Turnover in Other Services and Production in Construction

Improvement of Timeliness

Final Report

Subject: Community grants in the field of short-term statistics

Re: Grant agreement between the European Community and Statistics Finland – Tilastokeskus under the contract **Eurostat nr. 2003 44401014**

Index

1 Introduction

In the past few years a need for further improvement of the short term statistics of the European economy has been recognised. Both national and European users have expressed requirements for more timely, more frequent and more accurate data of the economy than the national statistical offices have been able to compile.

To achieve the needed improvements, European Commission prepared an amendment to the STS regulation 1165/98. The commission proposal for the amendment regulation was presented in May 2004. As the changes in requirements for the timeliness and reference period for other services and production in construction were mostly clear already in autumn 2003, Statistics Finland started this project to improve the compilation of these statistics. The main demands were:

1. To be able to compile turnover indices of the NACE 2-digit divisions in other services in less than 60 days after the end of the reference period instead of the 75 days possible at the time
2. To change the reference period of the production indicator for construction to a month instead of a quarter.

Partial survey based indices were already calculated between T+45 and T+52, mainly to give a one-month headroom for the treatment of most influential split-offs and mergers before the actual calculation at T+75 to T+82. The more timely indices were not disseminated, as their quality was not good enough. To improve quality, major changes were needed in the applied statistical methodology. Enlargement of the collection was not an option, since the statistics are mainly based on administrative data. The direct inquiries, though timely enough, were designed to complement the administrative data rather than to be used as stand-alone sources.

Two major new features were implemented:

1. Automatic imputation of missing data on the level of statistical units
2. Automatic time series estimation of missing data on the aggregate level

Both features required pioneer work as these types of methods were mainly unexplored in the context of short term business statistics. Introduction of estimation and imputation forced also redevelopment of some other parts of the compilation system. Perhaps even a greater development however, was the evolution of human capital, i.e. learning how to deal with less information than before.

As a result of the project, sufficient accuracy was reached for the statistics to be compiled in 45 to 52 days timeframe to be transmitted to Eurostat as preliminary figures. The variables affected as mentioned in the STS regulation are

For construction: 110 Production, 115 Production of building construction and 116 Production of civil engineering. For other services: 120 Turnover (NACE divisions 55, 60, 61, 62, 63, 64, 72, 74 and 72+74).

Transmission of NACE divisions 50-51 was already timely enough.

As agreed in the contract agreement, first delivery was in January 2005 (reference month November 2004). From this on, the production indicator for construction will be transmitted to Eurostat monthly with an approximately 50 days delay. The monthly turnover index for other services is transmitted with 45 days delay four times a year, after each quarter is full (transmissions on February, May, August and November). That is, for other services Finland does not transmit indicators with quarter as the reference period. Rather, the monthly indicators are sent every three months. The first transmission January covered the whole quarter IV/2004, when in the same time of the previous year the last month in transmission was November. In other services the regular monthly transmissions not on February, May, August or November are unaffected by the project and their delay is T+75.

2 Progress of the project

Planning of the project and some initial research work begun as early as October 2003. The project was set in December 2003.

The project was divided into preparatory and project phases. The preparatory phase begun in December 2003 and ended in late May 2004. The project phase begun in May 2004 and ended in February 2005. A project team of three senior statisticians and one statistician was formed to implement the project. A project management team was formed to support the project team.

This chapter is divided into two sections. Section 2.1 describes the preparatory and section 2.2 the project phase.

2.1 The preparatory phase (December 2003-May 2004)

According to the grant agreement the tasks for the preparatory phase were:
- To plan the entire project in detail
- To analyse the quality of the estimation method in quarterly indicators. To assess, whether the direct inquiry supplemented with estimation methods in quarterly indicators can be used as such in this project.
- To study possible improvements to the estimation system or alternative methods for estimating the missing observations
- To evaluate the quality of the currently available inquiry data and if required, to re-evaluate the selection of enterprises in the inquiry.
- To test the estimation methods for monthly indicators and to select the method to be implemented.

The execution of the tasks was overlapped because there was little interdependence between them.

2.1.1 Writing the project plan (January-February 2004)

The basic structure and timing of the tasks for the project was defined in the project plan, which was written in January 2004 and approved by the project management team in February 2004. Plans for the project phase were updated later in the spring before the project phase started.

The project had possible effects to the compilation of Finnish national accounts, since the indicators for production of construction and the turnover for other services are used as it’s source statistics or as reference. No other direct effects were identified and the expectations related to the compilation of national accounts did not contradict to the aims of the project.

It was obvious that the project may have indirect implications to other turnover indices, most importantly to the turnover of industry. A similar speeding up of this variable would become feasible, since the now improved compilation system is used also for the calculation of turnover of industry.

It was also identified that the project had some effects to direct data collection at Statistics Finland. The project would require some increase in the resources for data collection, as the direct inquiry would be the only source of data available in the accelerated schedule.

2.1.2 Analysis of the situation preceding the project (February-March 2004)

A part of the analysis was to estimate, how accurate the partial survey based T+45 indices would be without any modifications to the calculation system. Accuracy was estimated using historical revision data. The reasons of the revisions were also considered.

The project team investigated possible improvements to data validation methods and to sampling of enterprises in the direct inquiries. The structure and coverage of the inquiries were analysed as well as the effects of non-responses. Some minor modifications to the inquiries were made. The last subtask of the analysis was to evaluate the performance of the former imputation method developed for the calculation of quarterly turnover indicators at T+50. The results of these analysis, along with some elaboration, are presented later in chapter 3.

2.1.3 Definition of actions for the project phase (March-May 2004)

The direct inquiry could not be substantially expanded, so it's degree and effectiveness of utilisation had to be increased. The analysis suggested that it would be helpful to develop a new tool for selecting the enterprises into the inquiry. This action was one of the first improvements made to the compilation system during the project phase.

The existing imputation method (described in section 3.3.3) utilised a mixture of ARIMA-models and an *ad hoc* model of three month average year-on-year changes. Analysis showed that although the method improved accuracy of some indices, it was not usable in all cases. Another problem was technical implementation. The method was not integrated to the calculation system and turned out to be too slow to be used in the fast-paced calculation of the preliminary indices.

Due to the imperfections of the existing method, but also based on the encouraging results, it was decided that an improved imputation/estimation method should be designed. At this point became also evident that two different approaches would have to be taken: imputation of enterprise level data and ARIMA-model based estimation of aggregated data. The decision was made mostly for the sake of computational efficiency and to increase selection of possible ARIMA-models to be tested.

Finally it was decided that practices of data validation should be reconsidered. It was clear that not only the correctness of reported figures should be considered, but also how well the respondents represent the population at each point of time. The project team thus expressed a need to explore new ways to handle the data and to combine different sources of information.

The analysis made during the preliminary phase were collected to an intermediate report, which was approved by the project management team on 3rd of April 2004.

2.2 The project phase (April 2004-February 2005)

According to the grant agreement the preliminary tasks for the project phase were:

- To plan the changes required in the current production system with respect to estimation of turnover, and change from quarterly to monthly reference period in the deflator (prices) and production for construction. To implement a test version of the software needed and to draft the system documentation.
- To draft work guidelines to personnel. To plan the production timetable and the work process for the indicators.
- To test the entire system with the test version of the software using actual data. Testing must be started at least six months before the start of regular publish-

ing. This task includes quality control, finalisation of the system and documentation.

- To prepare the GESMES/TS data transfer procedure to Eurostat.
- To calculate monthly time series for production indicators for construction starting January 2000.
- To calculate the indices for the first transmission and to transfer them to Eurostat with GESMES/TS coding.
- To prepare the final report to Eurostat.

The execution of the project phase was straightforward. No changes to the planned tasks were made.

2.2.1 Planning the production timetable (April 2004)

The amendment STS regulation states that the transmission deadline for the production index for construction is one month and 15 days and for the turnover index for other services two months. The timetable for nationally published monthly turnover indices is approximately two months and 15-20 days. The monthly indices are calculated once a month. Therefore the natural timetable for the calculation of preliminary indices is one month and 15-20 days for both of construction and other services.

It was noted that this will require some more effort in the calculation. However, the increase in efficiency of the calculation system, data collection etc. will eventually cancel this effect.

2.2.2 Development and documentation of the imputation and estimation methods (April-October 2004)

Most of the effort was put into the development of the imputation and estimation methods and the working practises. First version of the ARIMA-estimation method was introduced in March 2004. The imputation method was integrated to the system in June 2004. Both methods were continuously evaluated and improved as new ideas emerged. User documentation for the methods was written as the methods were developed.

2.2.3 Updating and improving the selection of enterprises in the inquiries (April-October 2004)

No unified procedure or tools for following the selection of units in the inquiry existed before the project. As a remedy, the project developed a program, which lists the most and least influential units in different activity classes and suggests modifications to the inquiry. The inquiries were evaluated and modified, but the sizes remained unchanged at 200 units in construction and at 600 units in other services (excluding trade).

2.2.4 Defining the new and improving the old working practises and guidelines (April-December 2004)

Test calculations with the T+45 timetable were started as soon as possible, in March 2004. From there on, the project team had extensive discussions on data validation methods, selection of enterprises in the inquiry and other working practices. The discussions took place both in the official monthly meetings of the project team and unofficially. First draft of working guidelines for the imputation and estimation methods was written in summer 2004. As new features were added to the methods and new practises discovered, the document was updated up till winter 2004.

2.2.5 Testing the estimation and imputation methods (June 2004-January 2005)

Since there was little experience of the methodology of imputation and ARIMA-estimation in the context of the statistics in question, testing was seen as one of the most important tasks for the project. Preliminary tests for the estimation method started as soon as early spring 2004. Systematic evaluation of the methods started in June 2004, as the imputation method was introduced. During the tests the project team identified several points where the methods could be improved. The most important modifications were ARIMA-estimation via subaggregates, diagnostics of the imputation method and some enhancements to it.

Partly as a quality check and partly as a possible future development to the estimation method, dynamic regression models were tested as an alternative basis for the estimation method. The performance of these more general models was not significantly better than the ARIMA-models. Keeping reliability in mind the project decided to use the simpler, well established ARIMA-methodology. It was noted that a careful parameterisation of the dynamic regressions could improve accuracy, but this would require more research.

2.2.6 Preparation of the GESMES/TS transfer (January 2005)

The preparations for GESMES/TS transfers took place in December 2004 and in January 2005. Most of the work was done to modify the transmission of production indicator for construction from quarterly to monthly. The transmission of the turnover index for other services was changed so that the preliminary month is transmitted confidentially after each quarter is complete. Activity class 72+74 was added to the other services’ transfer at the same time.

3 The compilation methodology prior to the project

The compilation system which was improved in this project, has evolved continuously during the last 5-7 years. Also this project was evolutionary in nature. Therefore, it is necessary to describe the status of the methodology prior to the project.

The chapter is divided into six sections. Section 3.1 is a short history of the compilation system and section 3.2 an overview of it. Sections 3.3-3.6 focus on different aspects of the system, such as data processing and validation, index calculation et cetera. Section 3.7 presents the performance of the compilation system before the project.

3.1 Historical background

The methodology for both the production indicator for construction and the turnover indicator for other services was developed during the late 1990’s to fulfil the requirements of the EU’s regulation concerning STS-indicators (No. 1165/98). National legislation, on the other hand, regulated that data collection for statistical purposes shall primarily rely on administrative sources (Statistics Act 280/2004). For this reason, and also for the benefit of economic efficiency, the main data source for these statistics is the VAT data collected by the National Board of Taxes. The data contain monthly turnover figures from the beginning of year 1995 for all businesses, which are liable to pay VAT in Finland.

However, the VAT data could not be used solely without the direct inquiry conducted by Statistics Finland. The inquiry ensures the relevance of the figures by a more rigorous definition of the turnover variable and by dividing the most important enterprises to kind-of-activity units. Direct contact to the data providers is also essential to get information about changes in activity and/or firm’s structure.

When the index calculation method was being developed, it was obvious that the conventional methods designed for situations where data is collected by the statistical bureau were not an option. Most importantly, the developers were confronted by the massiveness of the VAT data and it’s relatively slow accumulation. The two main solutions for the problem were estimation of changes instead of levels and graphical data validation. Firstly, it was realised that estimation of changes can be easily carried out from the VAT data with reasonable accuracy. Secondly, it was noted that the data could be checked most efficiently by reviewing the most influential enterprises using a graphical interface.

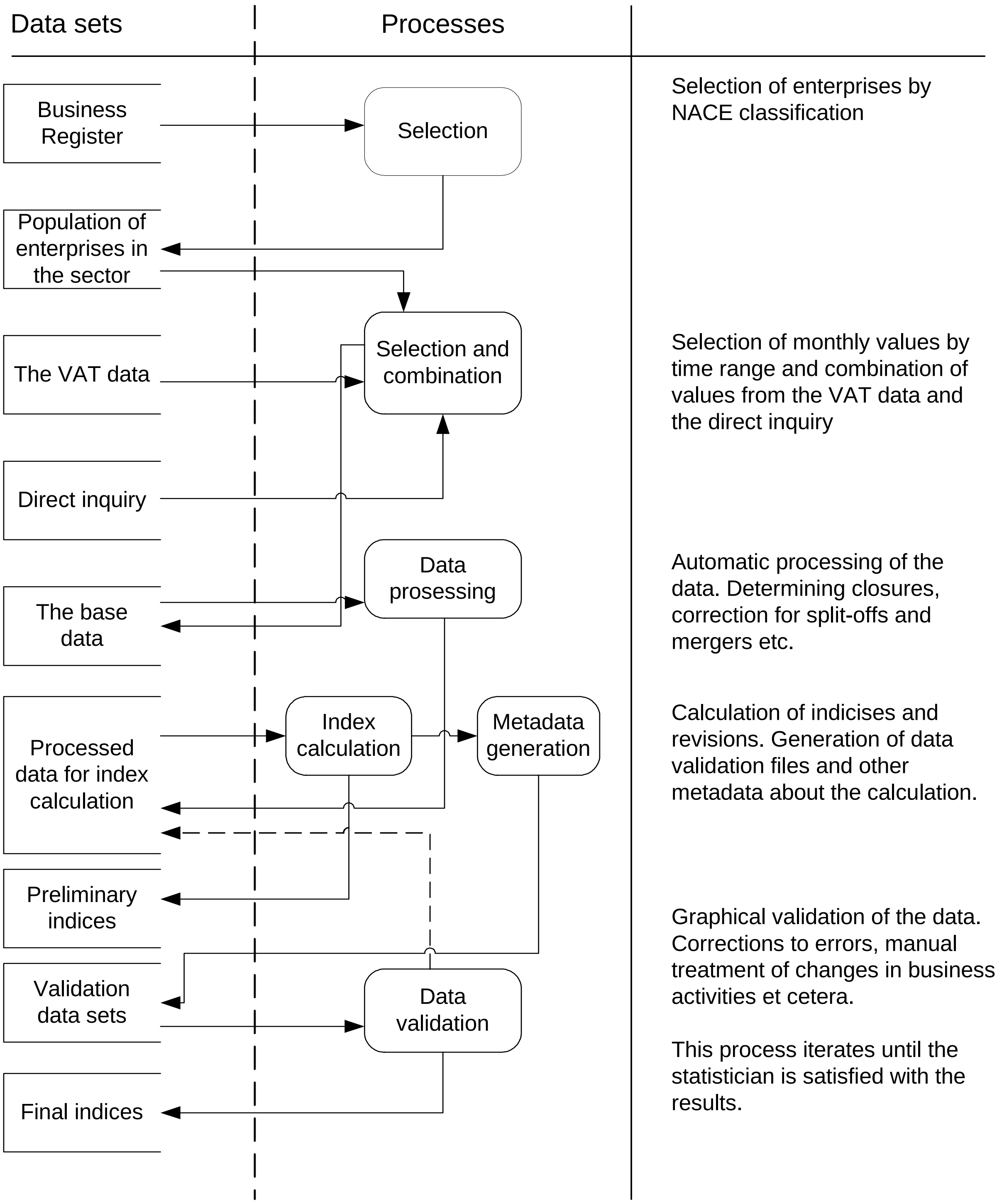
After the initial start of the methodology, development effort was used to improve quality and usability, but eventually the biggest pressure was to accelerate the production. In 2001 there was a project to develop an imputation method for calculation of quarterly turnover indicators with T+50 days delay. The indicators were to be used for quarterly national accounts. In 2002 Eurostat started a project to develop a country-stratified European sample for retail trade index. Finland took part in this project, and accelerated the turnover index for retail trade from approximately T+52 to T+27. Because of the structure of the trade sector, no imputation or estimation was necessary and the calculation could be based solely on the direct inquiry. Having said that, data validation methods and the treatment of mergers, split-offs etc. were improved considerably in this project.

Another notable development was the electronic data collection, which started in 2003. In 2004 approximately 70% of the enterprises in the inquiries responded via the internet. This has made improvement of timeliness much easier, because data collection has become less laborious and more resources have been released for quality analysis.

3.2 The compilation system in general

Picture 1 is a simplified illustration of the compilation process prior to the project.

Picture 1: The compilation process



The basic datum, i.e. Business Register, the VAT data and the data of the direct inquiries are stored in Sybase databases. The calculations are performed by SAS-programs run in an UNIX framework. Data validation is done in Excel with the help of a visual interface programmed in the VBA language.

The most characteristic parts of the process are described in the following sections. Some parts are skipped if they are not essential with respect to the work done in the project.

3.3 Data processing

Estimation of change has several good features. For example stratification of the data set is not essential, though it may in some cases improve accuracy, and non-responses or errors of less influential units can be treated by leaving them out of the calculations. However, because the estimation relies on information from two separate points of time, structural changes in business activity between the points of time has to be taken into account. The largest enterprises and the biggest changes of turnover are checked manually, but because the VAT database is very large, most data processing is automatic.

The next two sections describe the treatments of changes in business activity and the selection of data source. Section 3.3.3 describes the previously used imputation method, which was the base of the development of the new imputation method.

3.3.1 Treatment of changes in the structure of business activity

Structural changes in activities are stored in a special database table. Information of the events is collected directly from data providers and indirectly from the media.

The data used for index calculation contains for each month and for all the units the turnover T_t of the current month and turnover T_{t-12} of the same month of the previous year. Because of the index calculation method, the values have to be comparable – they have to represent the firm's turnover's change in the activity class in question. When, for example, there is a split-off, the previous year's turnover has to be correctly allocated to the new firms.

The problem is solved by dividing T_{t-12} by each firm's proportions of the sum of the current month's turnover. In the general case, when a change in business activity involves firms $A_1,...,A_n$ splitting and/or merging into $B_1,...,B_m$ (which are not necessarily in the same activity class), the formula for $T_{B_i,t-12}$, firm B_i:s previous year turnover is

$$T_{B_i,t-12} = \frac{T_{B_i,t}}{\sum_{i=1}^{m} T_{B_i,t}} \sum_{i=1}^{n} T_{A_i,t-12} = d_{B_i,t} \sum_{i=1}^{n} T_{A_i,t-12} .$$

As an example, consider a firm *A* which splits into three parts. If all parts have equal turnovers for month *t*, the previous year's turnover for each new firm is also equal: one third of *A*'s turnover for month *t*-12. On the other hand, if one of the new parts is twice as large as the other two, also the previous year's turnover allocated for it is twice as large as for the two other firms.

This solution is bit problematic when some of the firms have not reported the latest figures. Since the calculation of $d_{B_i,t}$ requires values from all parts B_i, a missing figure makes it impossible to calculate previous years' turnover for any B_i. Solution for this problem is simple. For those months, on which there are firms with missing values, ”old” coefficients are used to divide previous year's turnover sum for enterprises, which have reported turnover. That is

$$T_{B_i,t-12} = \frac{T_{B_i,s}}{\sum_{i=1}^{n} T_{B_i,s}} \sum_{i=1}^{n} T_{A_i,t-12} = d_{B_i,s} \sum_{i=1}^{n} T_{A_i,t-12} ,$$

where $s < t$ and all $B_{i,s}$ are non-missing.

Another mentionable procedure related to changes in business activity is detection of closures. Because the VAT data accumulate for six months, missing values do not usually mean that a business has ended. Administrative data of closures is not timely enough, so if no other source of information is available, a suspect of a closure is accepted only after a certain time has passed since the firm has stopped reporting turnover.

The project made no changes in the treatment of changes in business activity. However, the processing had to be taken into account in the development of the new methods.

3.3.2 Selection of source data for individual units

Another problem related to the comparability of turnover figures from the current and previous year is caused by variance in accounting practices in the two source datum. Enterprises in the direct inquiry have turnover in the VAT data also, and sometimes the VAT turnover is not comparable to the turnover reported directly to Statistics Finland. When an inquired firm has a turnover figure for the current month, but not for the previous year, it's not necessarily wise to compare the inquired figure to the figure from the VAT data. Especially in situations involving changes in business activity with several related enterprises, this selection procedure becomes complex.

3.3.3 The imputation method for quarterly indicators

The imputation method described below was an external program, which modified the index numbers calculated from the direct inquiry and the VAT data. The procedure had two steps. First six different ARIMA-models were fitted to the individual time series of the enterprises. If an admissible model was found, it was used for the estimation of the missing value. If no model was found, turnover was estimated using the formula

$$\hat{T}_t = \frac{T_{t-1} + T_{t-2} + T_{t-3}}{T_{t-13} + T_{t-14} + T_{t-15}} \times T_{t-12} \quad ,$$

i.e. by assuming that the year-on-year change of the turnover is approximately equal to the corresponding change of the three last months. Imputation was used only in those activity divisions, where it was clearly demonstrated to improve accuracy.

The new imputation method is in some ways similar to the method described here, but the program code is entirely new and is wholly integrated to the calculation system.

3.4 Index calculation

Once the unit data is processed, the turnover figures are aggregated to activity division level. Because the data for the base year is always final, index numbers for it are calculated directly from turnover sums. Figures for the years after and before the base year are, however, calculated using change estimation.

3.4.1 The index formula

For the years after the base year the index formula for change estimation is

$$I_t = \frac{\sum_{i \in C} T_{i,t}}{\sum_{i \in C} T_{i,t-12}} I_{t-12} + s_t - e_t,$$

where $\sum_{i \in C} T_{i,t}$ and $\sum_{i \in C} T_{i,t-12}$ are turnover sums of the current month and the previous year's corresponding month over the set of continuing businesses *C*, which have reported turnover. I_{t-12} is previous year's index number and the terms s_t and e_t are the effects of start-ups and closures.

$$s_t = 100 \frac{\sum_{j \in S} T_{j,t}}{\overline{T}_b} \text{ and } e_t = 100 \frac{\sum_{i \in C} T_{i,t}}{\sum_{i \in C} T_{i,t-12}} \frac{\sum_{k \in E} T_{k,t-12}}{\overline{T}_b},$$

where $\sum_{j \in S} T_{j,t}$ and $\sum_{k \in E} T_{k,t}$ are turnover sums of the sets *S* and *E* of start-ups and closures and $\overline{T}_b$ is the average monthly turnover of the base year.

As the data accumulates, the change estimate and the terms s_t and e_t will revise.

Revision depends on the accumulated turnover. As a rule of thumb, the expected proportional revision of the change estimate is directly dependent on the percentage of turnover not yet accumulated. (See Statistics Finland's final report for the project to develop the country-stratified index of retail trade, March 2004)

3.4.2 Outlier treatment

There is currently no stratification in the inquiries, nor is the index calculation based on stratum, but there is a treatment for outliers. When the flash estimate for retail trade was developed, it was noticed that when data is poorly accumulated, big firms can influence the index too much. To correct for this, statistician can decide to decrease the weight of the firm.

The method is as follows. When the outliers for a given month have been identified, an outlier-cleaned index for the continuing businesses is calculated.

$$I_{-O,t} = \frac{\sum_{i \in C} T_{i,t} - O_t}{\sum_{i \in C} T_{i,t-12} - O_{t-12}} I_{t-12}$$

where O_t is the turnover of the outliers in month *t*. After this, it's assumed that $\left(\sum_{i \in C} T_{i,t} - O_t\right) / \left(\sum_{i \in C} T_{i,t-12} - O_{t-12}\right)$ is roughly equal to the real proportional change of turnover in the whole activity class excluding outliers. Then outlier-cleaned turnover of the continuing businesses for the month *t*-12 is calculated. This sum contains also the enterprises which have not reported their turnover for month *t*.

Total turnover without outliers for month *t* is then estimated by multiplying the *t*-12 turnover by the outlier-clean change. Finally turnover figures of the outliers are added to the outlier cleaned total turnovers and the index is recalculated using these figures. To sum up, in the presence of outliers the term

$$\frac{\sum_{i\in C} T_{i,t}}{\sum_{i\in C} T_{i,t-12}} I_{t-12}$$

in the index formula is replaced by the term

$$\frac{\frac{\sum_{i\in C} T_{i,t} - O_t}{\sum_{i\in C} T_{i,t-12} - O_{t-12}}\left(\sum_{i\in C_{total}} T_{i,t-12} - O_{t-12}\right) + O_t}{\sum_{i\in C_{total}} T_{i,t-12}} I_{t-12},$$

where the set C_{total} is the set of all continuing businesses (given the available information at the time, of course).

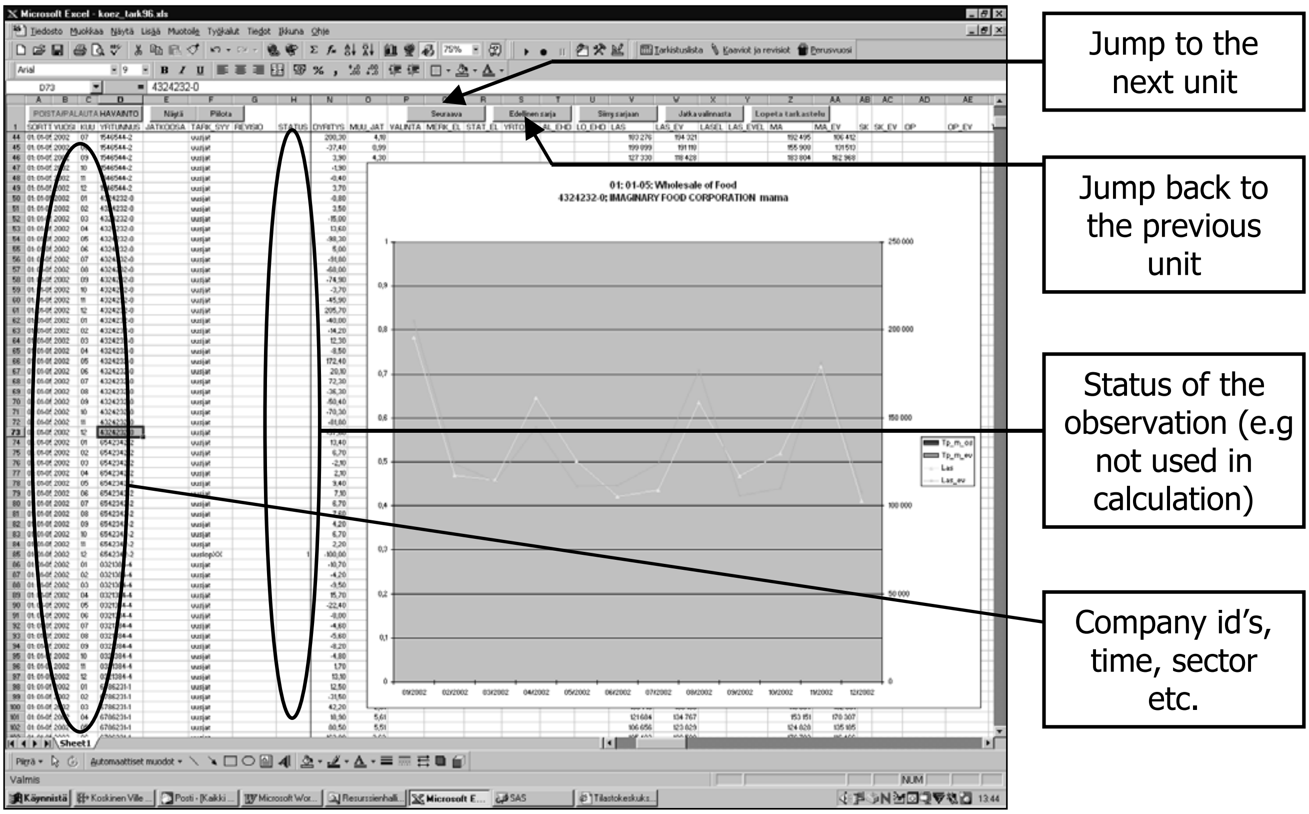
Index calculation formulas were not changed by the project. However, there have been outliers in construction and other services, so the method has been in use during the project.

3.5 Data validation

Statistics Finland has no control over the quality of the VAT data, and the number of units in the calculation is large, so the data has to be validated efficiently.

The most influential units are checked graphically. After preliminary index calculation, influence to the total change percent of the activity class is calculated for each firm. The data for the most dominant firms are then exported to an Excel-workbook which is processed with an Excel-VBA macro. A screenshot of the processed workbook is presented below (the data in the picture is not real).

Picture 2: Graphical data validation



Statistician checks each firm for possible errors, signs of split-offs, closures etc. Errors are corrected if possible. When an obvious error can not be corrected, the observation is omitted in index calculation. At this point split-offs and mergers may be stored into the database and their treatment can be checked.

Other data validation tools are revision tables, information on the accumulation of the data in terms of both numbers of units and percentage of turnover accumulated and indices of the 5-digit classes of the national specification to the NACE 2002.

Different parts of data validation cannot be thoroughly described in this report. However, it’s needless to say that it’s an imperative step in compilation of the statistics. In fact, the project team’s experience was that with the accelerated compilation schedule data validation will become much more determining than before.

3.6 Calculation of the production index for construction

The production index for construction is calculated by deflating turnover indices for building construction and civil engineering by their respective price indices. Price index for building construction is the implicit price index of the volume index of new building construction. For civil engineering, the cost index for civil engineering is used. The production index for construction in total is calculated by weighting building construction and civil engineering using the division’s value added proportions from national accounting statistics of the base year. Prior to the project the production index was already calculated from the monthly turnover index and only after this aggregated to form quarters, so the procedure was not changed by this project.

3.7 Accuracy

Because of the eventual total data, the final index numbers are practically unbiased in the statistical sense. The accuracy of the measurement of firm’s turnover in the VAT data and the treatment of structural changes, correction for erroneous observations etc. involves some error, of course. The most interesting question to the project was, however, not how well the final indices performed, but how accurate were the indices calculated solely from the direct inquiry with the T+45 timetable. That is: how much would the indices revise.

First, the project investigated accuracy of the monthly indices. Then the imputation method for quarterly indicators was analysed. Finally the inquiries were analysed considering both the effect of non-responses and possible bias caused by the selection of enterprises.

3.7.1 Revision of the monthly indices

For technical reasons, the analysis was conducted from the period of June to November 2003.

Table 1 presents the average, maximum and minimum absolute differences of the first change percentage estimates compared to the final change percentages for the turnover indices. That is: the figures are revisions of the change percentages in percentage units, had they been published at T+45.

In construction the average absolute revision of approximately two percentage units was considered satisfactory. The number was pleasantly surprising considering the accumulation of turnover. The project team was also surprised to find out that civil engineering seemed to be more prone for revision than building construction.

In other services, revisions of NACE divisions 55, 61 and 63 were satisfactory and bad only in class 64. The revision for division 64 was caused by a large enterprise that was missing from the inquiry during the period. Later, the enterprise was added to the inquiry.

Table 1: Revisions of the turnover indices 6-11/2003

Division	*Average absolute revision*	*Maximum absolute revision*	*Minimum absolute revision*	*Accumulation percentage of turnover*
Construction in total	2,1	3,4	0,7	41 %
Building construction	2,0	5,1	0,1	37 %
Civil engineering	3,9	7,6	2,1	53 %
55: Hotels and restaurants	1,6	3,0	0,2	42 %
60: Land transport; transport via pipelines	2,4	7,6	0,0	27 %
61: Water transport	1,8	3,6	0,1	85 %
62: Air transport	4,4	12,0	0,2	93 %
63: Supporting and auxiliary transport activities; activities of travel agencies	1,8	3,1	0,6	61 %
64: Post and telecommunications	12,3	17,3	1,6	89 %
72: Computer and related activities	3,1	8,0	0,4	50 %
74: Other business activities	3,0	7,0	0,3	38 %

Table 2 illustrates the revisions of the volume index for construction. Seemingly it revised less than the turnover index, but this is caused by measurement. Proportional revisions of the indices are equal, but because volume changes are generally smaller than changes in turnover, revision in percentage units are also smaller.

Table 2: Revisions of the volume index for construction 6-11/2003

Division	*Average absolute revision*	*Maximum absolute revision*	*Minimum absolute revision*
Construction in total	1,1	2,5	0,1
Building construction	1,1	2,3	0,0
Civil engineering	1,1	2,7	0,0

3.7.2 The performance of the imputation method for quarterly indicators

In construction the imputation method briefly described in section 3.3.3 was used for all the three quarterly volume indices. In other services the method was utilised in NACE classes 602-603 (land transport excluding railways) and 93 (other service activities). Imputation for rest of the activity divisions in other services did not improve accuracy sufficiently to make the method useful.

The method was evaluated by calculating how much revisions fall on average when the method is used and in how many cases the method improves accuracy. Revisions were calculated for monthly indices. The figures in table 3 summarise the situation from autumn 2003 to early 2004.

Table 3: Revisions when the imputation for the quarterly indicators is used

Division	*No. of months when imputation has improved accuracy vs. total no. of months when imputation was used*	*Average absolute revision on months when imputation was used*	*Average absolute revision on months without imputation*
Construction in total	6/7	1,8	2,5
Building construction	5/7	2,3	2,2
Civil engineering	5/7	1,6	4,3
60.2-3: Land transport excluding railways	2/3	1,6	3,0
93: Other service activities	3/4	2,6	3,3

Although imputation was utilised only in few divisions, it was clear that the accuracy was improved. Especially civil engineering benefited from imputation. The project team was very encouraged with the results of this analysis, but noted that the new method should be able to improve accuracy of all activity divisions in other services.

3.7.3 Analysis of the inquiries

Before the project, causes of revisions were not systematically charted. The project team listed following topics to be investigated:

- What is the effect of non-response? I.e. how much the revisions would decline if there was no non-response?
- Are revisions caused by the missing data as a whole or by individual firms? How do divisions differ in this sense?
- What would be the effect of always inquiring the largest enterprises (which might be reasonable considering the rule of thumb given in section 3.4.1) and leaving the smaller businesses out?

The effects of single enterprises were analysed with the tool described in section 4.1. In other services a large revision caused by a single enterprise is generally related either to non-responses or to unit's revised turnover figures. In construction, on the other hand, an increasing number of non-Finnish firms have participated and won biddings to take part in large construction projects. These firms may have large turnovers on single months, but in general cannot be contacted nor added to the inquiries. A common problem is also that some inquired firms can only estimate their turnover in the available time. The problems cannot currently be solved by Statistics Finland. Fortunately their effect has not been significant in most divisions.

The effect of non-responses is summarised in table 4. The results are trendsetting but not entirely exact because of technical reasons – for example changes in business activities could not be entirely taken account for. Again, the figures are revisions of the change percentages in percentage units. Negative difference values indicate that revisions without non-response are lower than with non-response.

Table 4: Revisions with the effects of non-response removed

Division	*Average absolute revision without the effect of non-response*	*Difference to the true average absolute revision*
Construction in total	2,79	0,69
Building construction	2,34	0,34
Civil engineering	3,77	-0,13
55: Hotels and restaurants	1,26	-0,34
60: Land transport; transport via pipelines	2,27	-0,13
61: Water transport	0,62	-1,18
62: Air transport	1,84	-2,56
63: Supporting and auxiliary transport activities; activities of travel agencies	1,11	-0,69
64: Post and telecommunications	0,77	-11,53
72: Computer and related activities	0,98	-2,12
74: Other business activities	2,86	-0,14

The analysis showed a strong dependence between response rate and revision in other services. Therefore, the number of e-mail and phone reminders was increased in other services. In construction however, no improvement could be expected. As shown in table 4, a 100% response percentage would have seemingly even impaired accuracy, but this is likely due to technicalities.

The analysis of the selection of units was based on the assumption that the better the coverage of the most significant units in the activity division would be, the better the accuracy would be. Firstly, a program for selecting significant units in was written (see section 4.1.2). Using this program, an alternative set of units was selected. Finally revisions for this set were simulated and compared to the simulated revisions of the true inquiry set.

The results of the simulations could not be compared directly to the true revisions, again because of technical reasons. Table 5 summarises the results.

Table 5: Revisions of an alternative inquiry

Division	*Simulated average absolute revisions of the alternative inquiry unit set*	*Simulated average absolute revisions of the true inquiry unit set*	*Difference*
Construction in total	2,15	2,79	0,64
Building construction	1,93	2,34	0,40
Civil engineering	2,57	3,77	1,20
55: Hotels and restaurants	1,25	1,26	0,01
60: Land transport; transport via pipelines	2,24	2,27	0,03
61: Water transport	0,43	0,62	0,19
62: Air transport	1,19	1,84	0,66
63: Supporting and auxiliary transport activities; activities of travel agencies	0,80	1,11	0,31
64: Post and telecommunications	0,85	0,77	-0,08
72: Computer and related activities	0,96	0,98	0,02
74: Other business activities	2,65	2,86	0,21

Re-selecting the units could cause a slight improvement in accuracy, especially in construction, but because adding firms to the inquiry always means taking the risk of non-response, no radical changes were made to the inquiries. Instead, it was decided to start to analyse and – if needed – update the inquiries more actively than before. Because of this, the tools for making the analysis in tables 3-5 were developed to be usable in monthly production.

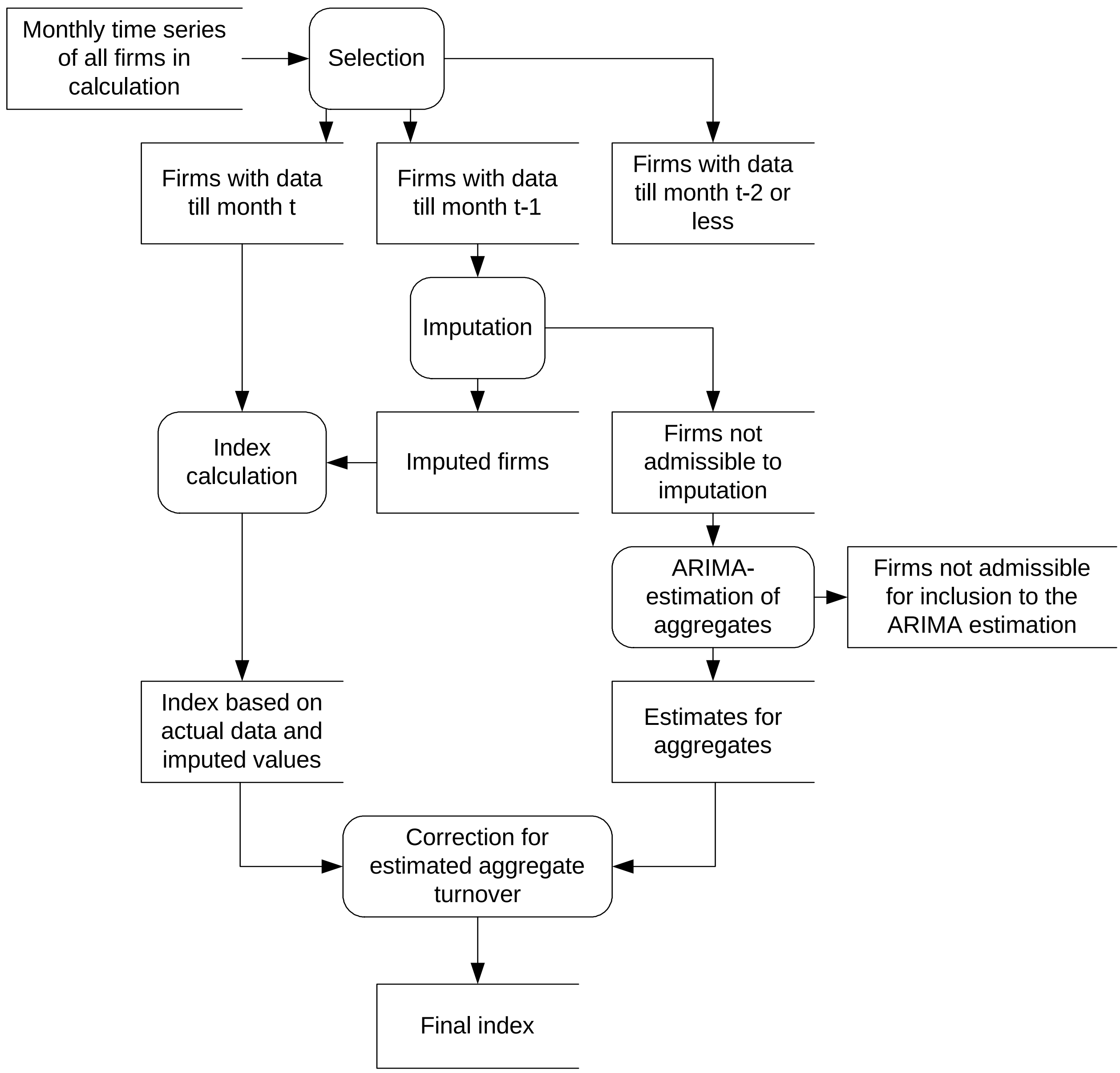
4 Improvements to the compilation methodology

In the terms of required time and resulted changes in the compilation system, the biggest developments of the project are the imputation and ARIMA-estimation methods. Some diagnostic tools were added to the compilation system as well.

A key issue for the project was the problem of how the new methods should be integrated to the compilation system. Tight integration would ensure transparency and minimise the need for adding new steps to the work process. On the other hand, integration would require more effort than programming separate programs for the methods. In the end, a technical issue dictated the solution. Imputation and ARIMA-estimation are done sequentially. This is because the model selection procedure required for ARIMA-estimation could not be computed efficiently enough to be enclosed to the massive imputation procedure. In fact, given a reasonable amount of time, model search can be only done for 20-40 time series. Therefore, only aggregated turnover series are ARIMA-modelled.

The flowchart below is an abstraction of how the imputation and ARIMA-estimation methods work together.

Picture 3: Overview of the imputation and estimation methods



This chapter is divided into three parts. Section 4.1 describes the tools developed to evaluate the selection of enterprises in the inquiries. In sections 4.2 and 4.3 methods for imputation and estimation are presented.

4.1 Tools for analysing the selection of units in the inquiries

Since there was a possibility of a need to change the inquiries, tools for analysing the selection of firms were needed. Three new tools were developed. A procedure for analysing revision effects of single enterprises and a simulation tool for making comparisons of different selections of units versus total data were developed in early spring 2004. Later in the spring of 2004 a new feature of printing out suggestions of updates to the inquiry was added to the compilation system.

4.1.1 Unit level contributions to revisions

If the number of units in calculation is large, the reason for the index revision is not always easy to infer. Obviously, there are no unambiguous causes on the unit level, but if datum from the both the first and second calculations are compared, it is possible to calculate how much a single unit contributes to the revision of the index. The principle is to calculate *ceteris paribus,* what would the index number have been, had the revised or new turnover figure been available initially. Technically this means that each firm's turnover values are sequentially replaced in the initial data by the values in the new data and the index number is recalculated and compared to the actual initial index number.

Once the contributions have been calculated, the largest are checked by a statistician. In practice, this tool has helped to identify problematic firms, i.e. firms which can not give reliable figures in the required time, firms that are not in the inquiry but perhaps should be, etc.

4.1.2 Simulation of inquiries

Because the last index number is always calculated entirely from the inquired data, an interesting question was how well different selections of units would represent the whole population. This could be assessed

1. by selecting units from the already accumulated, slightly older data,
2. by calculating indices from the selections to be compared and the population and
3. by comparing the results.

The simulation of inquiries does just this. The simulation is performed for the past half year to ensure that the change in the structural characteristics of the divisions do not affect the results. The tool relies on the processed data used for monthly index calculation. This makes the tool usable in monthly production. Because only accumulated data is used, the tool is also suitable for inferring the size of error caused by non-response.

4.1.3 Suggestions for updating the inquiries

Prior experience has demonstrated that it is reasonable to select large or in other ways influential units to the inquiries. Small units generally do not affect the index much, and because it is usually easier to collect data from a big firm than a small one, data collection is focused to the largest enterprises. The set of most influential firms varies, however. Therefore the inquiry must be constantly optimised. To help the process, a program was written to make suggestions to the statistician responsible for making the updates. The suggestions are based on the size of the enterprise and how much the firm affects the index on a typical month.

By comparing the change percentage estimates with and without the firm being in the calculation data set the program calculates an influence number, which describes the weight of the firm with respect to the change estimate. All enterprises are sorted by influence and size and then divided into two classes, inquired units and non-inquired units. From this sorted list the program makes suggestions based on the following rules:

- If a non-inquired unit is sufficiently larger and more influential than some inquired units, it is suggested to be added to the inquiry.
- Vice versa, if an inquired unit is smaller and less influential than some non-inquired units, it is suggested to be removed from the inquiry.
- If an unit meets these conditions partially, the program suggests the statistician to consider adding/removing the unit less emphatically.
- If a large and influential inquired enterprise has not responded yet, the statistician is advised to contact the enterprise promptly.

All the suggestions are made considering the smallest aggregate the unit is in. This is important for national statistics because in other services also three level NACE activity groups are published.

4.2 The imputation method

As noted above, the imputation method is an evolution to the old approach. The new method incorporates five different models and more rigorous quality checking and is also totally integrated to the system and expandable. That is: new models can be added in the future.

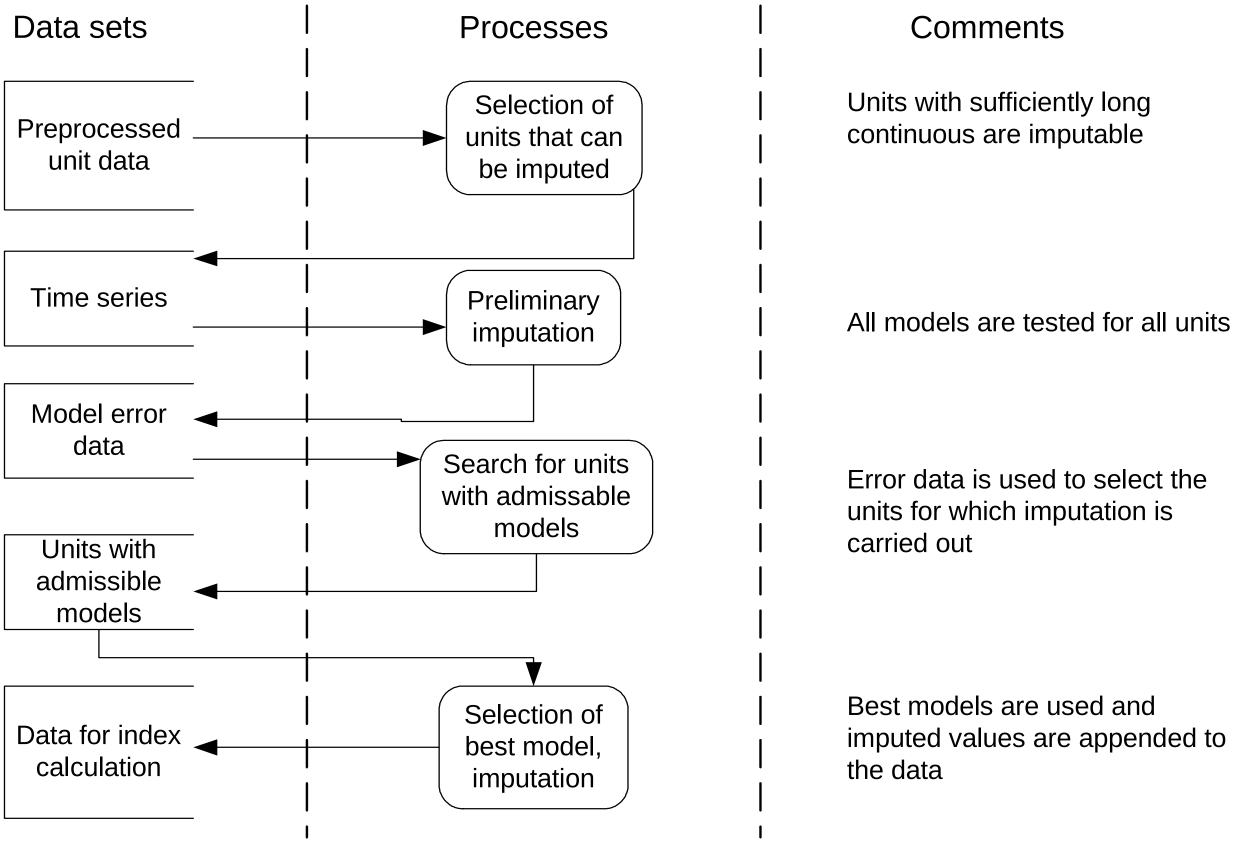
Before the procedure could be developed, unit data corrected for changes in business activity/structure and source data had to be recreated. Prior to the project, unit data was stored only in it's original, unprocessed form. Data was reprocessed from the beginning of 2000 and is now stored in a permanent data. The variables in the table are enterprise ID, possible kind-of-activity unit ID:s, time period, status (i.e. has the value been used in index calculation), time stamp (i.e. from which calculation the values originate), originating base data set, activity code at 5-digit level, possible ID number for event of change in business activity/structure relevant to the time period and turnover values for the time period and a comparable value from the previous year's corresponding month.

This data is updated after each month when the final indices have been calculated. In the imputation step this data is combined to the processed data of the seven previous month period which is being updated.

The philosophy of the method is that because there is a vast number of enterprises for which imputation can be tried, and because the results of the method cannot be checked manually, an imputed value can be accepted to index calculation only if the model is fairly reliable with historical data. Only the largest enterprises can be checked manually, in which case the statistician must be able to make corrections to the values.

Imputation is carried out just before index calculation. In the flowchart of Picture 1: The compilation process, page 9, imputation can be considered as a subprocess of the process "Data processing".

Picture 4: The imputation process



4.2.1 Methodology

Because imputation does not require much computer resources, all models can be tested for all enterprises with sufficiently long continuous time series.

Model (1): Estimate based on the year-on-year change rate of last three months summed

$$\hat{T}_t = \frac{T_{t-1} + T_{t-2} + T_{t-3}}{T_{t-13} + T_{t-14} + T_{t-15}}$$

Model (2): Estimate based on the geometric mean of the change rates of last two months

$$\hat{T}_t = T_{t-1}\sqrt{\frac{T_{t-1}}{T_{t-2}}\frac{T_{t-2}}{T_{t-3}}},$$

Model (3): Last month's turnover

$$\hat{T}_t = T_{t-1},$$

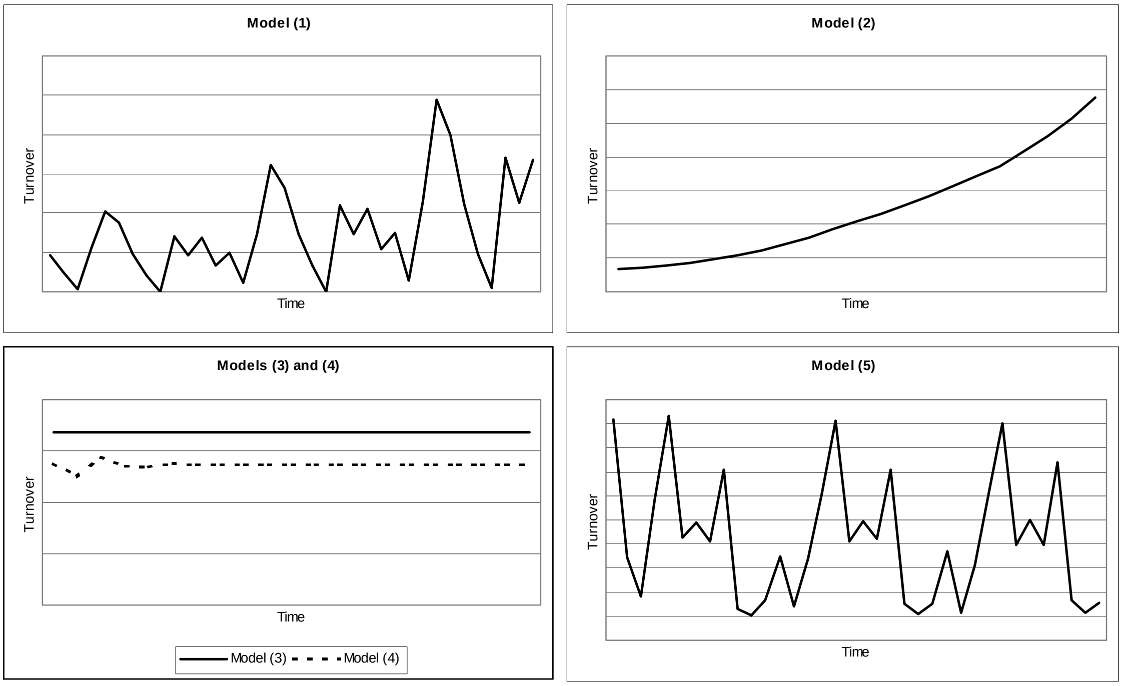
Model (4): Average turnover of the last three months

$$\hat{T}_t = avg(T_{t-1}, T_{t-2}, T_{t-3}),$$

Model (5): Previous year's corresponding month's turnover

$$\hat{T}_t = T_{t-12},$$

Picture 5: Illustration of the imputation methods



The time series presented in picture 5 are simulated.

The models are evaluated and compared by calculating maximum proportional forecast errors using data of the last six months. The selection rules are:

- A model with less than 20% maximum proportional forecast error is admissible.
- A model with greater than 20% and less than 50% maximum proportional error is non-admissible, but can be accepted in the data validation step by the statistician.
- The model with the smallest maximum error is considered best.

This means that a model which performs well on past data is assumed to work for current data. Also, model with the smallest average error is not always considered the best, because the cost of one large error is greater than the reduced profit of slightly smaller errors on several months.

The selection of the admissibility limit is of course not obvious. By allowing weak models there is a risk of reducing accuracy instead of improving it. On the other hand, too strict criteria cripples the method as the added coverage via imputation becomes insignificant to the index calculation.

The limit of 20% was chosen by simulating the proportional forecast error on the division level from September 2003 to February 2004. Table 6 recapitulates the results. It was observed that when the limit was risen from 15% to 20%, the number of units in imputation rose significantly on several activity divisions, but general accuracy remained unchanged. Limit of 20% could thus be preferred over 15%. On the other hand, rising the limit to 25% caused losses in accuracy on some divisions, so rising the limit any further was not seen as desirable.

The second limit of 50% was added to the method when it was noticed that in some cases good models were rejected because of extreme values in the period for which models were tested. Also the possibility to modify imputed values was seen necessary, so this feature was added too.

Table 6: Simulation of the admissibility limit

Division	*Admissibility limit: the maximum proportional forecast error*	*Imputed units*	*The average monthly proportional forecast error on division level*
Construction in total	10 %	85	0,67
	15 %	144	1,96
	20 %	272	2,92
	25 %	468	4,06
Building construction	10 %	67	0,59
	15 %	115	2,14
	20 %	217	3,09
	25 %	377	4,26
Civil engineering	10 %	18	1,57
	15 %	29	3,98
	20 %	55	2,05
	25 %	91	3,44
55: Hotels and restaurants	10 %	538	0,56
	15 %	1336	0,82
	20 %	2074	0,96
	25 %	2609	0,91
60: Land transport; transport via pipelines	10 %	520	2,09
	15 %	1597	1,99
	20 %	2905	2,56
	25 %	4044	2,86
61: Water transport	10 %	3	2,17
	15 %	5	6,96
	20 %	6	6,93
	25 %	11	7,02
62: Air transport	10 %	1	3,62
	15 %	1	3,62
	20 %	2	9,22
	25 %	2	9,22
63: Supporting and auxiliary transport activities; activities of travel agencies	10 %	37	5,22
	15 %	70	4,12
	20 %	127	3,66
	25 %	172	3,46
64: Post and telecommunications	10 %	13	2,13
	15 %	38	2,94
	20 %	60	4,17
	25 %	92	3,21
72: Computer and related activities	10 %	33	3,44
	15 %	59	4,11
	20 %	115	2,96
	25 %	183	3,81
74: Other business activities	10 %	341	0,83
	15 %	665	1,08
	20 %	1156	1,79
	25 %	1730	2,64

The project team was somewhat surprised to find out that the proportion of units for which a value is imputed is rather low. Situation in the end of the project is illustrated in table 7.

Table 7: Number of units in imputation in the end of the project

Division	*Approximate number of enterprises for which imputation is attempted*	*Approximate proportion of enterprises with admissible model*
Construction in total	28000	1 %
Building construction	22000	1 %
Civil engineering	6000	1 %
55: Hotels and restaurants	10000	20 %
60: Land transport; transport via pipelines	19000	15 %
61: Water transport	200	9 %
62: Air transport	100	8 %
63: Supporting and auxiliary transport activities; activities of travel agencies	2000	8 %
64: Post and telecommunications	500	13 %
72: Computer and related activities	4000	2 %
74: Other business activities	28000	4 %

Despite the low number of enterprises, simulations and practice indicate that the method works well in many cases. This is partly because what is relevant to the accuracy of the index calculation is not only the accuracy of imputation, but also how imputation works on firms with atypical growth rates – i.e. firms which should be inquired because they are growing or losing market share fast, but which for practical reasons have to be left not inquired. The imputation models seem to work well especially in situations when the turnover of a firm has a trend-like development. This is not surprising, considering models (1) and (2), which are suitable for estimating linear or exponential growth.

Finally, in table 8 is listed the distribution of models used in imputation at the end of the project.

Table 8: Distribution of the use of imputation models

Model	*Cases when used in imputation, percentages*
Model (1)	3 %
Model (2)	19 %
Model (3)	33 %
Model (4)	29 %
Model (5)	16 %

Surprisingly, model (1), which was the cornerstone of the earlier imputation method for quarterly indicators, is the least used model.

4.2.2 Data validation

Data validation of imputed values is similar compared to the figures from the VAT data and the direct inquiry. Values for the most influential units are checked manually. In practice there are some differences. The rows for imputed values on the data validation list are coloured red. It is also possible to make corrections to the values on the list.

Before accepting an imputed value to the calculation the statistician assesses how plausible the value is. Things to take into consideration are:

- Is the imputed value realistic considering the seasonal variation of the division?
- How long continuous increase or decrease in turnover is plausible?
- Has imputation been reliable for the particular enterprise in the past?
- Are there changes in business activity which could interfere imputation?

- Are there other sources, e.g. news or information about the wages than give extra information of the imputed month?
- If it's known that the firm's turnover's year-on-year change is rapid compared to the activity division, even a less accurate estimate of the missing turnover figure may correct the index to the right direction.

After each monthly calculation round, the statistician checks the biggest revisions of imputed values both on the unit level and on the division level. These listings, containing information also on selected models etc. are produced automatically.

4.2.3 Performance

Performance is measured by how much the revision of the change percentage decreases when the imputation method is used compared to the case with no imputation. This comparison is presented in the table below. In other services the figures are based on six months of data and in construction on four months of data. To ensure comparability between construction and other services, all figures are calculated from the turnover indices.

Table 9: Performance of the imputation method

Division	*Average number of enterprises with imputed values*	*Proportion of cases when imputation has improved accuracy*	*Average effect to revision of the change percentage estimate (in percentage units) **	*Best case*	*Worst case*
Construction in total	244	25 %	0,09	-0,32	0,34
Building construction	175	75 %	0,18	-0,37	1,72
Civil engineering	69	50 %	0,44	-0,51	2,05
55: Hotels and restaurants	1525	33 %	0,04	-0,75	1,21
60: Land transport; transport via pipelines	2255	67 %	-0,35	-1,31	0,23
61: Water transport	14	83 %	-0,10	-2,07	0,19
62: Air transport	5	67 %	-1,04	-2,75	0,30
63: Supporting and auxiliary transport activities; activities of travel agencies	112	67 %	-0,96	-2,46	0,59
64: Post and telecommunications	50	83 %	-0,66	-1,84	0,02
72: Computer and related activities	76	83 %	-1,15	-2,67	0,38
74: Other business activities	838	83 %	-1,37	-4,07	0,08

* if revision effect is below zero revision has decreased, if above, increased.

Especially in other services, the method works well. Imputation improves accuracy in all activity classes except 55: Hotels and restaurants. Divisions 62, 63, 64, 72 and 74 benefit most significantly. What is also pleasing is that in other services the worst case revision effects are relatively close to zero, whereas the best case effects are generally above two percentage units.

The method failed on one month both in building construction and civil engineering. In these cases the revision effect was close to two. When the unit level revision of the imputed values was investigated, it turned out that the method had worked as well as on other months, but because the imputed units had so different growth rates compared to those not imputed, index was corrected to the wrong direction. In construction the method thus seems to be somewhat biased towards modelling untypical development. This is taken into account in the current practice of validating the imputation results.

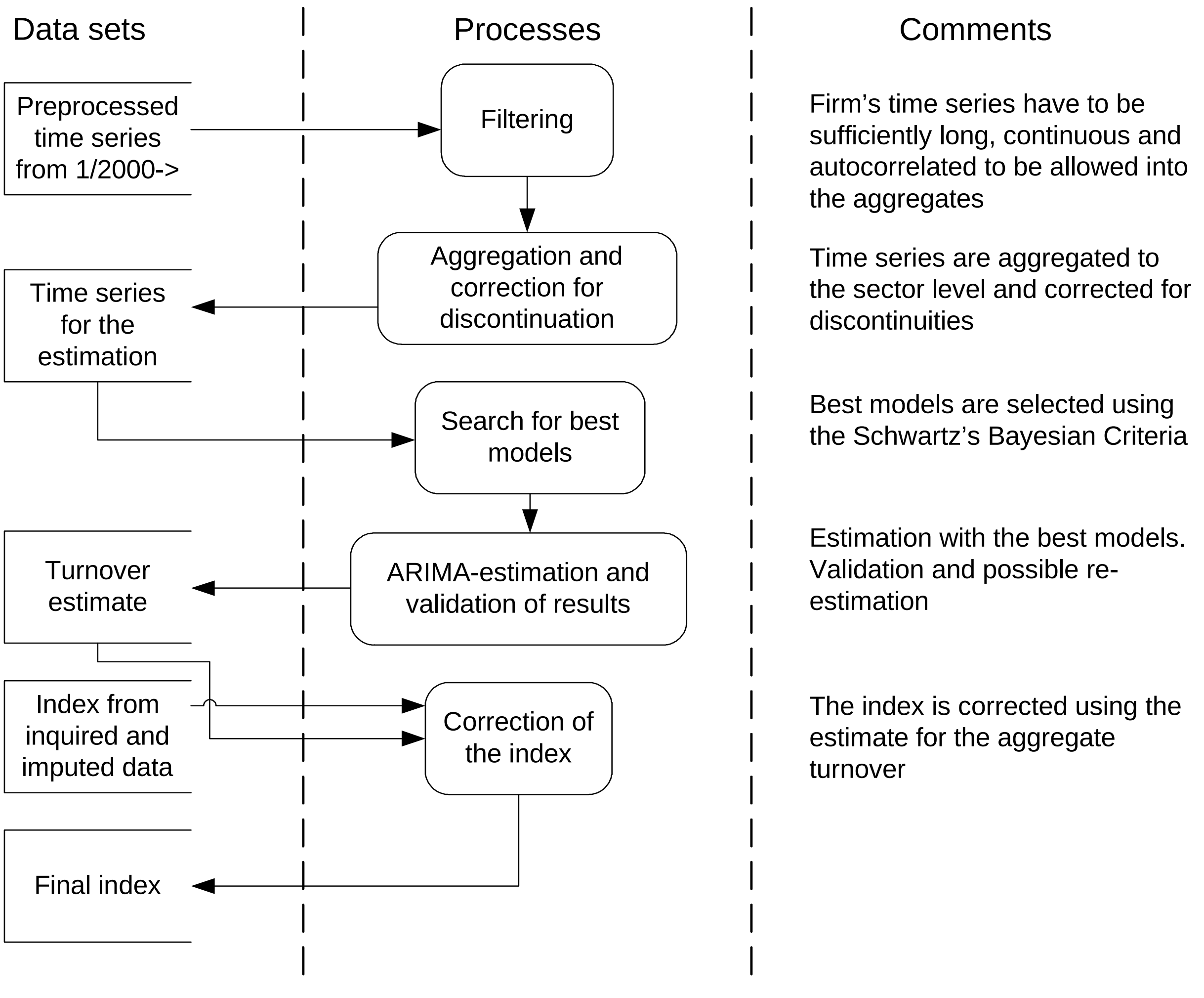
4.3 The ARIMA-estimation method

Unlike imputation, ARIMA-estimation cannot currently be performed efficiently on the unit level. It was also noticed that with the available data, estimation of aggregates was usually more accurate than combined unit level estimates. The reason for this has not been thoroughly investigated, but a possible explanation is that because of competition, partial mergers etc, there are crosscorrelations between the firm's turnover series, and that these effects make estimation difficult on unit level, but are cancelled in aggregation, so that they do not affect estimation of aggregates.

It was decided that ARIMA-estimation is performed on the activity division level. Because of this and the method's computational heaviness, it wasn't practical to integrate the estimation to the compilation system as totally as the imputation method.

The estimation procedure is executed after the indices based on the inquired and imputed data are ready. In the flowchart of Picture 1: The compilation process, page 9, ARIMA-estimation would lie after the process "Data validation". The picture below describes the estimation procedure generally.

Picture 6: The ARIMA-estimation process



4.3.1 Preprocessing

The source data for estimation is the same data which is used in imputation, but with some preprocessing. Because the aggregate time series should be as well-behaving as possible to make the estimation accurate, firms with undesirable stochastic properties are not allowed into the aggregates. As an abstraction, if

realisation of a stochastic process = stochastic effects + random effects, filtering is done to minimise the random effects by disallowing firms, which represent mostly randomness in their time series. To get through filtering, a firm must fulfil the conditions listed below.

- The firm must have reported a turnover figure for the month T-1.
- The turnover time series must consist of at least 24 observations.
- All the observations must be actual, not imputed.
- The time series must contain statistically significant autocorrelation with lags 1,2,3,6 or 12
- Turnover is not constant (variance is great enough).

These are all *ad hoc* rules. The justification for them is empirically observed positive effect to the accuracy of the estimation.

Estimation is not always done on the same aggregate level as the published indices. Some aggregates have subaggregates, which have entirely different business cycles and seasonal variation. (For example NACE division 63 contains classes like cargo handling and travel agencies.) The aggregation itself is straightforward. However, the data contains firms with different lengths of time series, and some time series are cut because of changes in business activity. Thus the directly calculated turnover sums contain level shifts.

To make the time series more suitable for ARIMA-modelling, the turnover sums of months *T*-1 to *T*-12 are chained backwards using year-on-year change percentages calculated for each month. To be more specific, the aggregate turnover series $X_{A_i,t}$ of aggregate $A_{i,t}$ used for estimation is formed by letting

$$X_{A_{i,t},t} = \sum_{j \in A_{i,t}} T_{j,t} \quad \text{for all } t \geq T-12$$

and

$$X_{A_{i,t},t} = X_{A_{i,t},t+12} \frac{\sum_{j \in A_{i,t+12}} T_{j,t}}{\sum_{j \in A_{i,t+12}} T_{j,t+12}} \quad \text{for all } t < T-12 .$$

The set $A_{i,t+12}$ in the latter equation contains firms with turnover data on the month *t*+12. That is, if a firm changes activity class after month *t* but before *t*+12, it is in the set $A_{i,t+12}$ but not in $A_{i,t}$.

Table 10: Illustration of correction for level shifts

	Firm A	*Firm B*	***Direct sum***	***Corrected sum***	
January 2005	100	500	600	600	
January 2004	200	400	600	600	
January 2003	300		300	**900**	=600*300/200
January 2002	350		350	**1050**	=900*350/300

The general idea is presented in the above table. Consider two firms, A and B. Assume B has changed activity class in the beginning of 2004. The change has been technical, so the level shift in the direct sums is not real. By correcting the time series for level shift as shown above, the time series become much more well-behaved and thus estimable.

4.3.2 Model search and estimation

The ARIMA-model used in estimation is determined by selecting the best model from a set of 288 parameterisations. They are formed by all combinations of ARIMA-models

- with or without log-transformation,
- with or without first differences,
- with or without yearly differences,
- with possible AR-polynomials of lags 1 to 3, 12, 13 and
- with possible MA-polynomials of lags 0 to 2, 12, 13

Best model is selected by minimising the Schwarz's Bayesian Information Criteria and by testing whiteness of the residuals. The criteria favours high likelihood (good fit) and low number of free parameters. It has been shown that the criteria asymptotically consistent, i.e. it selects the correct model almost certainly when the number of observations approaches infinity. Residuals are tested using the Ljung-Box test for whiteness of residuals at 10% significance level.

Model selection and estimation is entirely automatic. No user intervention is needed before data validation.

Currently no other explanatory factors than the time series' own history can be used in modelling. Dynamic regression models (or ARIMAX-models) were tested, but because they offered no evident improvement in accuracy, the project team decided to stick with the simpler ARIMA-models and to leave the development of more refined solutions for future projects.

4.3.3 Data validation

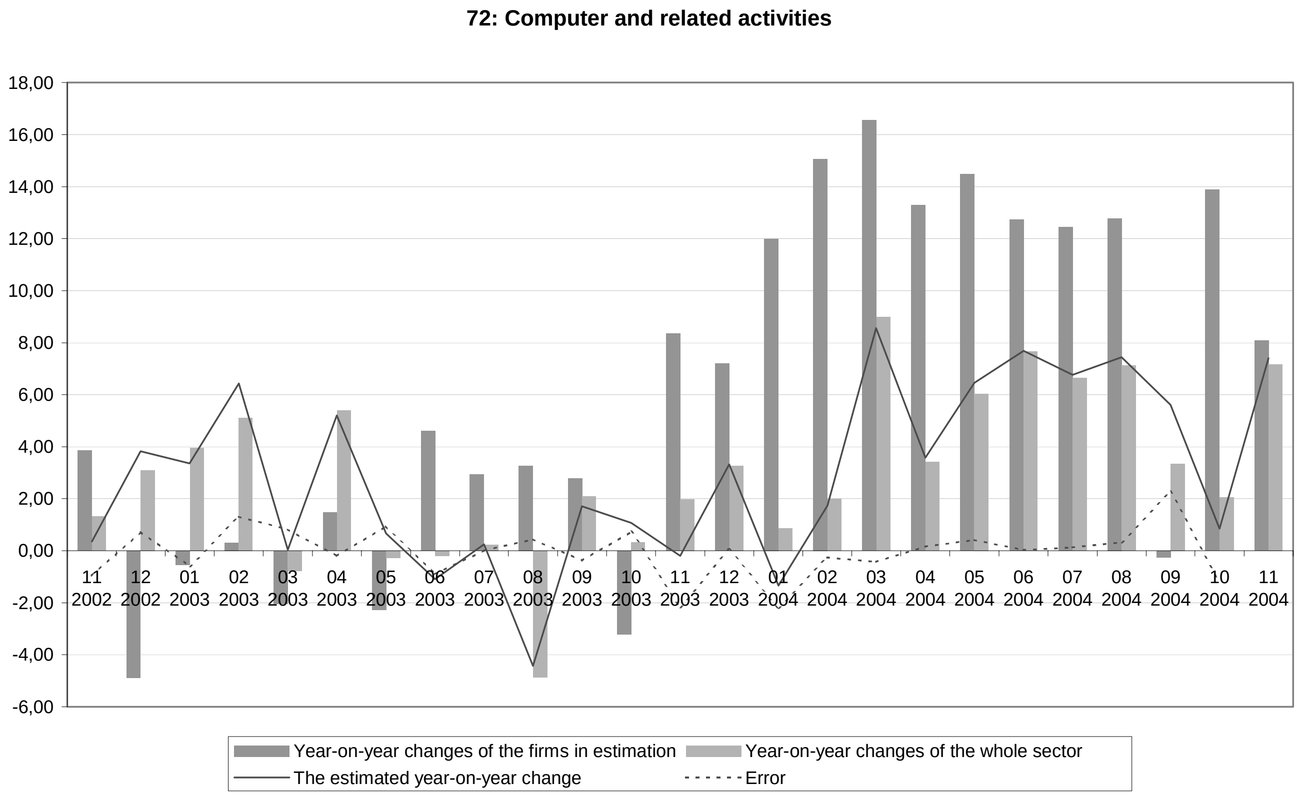
Results of ARIMA-estimation are checked using a similar solution as in unit level data validation. Statistician checks the plausibility of the estimates visually and using information of the largest enterprises in estimation. The guidelines for checking the results are:

- The estimated change percentage should be credible considering the general trend of the division. If the estimated change percentage differs considerably from the change percentage of the inquired and imputed enterprises, there should be a reason for that. For example, do the smaller not inquired units develop differently compared to the large enterprises or do some influential inquired units behave more volatility than the division in general?
- The change percentage should be credible considering the prior development of the largest enterprises in estimation. If it's known that changes in business activity have happened lately among the largest firms, the related firms may have to be removed from estimation. If it's known that there are some other exogenous events like strikes that cannot be modelled, the estimation should not be allowed to affect the results at all.
- Residuals are checked by investigating how much error the model would have had if used on prior months.
- Estimates are used only if they improve accuracy. If estimation increases revisions in two consecutive calculation rounds, only inquired and imputed data are used until the method has improved accuracy in two consecutive calculations.

The graphics are illustrated below. The picture contains the most important information from the check-list:

- Year-on-year changes of the firms in estimation
- Year-on-year changes of the whole division
- The best model's estimate of the whole division's change percentage if the firms in estimation would have been estimated with the model
- The error of the model

Picture 7 : Graphical data validation of ARIMA-estimation



The year-on-year changes of the firms in estimation can be compared to the whole division's changes to see if there is systematic difference in the development of the non-inquired firms compared to the inquired. The model's estimates and the errors are of course theoretical, but can give a hint of the model's performance.

In the case illustrated in the picture the non-inquired firms have since late 2003 systematically grown faster than the inquired firms. In this case estimation is likely to improve accuracy.

However in some cases the model can fail. ARIMA-models have been proved to be efficient in forecasting – or in this case nowcasting – various economic time series but they cannot take any exogenous or unique effects into account. Because of the naivety of the models, a statistician must evaluate the results based on his or her own knowledge of the industry and it's latest development.

4.3.4 Performance

Again performance is measured by the effect to revision. Table 11 summarises the results of estimations for February to October 2004.

The method has been very successful in construction, where the revision of the change estimate has decreased by almost two percentage units. This means that the revision, generally around 3-4 percentage units, has been halved. In other services the effect is not as significant, but the results are still generally positive.

Table 11: Performance of the ARIMA-estimation method

Division	*Effect to average revision of the change percentage estimate (in percentage units) **	*Estimation improves accuracy, proportion of cases*
Construction in total	-1,92	78 %
Building construction	-1,94	89 %
Civil engineering	-1,89	89 %
55: Hotels and restaurants	-0,31	67 %
60: Land transport; transport via pipelines	-1,07	67 %
61: Water transport	0,04	44 %
62: Air transport	0,35	44 %
63: Supporting and auxiliary transport activities; activities of travel agencies	-0,43	78 %
64: Post and telecommunications	0,37**	33 %
72: Computer and related activities	-1,12	78 %
74: Other business activities	-0,23	56 %

* If revision effect is below zero revision has decreased, if above, increased.
** The effect for class 64 is calculated from June to October. The estimation method would not have worked prior to this period of time because of a significant non-inquired unit.

The estimation method was in development during the time interval covered by the analysis. The method and working practices have been improved during the time, so it can be expected that the method can be utilised even more efficiently in the future than during the project. Having said that, it's possible that new challenges are encountered, so the method may need development and updating after the project.

Based on the experiences of the performance of the methods during the project, the expected effect of the method is as in the following table. The last column on the right is a calculational expected average revision – the effect of start-ups and closures has been estimated based on historic data and added to the expected revision for continuing businesses.

Table 12: Expected effect of the imputation and ARIMA-estimation methods

Division	*Average effect of the imputation method (in percentage units)*	*Average effect of the ARIMA-estimation method (in percentage units)*	*The expected average absolute revision of the change estimate for continuing businesses*	*Expected average revision (with start-ups and closures taken into account)*
Construction in total	none	-1,9	1,2	1,5
Building construction	none	-1,9	1,3	1,6
Civil engineering	none	-1,9	2,0	2,5
55: Hotels and restaurants	none	-0,3	1,2	1,6
60: Land transport; transport via pipelines	none	-1,1	0,8	1,0
61: Water transport	none	none	1,3	1,5
62: Air transport	-1,0	0,4	2,9	3,5
63: Supporting and auxiliary transport activities; activities of travel agencies	-1,0	-0,4	1,7	2,0
64: Post and telecommunications	-0,7	0,4	1,2	1,5*
72: Computer and related activities	-1,2	-1,1	1,8	2,2*
74: Other business activities	-1,4	none	1,6	2,1*

* The figure for divisions 64, 72 and 74 are ad hoc estimates because valid data for the revision caused by start-ups and closures was not available for these divisions.

These figures may be somewhat biased upwards, because the historic revisions caused by startups and closures have been directly added to the expected revision of continuing businesses. In fact, cancelling of different effects will occur.

5 Conclusions

In the grant agreement of the project (no 2003 44401014) the expected results of the action were described as follows:

The turnover indicator for other services will be transmitted to Eurostat within 60 days (instead of the current delay of 75 days). The production indicators for construction will be transmitted to Eurostat monthly within 60 days (instead of current quarterly indicator with 50 days delay). Both these changes are proposed in the STS Amendment regulation. The production system has been implemented and the necessary documentation and training of personnel has been carried out.

Results of this project may be applicable to more detailed activity levels than NACE 2-digit in services. Due to the relative good timeliness of the direct data collection, it might be possible that the system will allow Statistics Finland to produce the index series in faster timetable than the required 60 days. The results of this project may result in improvement in the estimation system for quarterly indicators. This would have a positive effect to the quality of quaterly national account figures.

The expected results have been achieved:

- The monthly turnover indicator for other services at 2-digit level and the monthly volume indicator for construction have been transmitted to Eurostat with 45 to 52 days delay instead of the expected 60 days delay.
- The production system has been finalised.
- The developed methods are documented, the personnel has trained the use of the new methods during the project and working guidelines and the modifications to the compilation system are documented in Finnish.
- The detailed activity levels (i.e. below NACE 2-digits) of other services are calculated but not yet transmitted, because their quality has to be assessed in long-term evaluation.
- Modification to the calculation of quarterly indicators for national accounts has been proposed by the project.

The developed methods improve accuracy of the indices significantly.

The project was one of Statistic Finland’s first efforts in increasing the use of statistical estimation and imputation methods in the field of short term economic statistics. It has benefited Statistic Finland’s knowledge base in time series modelling and in handling of massive time series datum. It has shown that even relatively simple models can be successfully utilised to refine the quality of statistics compiled from incomplete data.

The project has also laid a footing for further development of using estimation and imputation methods in conjunction with the VAT data. The imputation method is expandable, so new models can be added to it in the future. The estimation method can be improved by adding new regressors to the models.

The timely compilation schedule requires more effort in following the different auxiliary information available of the construction sector and other services. This will improve knowledge of business cycles and the structure of the divisions. In the end this will improve the quality of other statistics as well.

Statistics Finland considers the project finalised, since all the objectives are met and all the tasks have been carried out.