the sub-sample.

18. The Swedish preliminary estimate, published 28 days after the reference month, is based on a response rate of 75% which equates to a weighted response rate of 90%. A ratio estimation methodology is used, with a minor adjustment for the higher proportion of nil respondents (e.g. businesses in the sample which have ceased operating) contributing to the preliminary estimate. Non-responding large businesses

(i.e. those selected in the sample with certainty) are imputed using the most recent VAT data available for the business and / or previous survey responses. Data are released at aggregate and 3 sub-aggregate levels.

Final estimates

19. The US preliminary (published 6 weeks after the reference month) and final (10 weeks) estimates are usually based on the same sample of responding businesses, with a weighted response rate of approximately 80%. Estimation is based on inverse probability of selection weights, and explicit imputes are used for the non-responding businesses. These imputes are derived from a non-response model using administrative and other survey data (e.g. models combining historical data for the non-responding business and survey data for reporting businesses). Data are published for 57 categories, which represent various levels of the industrial classification.

20. Swedish final estimates, published 58 days after the reference month, are based on a response rate of 78% which equates to a weighted response rate of 92%. Ratio estimation is used with imputation for non responding large businesses as described above for preliminary estimates. Data are published for 18 categories of estimates which represent various levels of the industrial classification.

21. Final estimates for the Italian Retail Trade index, published 8 weeks after the reference month, are based on a response rate of approximately 50%. Estimation is in the form of an index, to minimise the bias caused by the high level of non response. Non-respondents are imputed each month using a model based on historical survey data where this exists, or the average change for responding businesses in the same stratum where no historical data exists. Data are published by 15 groups of products, type of outlet (5 types of large distribution outlet, small specialised stores, and total), 5 employment classes and 4 geographical macro-areas (north-west, north-east, centre, south-islands).

Revisions

22. For the US, the average absolute revision in the total retail sales estimate between the advanced estimate and the preliminary estimate is two-tenths of one percent. Significant effort has been expended over the years to refine the methodology for the advanced estimate to minimise this revision. Differences between preliminary and final estimates are very small as they are generally based on the same set of respondents and estimation methodology. Final estimates are then revised when new annual benchmarks become available. The annual total of retail sales in 2001 derived from the sum of the final monthly estimates was revised by 0.4% after the release of the annual benchmark for 2001.

23. The average revision between preliminary and final estimates in Sweden is approximately one-tenth of one percent. Only a small revision is expected given that estimates are based on a similar weighted response rate (i.e. 90% and 92%).

<u>Impact on timeliness and accuracy</u>

24. The estimation process used in Italy is dependent on data collected in the sample, to support the level of detail published (i.e. by type of product sold, type of outlet & employment class) and assist in imputation due to the high level of non response (e.g. collecting data from the same month of the previous year). The high level of non response is partly due to the different structure of the Retail Trade industry in Italy compared to Sweden and the US, where the share of small firms (who obtaining response from is generally more difficult) is much higher. Consequently the choice of estimation methodology (i.e. index form) is also influenced by the response rate, with the need to minimise bias. These requirements significantly impact on the timeliness of estimates that can be achieved.

25. The estimation process used for advanced estimates in the US is completely different to that used for preliminary (and final) estimates and demonstrates that reasonable quality estimates can be produced from relatively low response rates (i.e. 65%). The focus in the US has been on refining the estimation methodology for their advanced estimates to minimise revisions with preliminary estimates. Estimation methodologies for final (and thus preliminary) estimates are also reviewed to minimise revision between final estimates and subsequent annual benchmarks. The process of producing subsequent more reliable estimates for the same variable facilitates this kind of analysis, but of course has associated resource implications.

26. Sweden has opted to refine their operational procedures to improve timeliness, using very similar estimation methods for preliminary and final estimates due to the similar responses rates (i.e. weighted response rates of 90% and 92% respectively). This results in slightly lower average revisions at aggregate level compared to the US (i.e. 0.1 percentage points compared to 0.2). However Sweden's preliminary estimates are published approximately 2 weeks later than the US advanced estimates. This represents an interesting tradeoff between timeliness and accuracy when comparing the estimation processes for first estimates in Sweden and the US.

d. Conclusion: Important observations on timeliness derived from the case study

27. A number of issues related to improving timeliness within the statistical production process for short-term economic statistics (STES) in general can be drawn from this case study. Some of key points identified are:

- Sub-sampling can be effectively incorporated within the statistical production process to support the calculation of early estimates. That is, whilst businesses included in the sub-sample are processed according to an earlier timetable in comparison to the remaining sample, workflows can be designed to manage this efficiently.

- Ongoing improvements in businesses use of technology in their own reporting systems may result in the data required by statisitical organisations being available within an earlier time frame. It is important that statistical organisations are aware of these developments and exploit any opportunities for improving the timeliness of STES.

- Minimising the number of data items to ensure only those essential for estimation are collected can assist in reducing response burden and facilitate the use of more effecient forms of technology based processesing (e.g. touchtone data entry).

- The use of technology based collection techniques can create greater flexibility in the statistical production process and enable the streamlining of processes (e.g. combining data entry and data validation through touchtone data entry or computer assisted telephone collection).

- Concentration of staff for a short period of time (e.g. 3-4 days) on telephone reminders for non-responding businesses is an effective method to improve response rates to an acceptable level to publish early estimates at an aggregate level.

- Estimation techniques can be specifically designed to produce acceptable quality estimates at aggregate levels from relatively low response rates. This can support the publication of early estimates for STES (e.g. calculation of percentage movement estimates only, or calculation of an index).

- Ongoing analysis of revisions can be used to refine and improve methodologies used for early estimates of STES.

- Different users of STES have different requirements in regards to level of detail and timeliness of published estimates. It is possible to satisfy these different needs and maintain effeciency within the statistical production process.

28. The case study also revealed that each countries statistical organisation is influenced by a range of different external factors. For example in Sweden there is a greater tendancy for businesses to conform to legal requirements than in Italy, which assists in improving response rates. Conversely, many statistical surveys in the US are voluntary and the methodologies used for their estimation processes are specically designed to account for this.

29. A consequence of this is that methods for improving timeliness, such as those implied in the points above, may be more relevant or applicable to some countries than others. Understanding the methods available for improving timeliness, and assessing which methods could be implemented for minimum cost or have maximim effect within their country, is the key challenge facing national statistical organisations wanting to improve the timeliness of the STES they produce.

Appendix Summary of differences in statistical functions for retail trade in Italy, Sweden and the United States

Country Statistical function	Italy	Sweden	United States
Sample size and scope	Sample 7,100. Full coverage of population.	3,100. Full coverage of population.	Full coverage of population. Sample size of 12,000, sub-sample size of 5,000
Sample selection method	Stratified simple random sampling, annual sample.	Stratified simple random sampling, annual sample.	Stratified probability proportional to size sampling, annual sample supplemented for births each quarter
Number of data items collected	6 – turnover (current month & same month previous year), employees, selling surface, number outlets & number of working days.	1 - turnover	1 – total value of sales
Reference period	Month	Month	Month
Data collection method	2/3 mail, 1/3 telefax	60% touchtone data entry, 40% combination of mail, phone & telefax	Virtually all mail
Use of administrative data	No	VAT data can be used to assist imputation of large businesses	Administrative data used in imputation models for non respondents (main sample only)
Due date	10 days after the reference period	15 days after reference month (specifically legislated)	3 working days after reference month for sub-sample, 12 days for remaining sample
Follow up of non-respondents	Telephone reminders 1 week after due date over a 5 week period.	Mail reminder 1 week prior to due date. Intensive telephone reminders of 3-4 days 1 day after due date. Large businesses receive priority to ensure their responses.	Intensive telephone reminders of 3-4 days 1 day after the due date for sub-sample. Telephone reminders after due date for remaining sample over 3 week period.
Data validation methods	Coherence rules based on historical data and structural data (i.e. for Retail industry) applied to each questionnaire received.	Automated querying from touchtone system based on fixed coherence rules (also applies to data received by mail etc.). Graphical analysis.	Coherence rules applied to each questionnaire received.
Response rates for first estimates & estimation method	Only one version of estimates produced	75% response, 90% weighted response. Ratio estimation used with adjustment for higher proportion of nil respondents.	65% weighted response. Ratio of current to previous months sales for the sub sample, applied to previous months estimate
Level of detail released & timing for first estimates	Not applicable.	Aggregate and 3 sub-aggregate estimates released, 28 days after reference month.	Aggregate and 18 sub-aggregate estimates released, 9 work days after reference month.
Response rates for final estimates & estimation method	50%. Index estimation methodology used to compensate for low response rates.	78% response, 92% weighted. Ratio estimation	80% weighted response. Inverse probability of selection weights.
Imputation methods	If previous data available, apply stratum average growth rate since last data point. If no previous data, impute using stratum average.	Large businesses are imputed using a combination of previous survey data and VAT records. This applies for both first and final estimates.	Complex non-response models are applied using administrative and survey data. This is only done for estimation based on the main sample (i.e. not for the sub-sample).

Appendix (cont.) Summary of differences in statistical functions for retail trade in Italy, Sweden and the United States

Country Statistical function	Italy	Sweden	United States
Level of detail released & timing for final estimates	15 groups of products by 6 types of outlet, 5 employment classes and 4 geographical areas released 8 weeks after the reference month	18 categories of estimates released, 58 days after the reference month.	57 categories of estimates released, 6 weeks after the reference month (official final estimates are at 10 weeks).
Average revision between first and final estimates	Not applicable	One tenth of one percent (i.e. 0.1 percentage points of reported growth rate)	Two tenths of one percent (i.e. 0.2 percentage points of reported growth rate)
Methods of dissemination	Web / PDF & database. Annual paper publication.	Web / PDF & database.	Web / PDF & database. Annual paper publication.
Delays between management approval and release of estimates	2 days	1-2 days	1 day for advanced estimates.