WP.11
ENGLISH ONLY

**UNITED NATIONS STATISTICAL COMMISSION and**
**ECONOMIC COMMISSION FOR EUROPE**

**CONFERENCE OF EUROPEAN STATISTICIANS**

**Work Session on Statistical Data Editing**
(Ljubljana, Slovenia, 9-11 May 2011)

Topic (ii): Editing of electronic collections

# Editing of Mixed Source Data for Turnover Statistics

**Invited Paper**

Prepared by Jeffrey Hoogland, Statistics Netherlands

**Abstract:** At Statistics Netherlands a new statistical process for quarterly and yearly turnover statistics is implemented. This process is based on mixed source turnover data. VAT data are used, instead of survey data, for all enterprises except those belonging to one of the 1900 largest large groups of enterprises or branches where VAT data is useless. The paper focuses on automatic editing of VAT data and a top-down editing strategy for mixed source turnover data.

## I. INTRODUCTION

1. The Dutch government and the business community in the Netherlands demands that Statistics Netherlands reduces the response burden for companies by avoiding inquiring data that could be collected from Administrative registers. Under the project name DRT we therefore developed a statistical process for turnover statistics which uses value-added tax (VAT) data for most enterprises. Questionnaires are used for large groups of enterprises and branches where VAT data are useless. Selective macro-editing is used to improve the efficiency of the estimation proces and the plausibility of the estimates. Last, the quality of turnover statistics is improved by making the short-term statistics (STS) turnover consistent with that of the structural business statistics (SBS) on a yearly basis.

2. In the new process, all enterprises belonging to the 1900 largest groups of enterprises are surveyed. The 350 largest groups of enterprises are profiled, observed and edited by a separate department at Statistics Netherlands for all quarterly and yearly business statistics. The remaining 1550 large groups of enterprises are edited within the DRT system for quarterly and yearly turnover statistics. For the next few years monthly turnover statistics are still produced by the statistical process IMPECT2[1] which is in use since 2002 (De Jong, 2003).

3. In this paper we focus on editing of mixed source turnover data for Dutch enterprises besides the 350 largest groups of enterprises. In section II, we discuss the use of VAT data for turnover statistics. Section III concerns two situations were automatic editing of VAT data is necessary. Section IV deals with the statistical process for the macro-editing phase of DRT. In section V, we discuss macro-editing

[1] The IMPECT2 system is designed for short-term enterprise-based statistics and focuses on turnover. Depending on industry and size-class, one or more variables are collected on a monthly or quarterly basis. For most industries, if possible, VAT-data is used for small enterprises (less than 10 persons employed) instead of questionnaires.

methods that are used for DRT. Interactive editing of VAT data and STS questionnaires is discussed in section VI.

## II. TURNOVER STATISTICS BASED ON VAT-DATA

4. VAT data can be very useful to produce statistics about turnover. However, the use of Dutch tax data is not a straightforward process. VAT data are only available for VAT units and several matching procedures are necessary to obtain VAT data for enterprises. Furthermore, VAT declarations have to be edited quite often, because of measurement, population frame, and linking errors. VAT declarations are submitted on a monthly, quarterly, or yearly basis. Especially large enterprises declare every month. VAT turnover for a specific VAT unit and time period is derived from variables in the VAT declaration.

5. For STS, we are mainly interested in the yearly development of turnover in a specific period, e.g. a month or a quarter. For SBS, we are interested in yearly development and level of yearly turnover. Yearly turnover that is estimated on the basis of the new STS process is leading for SBS. SBS turnover is the sum of quarterly STS turnover within that year. Therefore, we check the turnover level per quarter. In a final release for STS the yearly declarations are used as well and the sum of the quarterly turnover in a reporting year is made consistent with the yearly turnover. In exceptional cases a new quality version of the final STS release is made if the SBS department concludes that STS turnover is totally wrong.

6. VAT turnover for an enterprise can not be used if one of the related VAT units is also related to another enterprise. Since respondent burden can be reduced if VAT data can be used more often Statistics Netherlands (SN) changed the composition of enterprises (Beuken and Vlag, 2010). If a VAT unit is related to several enterprises they are combined. A disadvantage of this approach is that enterprises may become more diverse in their economical activities.

7. For some branches it is expected that VAT data can not be used, because of VAT regulations. Therefore, the usability of VAT data is assessed for each NACE code. For past reporting years, the yearly VAT turnover of an enterprise is compared with the yearly turnover observed by SBS. A decision tree is used to decide whether VAT turnover can be used for a specific branch. For instance, VAT turnover is not used when a regression analysis shows that there is no linear relationship between yearly VAT turnover and SBS turnover. A list of VAT regulations is also used to decide on the usability of VAT data. For example, mortuaries have a VAT dispensation. However, some mortuaries do declare VAT, because of sidelines such as photo reports. There is no linear relationship between sidelines of mortuaries and principal turnover. VAT turnover of mortuaries is therefore not used to estimate turnover of mortuaries.

## III. AUTOMATIC EDITING OF VAT DATA

8. Several suspicous patterns are found in monthly and quarterly VAT turnover data. For example, the turnover pattern of a VAT unit for four subsequent quarters in a year may follow a $(0, 0, 0, x;\ x>0)$ pattern, i.e. zero turnover in the first three quarters, positive turnover in the last quarter. This pattern often means that the VAT unit declares the yearly turnover in the last quarter of the year, especially if the same pattern arises each year. The final release for quarterly turnover for a specific year is made after the end of the year, i.e. when almost all declarations are received. For the final release the turnover value $x$ of such a unit is therefore used as its yearly turnover and its turnover per quarter is imputed.

9. Table 1 summarizes the quarterly turnover patterns that are most frequently found in VAT data. This table also includes the percentage of units per pattern and their share in the yearly turnover of quarterly declarations. Table 1 shows that about 13% of VAT units that declare per quarter show a suspicious turnover pattern. Especially patterns with zero, or the same turnover values occur frequently. For patterns 1 till 8 a correction rule is applied, only if it occurs two years in a row for a VAT unit. For most patterns the correction rule implies that only the yearly turnover is used. For patterns 5-9 a turnover

correction depends on the NACE code of a VAT unit. For instance, the same turnover in each quarter is plausible in case of rent of real estate. The NACE code is not available in VAT data and corrections therefore take place after VAT units are matched to the population frame. The NACE code of a VAT unit is approximated by the NACE code of an enterprise that is linked to the VAT unit.

*Table 1. Impact of VAT units with a suspicious quarterly turnover pattern. The percentage of these units among all units that report per quarter is shown. Also the share of these units in the total yearly turnover of units that report per quarter is given.*

| | Year | 2007 | 2008 | 2007 | 2008 |
|---|---|---|---|---|---|
| No | Pattern | Percentage of units | | Share in yearly turnover | |
| 1 | (0,0,0,x) , x > 0 | 1,59% | 1,53% | 0,47% | 0,40% |
| 2 | (0,0,x,0) , x > 0 | 0,42% | 0,45% | 0,12% | 0,14% |
| 3 | (0,x,0,0) , x > 0 | 0,52% | 0,52% | 0,14% | 0,10% |
| 4 | (x,0,0,0) , x > 0 | 0,99% | 0,97% | 0,24% | 0,15% |
| 5 | (w,x,y,0), w,x,y>0 | 1,27% | 1,27% | 0,78% | 0,63% |
| 6 | (w,x,0,z), w,x,z>0 | 1,05% | 1,04% | 0,85% | 0,43% |
| 7 | (w,0,y,z), w,y,z>0 | 0,81% | 0,79% | 0,66% | 0,62% |
| 8 | (0,x,y,z), x,y,z>0 | 1,99% | 1,99% | 2,38% | 1,52% |
| 9 | (x,x,x,x), x ≠ 0 | 2,46% | 1,89% | 0,71% | 0,57% |
| 10 | (x,x,x,y), x ≠ 0 and y ≠ 0 | 0,98% | 0,73% | 0,37% | 0,28% |
| 11 | (x,x,y,x), x ≠ 0 and y ≠ 0 | 0,16% | 0,12% | 0,06% | 0,05% |
| 12 | (x,y,x,x), x ≠ 0 and y ≠ 0 | 0,17% | 0,13% | 0,06% | 0,05% |
| 13 | (y,x,x,x), x ≠ 0 and y ≠ 0 | 0,51% | 0,42% | 0,16% | 0,14% |
| 14 | (w,x,y,z), z<0 | 0,17% | 0,17% | 0,14% | 0,17% |
| 15 | (w,x,y,z), y<0 | 0,08% | 0,08% | 0,06% | 0,05% |
| 16 | (w,x,y,z), x<0 | 0,05% | 0,06% | 0,05% | 0,04% |
| 17 | (w,x,y,z), w<0 | 0,04% | 0,04% | 0,04% | 0,10% |
| | Other | 86,74% | 87,81% | 92,73% | 94,56% |

10. Monthly turnover of VAT units can refer to a 4-, 5-, or 8-weeks period. This occurs for enterprises with a 4-weekly or weekly administration, such as supermarkets. The Tax Authority does not offer the possibility to declare VAT on a 4-weekly basis and VAT units are therefore inclined to declare VAT turnover for a 4-, 5-, 8-, or 9-weeks period instead of a month. For instance, supermarkets frequently declare turnover for eight or nine weeks for November or December, and turnover for four or five weeks for other months. This also distorts the turnover per quarter. An automatic editing method was therefore developed to obtain a correct turnover per month.

11. For detection of a so called *month peak* due to a 4-weeks administration we can use a dummy variabele $D_j^t$ for any VAT unit *j* which is reporting turnover in month *t* (*t* =11 of 12)

$$D_j^t = \begin{cases} 1 & \text{if } O_j^t > 0 \text{ and } O_j^{t-1} > 0 \text{ and } O_j^t > \alpha^t O_j^{t-1} \text{ and } O_j^t < \beta^t O_j^{t-1} \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

where

$O_j^{t-1}$, $O_j^t$ : net turnover of VAT unit *j* in month *t-1* and *t*, respectively

$\alpha^t$ , $\beta^t$ : parameters that determine how much $O_j^{t-1}$ and $O_j^t$ may differ from each other.

If $O_j^t$ is bigger than $\alpha^t O_j^{t-1}$ we assume that $O_j^t$ refers to a period of at least 8 weeks. However, if $O_j^t$ is bigger than $\beta^t O_j^{t-1}$ we assume that there is some other kind of error in the turnover. Note that $1 < \alpha^t < \beta^t$ .

12. For the development of detection rules for month peaks we used a dataset containing both VAT data on a monthly basis and survey data on a 4-week basis for about 60 franchisees of a large chain of

supermarkets. About half of the franchisees reported a 8-week turnover for November or December. An example is given in Table 2. In 2009 the thirteen '4'-weekperiods covered 53 weeks. The VAT unit reported an 8-week turnover for November and a 5-week turnover for December. The turnover in November is therefore 2,06 times larger than in October. We optimized the parameters for supermarkets such that almost all month peaks were detected without a false alarm. This resulted in $\alpha^{11}$ =1,8 and $\beta^{11}$ =3,6; $\alpha^{12}$ =2,0 and $\beta^{12}$ =4,0.

*Table 2. Reported monthly turnover of a VAT-unit in the period August till December 2009.*

| Month | VAT-turnover | '4'-week period | Number of weeks |
|---|---|---|---|
| August | 544 | 13 Jul – 9 Aug | 4 |
| September | 539 | 10 Aug – 6 Sep | 4 |
| October | 544 | 7 Sep – 4 Oct | 4 |
| November | 1118 | 5 Oct – 1 Nov | 4 |
| | | 2 Nov – 29 Nov | 4 |
| December | 682 | 30 Nov – 3 Jan | 5 |

13. For correction of monthly and quarterly turnover due to a month peak for a VAT unit *j* in month *t* (*t* =11 of 12) we assign a weight to each sale day in the week. For supermarkets these weights are given in Table 3. These weights are used to map a 4-, 5-, 8-, or 9-weeks turnover on a monthly turnover. For details we refer to Van Delden et al (2010). Notice that the weights shown in Table 3 need an update, because visiting a supermarkt on Sunday has become more common in the Netherlands.

*Table 3. Sale day pattern of Dutch supermarkets as percentage of the weekly turnover ( May 2006).*

| Mon | Tue | Wed | Thu | Fri | Sat | Sun | week |
|---|---|---|---|---|---|---|---|
| 8,2 | 11,7 | 12,0 | 19,4 | 25,3 | 22,0 | 1,5 | 100,0 |

## IV. STATISTICAL PROCESS FOR MACRO-EDITING PHASE

14. Each time a part of the statistical process of the DRT system is completed, the resulting data are saved in one or more databases. In this way, processed data are available for users within Statistics Netherlands. We have several so called *rest stages*, such as

- rest stage 3: Improved input data matched to population frame
- rest stage 4A: Authorized publication figures
- rest stage 4B: Edited enterprise data related to publication figures
- rest stage 4C: Edited input data related to enterprises

For each reporting period and release, the macro-editing phase starts with data from rest stage 3. This phase results in edited data in rest stage 4A, 4B, and 4C. In the following paragraphs each part of the macro-editing phase is briefly described.

15. The statistical processes which are executed during the macro-editing phase are depicted in Figure 1. Macro-editing is, basically, an iterative feedback process which is represented graphically as a loop. The process steps are run per processing cell, i.e. a group of publication cells is processed simultaneously. The loop is repeated if an analist concludes that some publication figures within a processing cell are not plausible. In this case we speak of a new iteration for the macro-editing loop.

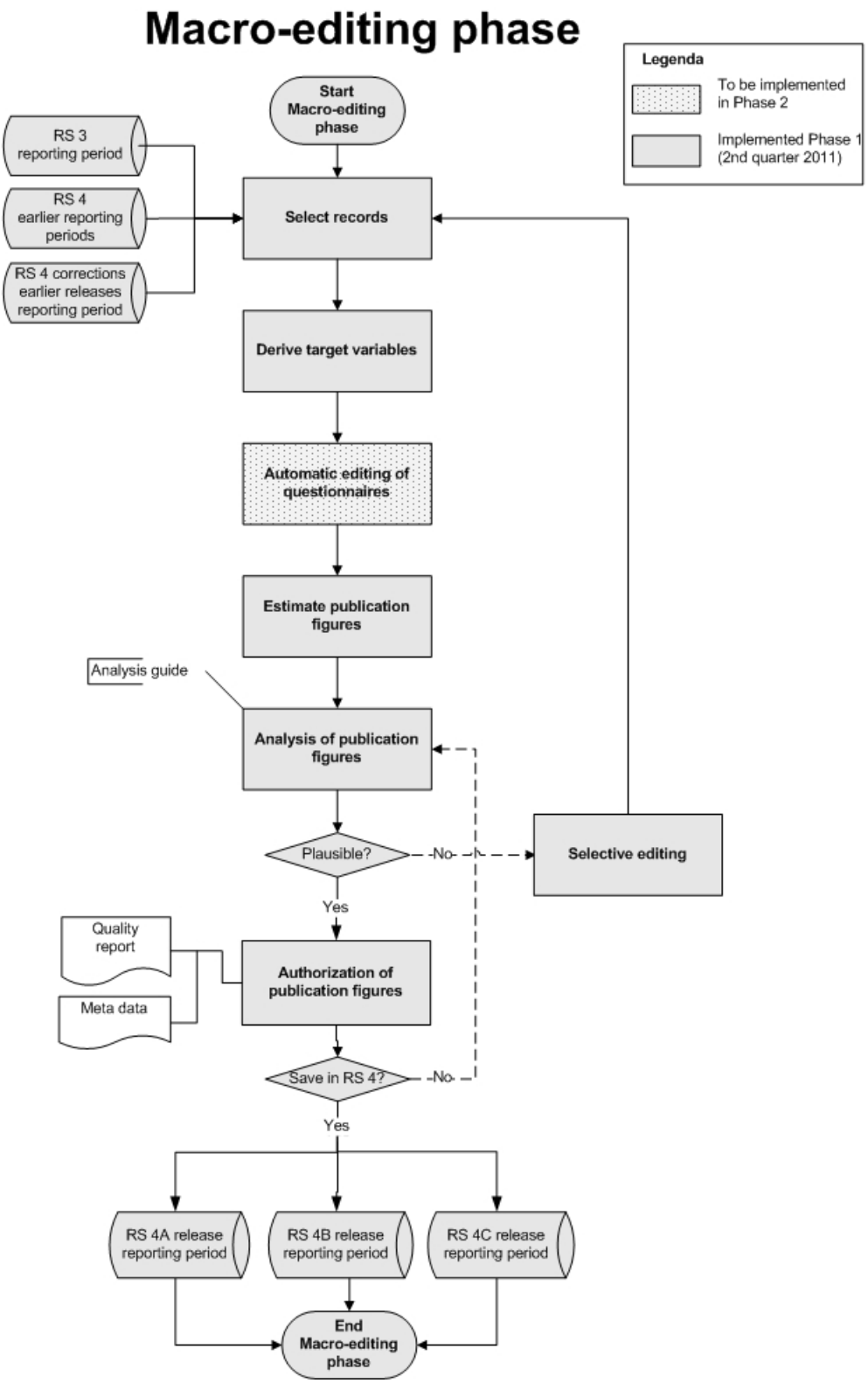


Figure 1. Macro-editing phase of DRT system.

16. *Select records*: An iteration of the macro-editing loop for a reporting period and release starts with

- data in rest stage 3 for the reporting period
- data in rest stage 4A, 4B, and 4C for earlier reporting periods
- corrections in rest stage 4B and 4C from earlier releases for the reporting period
- corrections made in earlier iterations

An analyst for a specific branch decides if data that became available after the start of the previous iteration are included. As an aid several indicators are available per data source. For each enterprise the system checks whether survey data or VAT data are used. Corrected values resulting from interactive editing in previous iterations are kept. Corrections made in earlier releases are applied to data that are loaded from rest stage 3. Some corrections relate to process steps before rest stage 3 and should therefore result in a new version of rest stage 3. For example, a correction of the NACE code.

17. *Derive target variables*: VAT data for a specific period does not necessarily match to the target period. Several automatic editing procedures are therefore needed, see section III. Furthermore, the definition of VAT turnover does not necessarily match to SN's definition of turnover. In publication cells where the difference between VAT turnover and SN turnover is somehow bridgeable the SN (target) turnover is derived from the VAT (source) turnover. This is done via a linear model.

18. *Automatic editing of questionnaires*: Unit errors (e.g. euros instead of thousand euros) and addition errors may occur in completed questionnaires. Furthermore, depending on the branche, there are one or more target variables for turnover. Possible target variables are turnover incl VAT, turnover excl VAT, inland turnover excl VAT, and foreign turnover excl VAT. When paper questionnaires are used it can happen that turnover excl VAT is not equal to the sum of inland turnover excl VAT and foreign turnover excl VAT. These cases are automatically detected and corrected. In Phase 1 of the DRT project this is done by means of the software for the current STS process IMPECT 2.

19. *Estimate publication figures*: An important step is the imputation of missing values for VAT units and enterprises. A missing value for turnover is only imputed if turnover is expected. Earlier edit actions are also taken into account. For example, if an editor estimates a missing value it is not imputed in the next iteration. Publication totals are computed by adding up all relevant observed, estimated and imputed values. Indices and developments are computed afterwards.

20. *Analysis of publication figures*: On the basis of the turnover level and development per publication cell, a set of indicators, tables with sorted micro data, and graphics an analyst decides whether a set of publication figures is plausible. See section V for details and Ossen et al (2011) for examples of graphics and tables.

21. *Selective editing*: The detection of records with an influential and suspicious turnover or yearly development is executed automatically, using principles of selective editing, cf. Granquist and Kovar (1997). That is, by detecting potential errors that influence the required output. To detect these potential influential errors (PIE), the influence and the risk associated with a turnover value in a publication cell is assessed. The method used strongly resembles the method described in Hoogland (2009). An analyst has a PIE list, graphics and tables available to select records for interactive editing.

22. *Authorization of publication figures*: When all processing cells for a specific branch are approved by analysts a branch project manager assesses the publication figures and the quality reports for each part of the macro-editing process. He/she then decides if there are any problems with process steps and/or additional macro-editing is necessary. If everything is found to be in order then publication figures, edited enterprise data, and edited input data can be loaded into rest stage 4A, 4B, and 4C, respectively.

## V. MACRO-EDITING METHODS

23. For analysis of DRT publication figures, selective editing, and authorization of publication figures we use a tool developed in Macroview[2], see Ossen et al (2011). Following a top-down strategy, we first take a look at the total picture. Macroview contains a screen where turnover developments are shown per publication cell for all branches. Those publication cells for which we expect about the same development are grouped into clusters. If developments of publication cells within the same cluster vary too much then this cluster is highlighted. That is, if

$$\max_{h\in c}\{\hat{g}_{h,r,v}^{k,k-s}\} - \min_{h\in c}\{\hat{g}_{h,r,v}^{k,k-s}\} > d_c^s \text{, where} \qquad (2)$$

$\hat{g}_{h,r,v}^{k,k-s}$: estimated turnover development for publication cell *h*, quarter *k* versus quarter *k-s*, release *r*, and iteration *v*

$d_c^4$: marginal value for the maximal difference between year-to-year development of quarterly turnover for two publication cells within cluster *c*

$d_c^1$: marginal value for the maximal difference between quarter-to-quarter development of quarterly turnover for two publication cells within cluster *c*.

If at least one of the marginal values is exceeded, the publication cells in the cluster deserve extra attention in the continuation of the analysis.

24. A clusters of publication cells is illustrated in table 4, including developments for the fourth quarter of 2009. These figures were determined on the basis of unedited data for illustration purposes only. Suppose that the maximal difference between the year-to-year and quarter-to-quarter development of publication cells within the cluster could be 3 percent. However, the maximal difference is 8,7 percent for the year-to-year development and 15,7% for the quarter-to-quarter development. Hence, the cluster shall be red-coloured in the Macroview tool and should be checked

*Table 4. Cluster of publication cells w.r.t. transport by water and service and developments for fourth quarter of 2009.*

| Cluster | Publication cell | Description | Year-to-year development | Quarter-to-quarter development |
|---|---|---|---|---|
| 50 | 50120 | Sea shipping | -27,7% | -3,3% |
| 50 | 50340 | Inland shipping | -19,0% | 10,8% |
| 50 | 52220 | Service for transport by water | -20,3% | -4,9% |

25. Population dynamics are also checked. Suspicious developments can be caused by substantial changes in the population frame. These changes can be realistic or can be caused by problems with the population frame. A distinction is made between new, disappeared and permanent enterprises. New enterprises just started or came from a different publication cell, i.e. the NACE code changed. Disappeared enterprises closed down or went to a different publication cell. Characteristics that are computed for each publication cell are

- number and turnover share of starting enterprises
- number and turnover share of new enterprises due to change of NACE code
- number and turnover share of permanent enterprises
- number and turnover share of closed enterprises
- number and turnover share of disappeared enterprises due to change of NACE code

[2] Software developed by SN that combines features of Blaise and R to analyse data coming from SQL-databases.

26. The imputation methods used in the DRT system are rather complex. An important question is whether enterprises with imputed turnover have in general the same development as enterprises with fully observed turnover. To check this, the following indicator is computed.

$$\frac{G_{h,r,E}^{k,k-s} - G_{h,r,O}^{k,k-s}}{G_{h,r,O}^{k,k-s}}, \text{ with} \qquad (3)$$

$G_{h,r,E}^{k,k-s}$ the growth rate of the panel of enterprises that exist in both quarter *k* and quarter *k-s* for publication cell *h* and release *r*

$G_{h,r,O}^{k,k-s}$ the growth rate of the panel of enterprises with a fully observed turnover in both quarter *k* and quarter *k-s* for publication cell *h* and release *r*

If indicator (3) exceeds a certain threshold value then it is highlighted in the Macroview tool.

27. Several indicators are used to assess the growth rate per publication cell

- Difference between estimated growth rate and expected growth rate
- Difference between estimated growth rate and growth rate for previous release
- Turnover share and growth rate per data source
- Turnover share of potential influential errors

28. A growth rate is suspicious if $|G_{h,r}^{k,k-s} - E(G_h^{k,k-s})| > d_h^E$ with $E(G_h^{k,k-s}) = G_h^{k-1,k-1-s}$. That is, the difference between the estimated growth rate and expected growth rate is too large. The expected year-to-year growth rate for a specific quarter is the year-to-year growth rate for the most recent release for the previous quarter. A growth rate is also suspicious if $|G_{h,r}^{k,k-s} - G_{h,r-1}^{k,k-s}| > d_h^G$. That is, the difference between the estimated growth rate for the current release and the estimated growth rate for the previous release is too large.

29. To assess the yearly turnover development and turnover level in quarter *k*, the turnover share and development are computed for each data source and publication cell. If the development for a specific data source is deviant with respect to other data sources special attention should be paid to the deviant data source. Furthermore, if the turnover share for a specific data source shows a large shift from one period to another this should be investigated.

30. To assess the yearly turnover development and turnover level in quarter *k* we also compute the turnover share of specific potential influential errors. We consider enterprises on the PIE list and enterprises that are selected for editing by an analyst. For this subset, we consider enterprises that are not checked by an editor, or where an editor indicated that there was insufficient information for editing.

31. The following indicator can be used to asses microdata that is used to estimate the yearly turnover development and turnover level for publication cell *h* in quarter *k.*

$$\frac{\sum_{j \in h} V_j^{k-4,k} \max\{O_j^{k-4}, O_j^k, \}}{\sum_{j \in h} \max\{O_j^{k-4}, O_j^k\}}, \qquad (4)$$

where

$V_j^{k-4,k}$ = 1, if a turnover value for enterprise *j* in quarter *k* is a PIE and it is not edited or an editor indicated that there was insufficient information for editing

$V_j^{k-4,k}$ = 0, otherwise.

## VI. INTERACTIVE EDITING OF MIXED SOURCE DATA

32. When a publication figure is not plausible, the PIE list and tables with microdata can be sorted. Bar diagrams, and scatter plots are also used to select records for editing. Scatter plots are used to detect outliers which can represent errors. Outlier detection serves as a check for the PIE list. For instance, it is useful to examine a scatter plot with turnover for quarter *k* versus turnover for quarter *k-4* or *k-1*, where enterprises on the PIE list have a different colour. When an enterprise needs to be edited this is indicated by a Macroview plugin. For such an enterprise turnover series are shown in the Macroview tool.

33. A record that is selected for editing can be edited via a Macroview plugin. There are several error types, depending on the data source, unit type, and branch

a) an enterprise has an erroneous feature ‘Belongs to target population’ (yes/no)
b) an enterprise has an erroneous feature ‘Temporary building consortium’ (yes/no)
c) an enterprise has an erroneous NACE code
d) an enterprise has an erroneous size class
e) an enterprise has an erroneous VAT turnover
f) an enterprise is wrongly matched to VAT data
g) a VAT unit has an erroneous VAT turnover
h) a VAT unit has an erroneous feature ‘foreign’ (yes/no)

34. In the building branch there are *temporary building consortiums (TBC)*. These are temporary joint ventures of several construction firms that work on a large building project. Building companies that receive an STS questionnaire have to fill in turnover including turnover obtained via TBC. We may receive VAT data, or a questionnaire for a building company that represents a TBC. In this case the turnover for this enterprise should not be used to prevent double counting. Statistics Netherlands makes a list with TBC and they are removed when records are selected in the macro-editing phase. However, the list may be incorrect. An editor therefore has the possibility to change the feature TBC for an enterprise

35. There are VAT units that have to declare turnover in the Netherlands, while SN considers them as foreign. The department International Trade of Statistics Netherlands makes a list with foreign VAT units. These units are removed when records are selected in the macro-editing phase. However, this list may be incorrect as well. An editor therefore has the possibility to change the feature ‘foreign’ for a VAT unit in consultation with the department International Trade.

36. The macro-editing phase for mixed source turnover data is tested and implemented in 2011. Initially, macro-editing is done for each quarter in 2010 and the first quarter of 2011 for the branches wholesale trade and car trade. In July 2011, we start producing a quarterly statistic for wholesale trade and car trade for the second quarter of 2011 based on the new statistical process. For other branches where quarterly turnover statistics are needed, implementation will start later this year.

**References**

Beuken, Y., and P. Vlag, 2010, *Business Register: the Dutch experience*, Paper for ESSnet Administrative data Seminar in Rome, March 2010.

De Jong, A., 2003, *IMPECT: recent developments in harmonized processing and selective editing*. Paper for the Work Session on Statistical Data Editing in Madrid, Spain, 20-22 October 2003.

Granquist, L. and J. Kovar, 1997, Editing of Survey Data: How Much is Enough? In: *Survey Measurement and Process Quality* (ed. Lyberg, Biemer, Collins, De Leeuw, Dippo, Schwartz, and Trewin), John Wiley & Sons, pp. 415-435.

Hoogland, J., 2009, *Detection of potential influential errors in VAT turnover data used for short-term statistics*. Supporting paper for the Work Session on Statistical Data Editing in Neuchâtel, Switzerland, 5-7 October 2009.

Ossen, S., W. Hacking, R. Meijers, and P. Kruiskamp, 2011, *MacroView: a generic software package for developing macro-editing tools*, Supporting paper for the Work Session on Statistical Data Editing in Ljubljana, Slovenia, 9-11 May 2011.

Van Delden, A., R. Banning, K. van Bemmel, A. de Boer, T. Carolina, J. Hoogland, M. van der Loo, M. Slootbeek, P. Ouwehand, H. van der Velden, and P.-P. de Wolf, 2010, *Methodology description DRT (In Dutch)*. Internal report, Statistics Netherlands.