WP. 4
English only

UNITED NATIONS STATISTICAL COMMISSION and ECONOMIC COMMISSION FOR EUROPE

CONFERENCE OF EUROPEAN STATISTICIANS

Work Session on Statistical Data Editing
(Madrid, Spain, 20-22 October 2003)

Topic (i): Development and use of data editing quality indicators

EVALUATION OF AUTOMATIC VERSUS MANUAL EDITING OF PRODUCTION STATISTICS 2000 TRADE & TRANSPORT

Invited Paper

Submitted by Statistics Netherlands, the Netherlands[1]

Abstract: Most production statistics 2000 of Statistics Netherlands were edited selectively. Each record with raw data was considered on the basis of a plausibility indicator. The implausible records were edited by hand. The plausible records were edited automatically with the computer program SLICE 1. We examined 12 of the 54 publication cells of Trade and Transport to check for differences between manual and automatic editing and see the influences of selective editing on the publication totals. The differences were small for most variables. Sometimes there were greater differences, also for key variables. In some publication cells for transport we found major deviations. We expect these differences can be reduced by improving the plausibility indicator and SLICE 1.

I. INTRODUCTION

1. The production process of most production statistics 2000 (PS 2000) of Statistics Netherlands was completely overhauled. Previously the statistical processes of the various statistics differed greatly. All these different procedures are now standardised. This increases the efficiency and manageability of statistical production. The new process consists of a uniform questionnaire, which is the same for all lines of business. It contains only a small amount of branch-specific questions. Furthermore, there are uniform weighting and editing procedures which are identical for all production statistics.

2. One aspect of the new editing procedure (UniEdit 1) is selective automatic editing (De Jong, 2002). Each questionnaire received is checked on the basis of a plausibility indicator for the quality of the data entered and the influence of the questionnaire on the weighted totals. Both the implausible and the very important questionnaires are edited manually, as was the case before. The other questionnaires are sent to SLICE 1 for automatic editing. This yields a significant gain in efficiency.

3. The question is: What is the effect of this change on the figures that will be published. SLICE 1 tries to correct the records on the basis of a set of editing rules. We expect the quality of the result to be lower than the optimal result that can be obtained by 100% manual editing. This is because the editors know more about the enterprises in a publication cell, and have a better sense of how plausible the reported values of the variables are.

[1] Prepared by Jeffrey Hoogland (jhgd@cbs.nl) & Eugène van der Pijll (epjl@cbs.nl).

4. In this study we measure the consequences of partly automatic editing for the published totals of the variables. For this purpose all records of twelve publication cells from trade and transport of PS 2000 are edited both automatically and by hand. The outcomes are compared, and the effect of selective automatic editing is calculated. This is done for different percentages of automatically edited records. Based on these results we can determine whether SLICE 1 can be used to do more of the work.

5. The two components of the editing process that are important for us are the plausibility indicator and SLICE 1, which we both describe in section 2. We describe the data that formed the basis of the evaluation in section 3. Our results are shown in section 4. Here we indicate what the effect is of selective automatic editing on the weighted totals. We mainly look at the ideal situation where there is a plausibility indicator available for each record. We also describe what happens when the percentage of automatically edited records changes. On the basis of these results we can determine what part of the records can be processed automatically in the future. We draw our conclusions in section 5.

II. THE NEW EDITING PROCESS FOR PRODUCTION STATISTICS

A. UniEdit

6. UniEdit 1 is the latest statistical process for micro-editing of production statistics of Statistics Netherlands. For many branches of industry it was first applied to the PS for the year 2000. UniEdit aims at a uniform editing process that is identical for all branches of industry so that the efficiency of production can be optimised.

7. After removing obvious mistakes, it is determined whether a record must be edited by hand or automatically. This is done with the plausibility indicator. We use the principle of selective editing, cf. Granquist (1995), Granquist and Kovar (1997), Hidiroglou and Berthelot (1986), assuming that records with influential errors should be edited by hand. The records are changed by either the editors or by SLICE 1 in such a way that all editing rules are satisfied.

8. UniEdit 2 is the latest statistical process for weighting and macro-editing of production statistics of Statistics Netherlands. It includes unit-imputation of large companies that did not respond, automatic outlier detection, weighting with auxiliary information, and authorisation of figures. Weights of records can be adjusted manually in the authorisation step when figures do not seem plausible.

B. Plausibility indicators

9. For PS 2000 the selection of records for automatic editing in a processing cell is based on a plausibility indicator (PI). A processing cell contains the companies in a specific publication cell and size class. The PI is a grade showing whether the values recorded are plausible. The PI is calculated on the basis of a number of partial plausibility indicators (PPI). Four of these PPI indicate per block of questions whether the recorded values deviate much from the medians of the previous year in that processing cell. The other three PPI are determined on the basis of external sources (VAT data, short-term statistics, data of the previous year), the quality of the response (hard errors, empty fields) and a number of ratios of important variables. For each of these indicators we first calculate a partial plausibility formula (PPF) that ranges from 0 to 8 and shows to what extent the raw values match expected values. Then the PPF are converted to PPI, which are grades ranging between 0 and 10. The PI is a weighted average of these grades, in which the low grades are given more weight than the high grades. The details of the PI are discussed in Hoogland (2002) and Hoogland & Van der Pijll (2003).

10. The records that get a failing grade (PI less than 6) are sent to the editors. The other records are plausible enough for us to assume that they can be edited automatically. The percentage of records that is edited automatically can be increased or decreased by shifting the cut off point for failing grades for the PI.

11. In the calculation of the PI we take into account what the influence of the record is on the weighted total. Because a relatively large error in a record of a major enterprise has a quite substantial effect on the weighted total, such an error is serious. Therefore major enterprises generally have a lower PI than small companies, and the percentage of automatically edited records is higher for major enterprises.

C. SLICE 1

12. When a record is considered plausible on the basis of the calculated PI, it is edited with the SLICE 1 program developed at Statistics Netherlands. SLICE 1 contains the module CherryPi, cf. de Waal (2000) and de Waal and Wings (1999). This module localises errors in a record on the basis of a number of editing rules that show the mathematical relationships between variables. It also contains an imputation module, which can apply mean, ratio, or regression imputation in the case of erroneous values. Finally, it contains a module that corrects imputed values when they violate edit rules.

13. The editing rules violated by a record have to be satisfied by CherryPi by changing the values of one or more variables. Usually there are several possibilities for editing a record. This requires a choice, which is made according to the *principle of Fellegi and Holt*, see Fellegi and Holt (1976). The principle states that it is more likely that there is one major error in one variable, than that there are smaller errors in more than one variable. The best solution for editing a record is to change as few variables as possible.

14. In practise some variables contain fewer errors than others, or are not allowed to be changed by as much as other variables. That is why each variable in CherryPi is weighted for reliability. This implies that a change in one variable can weight a number of times heavier than a change in another variable. The best solution for the record is obtained by minimising the sum of the reliability weights of the variables changed, such that all edit rules are satisfied. The resulting constrained minimisation problem is solved with the Chernikova algorithm, cf. Chernikova (1965).

15. When several variables have the same weight, CherryPi often finds several equally good solutions. The output given by CherryPi contains all these solutions, from which one must be selected. In this case the first solution is always chosen. When many editing rules are violated, many variables must be changed. In such cases CherryPi cannot find a solution because there are too many possibilities to check and the record must be edited by hand anyway.

III. DATA USED

A. Selected publication cells

16. The returned questionnaires are divided into a number of publication cells depending on the SIC (Standard Industrial Classification; dutch version of the NACE) of the respondent. These publication cells contain one or more SICs. During the editing process we no longer distinguish records with different SICs within a publication cell. This implies that the publication cells must be as homogenous as possible for the editing process to work well. In contrast with this is the demand that a publication cell must be big enough to allow the application of statistical methods.

17. The twelve publication cells used for evaluation are given in table 1. They contain 4651 records in total for PS 2000. Some 45% of these records have a sufficient PI. Because the PI was not ready in time for production, records were not edited automatically the first few months. This is why only 31% (1454 records) were edited by SLICE 1 during production. These records were later edited by hand in the course of this study. The records that were edited by hand during production were later edited automatically.

Table 1. The publication cells selected for evaluation.

	Wholesale trade
151200B	in flowers and plants
151300C	in food, beverages and tobacco excl. fruit, vegetables and potatoes
151600C	in tool, machinery for agriculture/textile production
	Retail trade
152110	in food, beverages and tobacco in shops; super markets
152121C	in furniture, household textile, lights, and household articles
152121E	in hardware, tools, paint and construction materials
	Transport
160220	Irregular transport of people by taxi
161100	Shipping at sea
161200	Inland shipping
163110	Loading, unloading, warehousing
163300	Travel organisation and mediation; information for tourism
163400	Shipping agents, cargo insurance and chartering brokers; weighing and measuring

18. The enterprises in each publication cell are categorised by size class in three processing cells: size class 1 (less than 10 employees), 2 (from 10 to 100 employees), and 3 (at least 100 employees). The questionnaires consist of cover pages and inserted pages. Cover pages are less extensive for size class 1 than for size class 2-3, but they do not contain specific questions on publication cells. These are on the inserted pages. In the publication cells we studied there are five different inserted pages.

19. We encountered a problem when we re-edited the wholesale trade records in size class 1 by hand. We could not produce all necessary files for the interactive editing program. So for the corresponding processing cells we only have part of the records available that contain both automatically and manually edited data. These processing cells were therefore eliminated from the study. After removing the records of wholesale trade size class 1 we had 4162 records left. All weighted totals were determined on the basis of these records.

20. We have both manually and automatically edited data available for these 4162 records, except for 247 records for which SLICE 1 could not find a solution. This is mainly due to the poor quality of these records, which means too many variables had to be adjusted. Although it may be possible to edit these records automatically anyway by improving SLICE 1, this is not the way we would go. The quality of these records is so low that they should be checked by an editor in any case. Table 2 shows that most of these records received an insufficient PI and would not end up with SLICE 1 during production. The percentage of records with a PI of 4 or more for which SLICE 1 could not find a solution was very small.

21. Another group of records, which is completely edited by hand, consists of businesses in size class 3. These major enterprises attribute so much to the published total, that it is very important that they have optimal quality data. This concerns 216 enterprises. Together with the group that cannot be edited automatically, there are 449 records (11% of the total in the twelve publication cells) that have to be edited by hand in any case.[2] This means that the percentage of automatically edited records cannot exceed 90%.

[2] Because the two groups of records overlap somewhat, the total number is smaller than the sum of the sizes of the two individual groups.

Table 2. Records for which SLICE 1 cannot find a solution, by PI value.

PI	Total number of records	No solution	
		number	percentage
0	222	66	29.7%
1	560	76	13.6%
2	367	30	8.2%
3	303	20	6.6%
4	364	14	3.8%
5	373	11	2.9%
6	395	12	3.0%
7	593	14	2.4%
8	685	4	0.6%
9	292	0	0%
10	8	0	0%
Total	4162	247	5.9%

B. Publication totals in evaluation versus actual publication totals

22. Part of the manually edited data in this evaluation come from the production of PS2000 and part were edited especially for this evaluation. We assume that the editing quality of the records later edited by hand is as good as the editing of the records that were originally edited by hand.
There are still differences between the PS2000 production and our evaluation. This is mainly due to the following points:

1. The implementation of the PI was not ready when manual editing was started for production. So at the start of the editing, records with unknown PI's went to the editors. Some of these may have had a sufficient PI, and could have been edited automatically. In the analysis we assume that all records that would have had a sufficient PI went to SLICE 1. So this means we could overestimate the inaccuracy of the published totals for PS 2000.

2. In the production of PS 2000 we used a new weighting method. However, during this analysis we used an old weighting method, because we were only interested in the trend breach as a consequence of the introduction of selective editing. In the old method direct weighting is done per SIC and company size on the basis of the population and survey sample sizes, excluding outliers. A number of records were already identified as outliers in the editing phase. These records were given a weighting factor of 1.

3. For the evaluation we used the production database as it was right after micro-editing and before the UniEdit 2 process. So we did not have the outlier indications that were determined in that process.

23. The weighted totals we calculated are therefore not identical to the published totals. Furthermore, aggregated deviations between records that were edited manually and automatically, which are calculated in the next section, can differ from these deviations in practice.

24. Because more records were edited manually than was necessary on the basis of the PI, and because the greatest deviations from the PS 1999 figures were smoothed out in the authorisation step, the bias in the published totals will be lower than the values we calculated. We are mainly interested in the effect of automatic editing as it will take place in the future. We therefore feel that it is more important to look at the ideal case, in which the PI is available from the very beginning, and no work has to be done during authorisation.

IV. DEVIATIONS AS A RESULT OF AUTOMATIC EDITING

A. Pseudo-bias

25. The bias of a weighted total after partly automatic editing indicates how big the deviation is from the real value. Because the real total is unknown we cannot determine the bias. However, we are concerned with the difference between automatic and manual editing. So we can approximate the bias with the difference between the weighted total of 100% manually edited data and that of partly automatically edited data. We call this value the pseudo-bias. It is expressed as a percentage of the manually edited weighted total.

26. The pseudo-bias of a weighted total is defined as follows for each variable y_j and for each publication cell

$$\text{pseudo} - \text{bias}(y_j) = \frac{\sum_A w_i(\tilde{y}_{ij} - \hat{y}_{ij})}{\sum_C w_i \hat{y}_{ij}},$$

with

A : set of automatically edited records in a publication cell;
C : set of all records in a publication cell;
$\tilde{y}_{ij}$: automatically edited value of variable y_j in record i;
$\hat{y}_{ij}$: manual edited value of variable y_j in record i;
w_i : weighting factor of record i.

27. The pseudo-bias depends on the selection of records that have to be edited automatically (and therefore on the PI) and on the quality of automatic editing with SLICE 1. The aim is to keep the influence of automatic editing on the weighted totals to a minimum. So the pseudo-bias must be as small as possible.

B. Pseudo-bias as a result of selective editing

28. The pseudo-bias is calculated for all variables and twelve publication cells on the basis of the PI as it could have been used in 2000. Due to problems, specified above, with the PI calculation during production of PS 2000, the selection of automatically edited records in our study is not identical to the selection of automatically edited records during production.

29. There is a great deal of variation in pseudo-bias. For most variables the pseudo-bias is close to zero. This is because many variables are hardly changed by editors or SLICE 1. This is mainly true for less important variables. Variables that show major deviations, however, are also usually unimportant ones. For important variables the pseudo-bias is at most 15%. The pseudo-bias of 15% is caused by errors in the program that corrects obvious mistakes. These errors have been removed for PS 2001. Details of this research are given in Van der Pijll & Hoogland (2003).

30. Pseudo-biases in publication cells for transport are usually greater than those for trade. This is because questionnaires for transport are not filled in as well as those for wholesale and retail trade. One way to explain this is that cover pages of the questionnaires are the same for all publication cells trade and transport, while they are based on questionnaires for trade PS 1999. This may confuse respondents from transport, because definitions in variables may differ.

31. For some variables we found the cause of deviant editing by SLICE 1. These deviations were mainly found in the publication cells for transport. It will be possible to remove major deviations in these variables in the future by

- improving questionnaires
- splitting heterogeneous publication cells
- adapting the software that corrects obvious mistakes
- improving the PI
- adding a number of edit rules
- adjusting reliability weights
- improving the error localisation module within SLICE
- improving the imputation module within SLICE
- building an extra step in the statistical proces before SLICE, which removes systematic mistakes that do not follow the Fellegi-Holt principle

C. The effect of more automatic editing

32. One key question is whether the percentage of records that is edited automatically can be increased or must be decreased. We can study this by varying the selection of records to be edited automatically. We can again calculate the pseudo-bias for each selection. This will generally be larger as the percentage of automatically edited records increases. In some cases the errors of the added automatically edited records can cancel out some of the existing deviation, so that the pseudo-bias is reduced. However, these are only incidental cases, and we should not count on them.

33. We varied the threshold for sufficient grades. This means we varied the number of records for which the PI is sufficient. Table 3 shows the percentage of records that is automatically edited at the given threshold. When the threshold is 6, as was the case for PS 2000, the percentage of automatically edited records is between 43 and 62 percent for most publication cells. It is impossible to determine the percentage in advance because the PPI are calibrated on the basis of raw and edited values of the previous year. Therefore the percentage of automatically edited records fluctuates for each publication cell. Apparently the PI was very severe for publication cell 152110, because only 28% of the records were deemed plausible enough for automatic editing. When the threshold is 4 the percentage of automatically edited records will exceed 60% in most publication cells, whereas a threshold of 2 will generally correspond to an automatic editing percentage from 72 to 90 percent.

34. In appendix A we show how the pseudo-bias depends on the threshold for three key variables. Table 4 shows that it is very difficult to automatically edit the number of employed persons in publication cell 161200 (inland shipping) when the threshold is set at 6. The deviation between automatic and manual editing is over 5% here. This is partly due to a systematic error made by the respondents, which can not be corrected by SLICE 1. There are no significant problems in the other publication cells. When the percentage of automatically edited records exceeds 60%, the pseudo-bias in the total number of employed persons in publication cells 152121C and 163300 reaches more than 2%.

Table 3. Percentage of automatically edited records per threshold.

Publication cell	1	2	3	4	5	6	7
151200B	83%	72%	64%	60%	55%	48%	38%
151300C	83%	79%	73%	65%	58%	53%	44%
151600C	85%	74%	62%	52%	46%	43%	36%
152110	81%	65%	54%	46%	37%	28%	22%
152121C	88%	77%	68%	62%	53%	43%	41%
152121E	91%	82%	74%	65%	55%	48%	45%
160220	83%	73%	64%	59%	55%	44%	36%
161100	94%	90%	85%	80%	73%	62%	43%
161200	93%	86%	80%	77%	71%	62%	53%
163110	89%	81%	77%	77%	70%	61%	48%
163300	87%	72%	66%	59%	49%	38%	34%
163400	83%	72%	67%	62%	53%	44%	30%
All twelve cells	86%	76%	68%	61%	54%	45%	36%

35. The variable net turnover (table 5) is correctly edited automatically in almost all publication cells. Raising the percentage of automatically edited records to 80% causes virtually no problems for these variables. The only publication cell in which there is a major difference between manual and automatic editing is cell 163300 (travelling organisations and travel intermediation). This is because respondents often filled in some purchase value while the variable in this publication cell should almost always be zero. The editors usually removed the purchase value and balanced it with net turnover, whereas SLICE 1 left the records unchanged. This leads to major pseudo-bias in these variables, which cannot be avoided by applying extra editing rules. The problem will continue to show up in future, because the Fellegi-Holt principle does not hold. One long-term solution for this problem is an extra editing round focusing on some specific errors such as this.

36. SLICE 1 does a fairly good edit of total operating result (table 6) for all publication cells. However when SLICE 1 starts to edit more records some difficulties show up in various publication cells. For publication cell 163300 these occur when the threshold is set at 5 or less. This corresponds to a percentage of records to be edited automatically of over 50%. In publication cell 163110 (loading, unloading, warehousing) the problems start around 75%.

37. The pseudo-bias for most variables and publication cells is small. It does not get much higher either when the number of records edited by SLICE 1 increases up to 70%. This is not true for variables with a large pseudo-bias in a publication cell. For these variables and publication cells we can see a rapid increase in pseudo-bias when the number of automatically edited records increases.

38. Tables 4-6 show several high percentages. This does not mean that these major deviations ended up in the published totals, though. The weighted totals of these variables may have been corrected by the automatic outlier detection, the new weighting method, or during authorisation, substantially reducing these deviations.

V. CONCLUSIONS

39. We examined differences between manually and automatically edited weighted totals of PS 2000 prior to authorisation. We assumed that the plausibility indicator was operative during the entire editing period. The evaluation was made for twelve publication cells in the wholesale and retail trade, and in transport.

40. The effect of selective editing differs per variable. Most variables are hardly changed during manual or automatic editing. This is true both for less important variables and for some key variables such as total operating costs, total labour costs and total operating profits. The weighted totals for these variables hardly change in most publication cells when selective editing is used. Even when the percentage of automatic editing increases to 80% the deviation in the weighted totals for these variables stays under 2%.

41. However, for some variables the deviations of selective editing are large. These are mainly variables from the results block of the questionnaire, such as the result before taxes and the financial result. The current 45% threshold for automatic editing already yields many deviations of more than 5% in the weighted totals of these variables. When the percentage of records for automatic editing increases, the quality of these variables will plummet.

42. We found the greatest deviations in publication cells in transport. The problems are such that these publication cells will have to be edited more by hand, rather than less. The deviations in wholesale and retail trade are smaller. It depends on the level of bias in the published figures that is considered acceptable whether we can gain in efficiency by more automatic editing. The biases mentioned in this paper are based on how the PI and SLICE 1 worked during PS 2000. There may well be less bias in most variables in the future when the PI and SLICE 1 are improved.

43. We have found room for improvement on a number of points. By adding a few editing rules, by developing software that removes systematic errors, and by improving the questionnaires we can come up with considerable improvements for the variables. Some of these improvements can be applied to large numbers of publication cells, also well beyond the twelve cells we studied here. Other improvements are publication cell specific (for instance pertaining to inland shipping and travel organisations). When these improvements are implemented those publication cells where SLICE 1 currently produces major deviations may well be edited without major problems.

References

Chernikova, N.V., 1965, Algorithm for finding a general formula for the non-negative solutions of a system of linear inequalities. *USSR Computational Mathematics and Mathematical Physics,* **5**, pp. 228-233.

Hoogland, J.J., 2002, Selective editing by means of Plausibility Indicators. Research paper no. 227, Statistics Netherlands, Voorburg.

Hoogland, J. J. en E.C. van der Pijll, 2003, Evaluation of the plausibility indicator for production statistics 2000 Trade & Transport (In Dutch). Internal paper, Statistics Netherlands, Voorburg.

Granquist, L., 1995, Improving the Traditional Editing Process. In*: Business Survey Methods* (ed. Cox, Binder, Chinnappa, Christianson, and Kott), John Wiley & Sons, pp. 385-401.

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.

Hidiroglou, M.A., and J.-M. Berthelot, 1986, Statistical Editing and Imputation for Periodic Business Surveys. *Survey Methodology*, **12**, pp. 73-83.

Jong, A.G. de, 2002, UniEdit: Standardised processing of structural business statistics in The Netherlands. Statistics Netherlands, Voorburg.

Van der Pijll, E.C. en J. J. Hoogland, 2003, Evaluation of automatic versus manual editing of annual structural business statistics 2000 Trade & Transport (In Dutch). Internal paper, Statistics Netherlands, Voorburg.

Waal, T. de, 2000, SLICE: generalised software for statistical data editing and imputation. In: *Proceedings in computational statistics 2000* (ed. J.G. Bethlehem and P.G.M. van der Heijden), Physica-Verlag, Heidelberg, pp. 277-282.

Waal, T. de, and Wings, 1999, From CherryPi to SLICE. Report BPA-no 461-99-RSM, Statistics Netherlands, Voorburg.

Appendix A. Pseudo-bias in publication totals for several key variables.

A column in a table shows the pseudo-bias resulting from automatic editing of records with a PI equal to the threshold or higher. When the pseudo-bias exceeds 5% it is printed in bold character. Pseudo-biases between 2% and 5% are underlined.

Table 4. Pseudo-bias in the total number of employed persons.

	Threshold for plausibility indicator						
Publication cell	1	2	3	4	5	6	7
151200B	1.1%	1.1%	0.5%	0.6%	0.6%	0.1%	0.3%
151300C	0.3%	0.1%	0.2%	0.3%	0.2%	0.2%	0.1%
151600C	0.8%	0.2%	0.1%	0.1%	0.1%	0.1%	0.1%
152110	0.8%	0.3%	0.1%	0.0%	0.0%	0.0%	0.0%
152121C	3.6%	2.4%	2.4%	2.5%	1.8%	1.5%	1.4%
152121E	0.1%	0.5%	0.4%	0.5%	0.8%	1.0%	1.1%
160220	2.5%	0.7%	1.2%	1.0%	1.0%	0.9%	0.8%
161100	1.5%	1.4%	1.4%	1.4%	1.4%	1.4%	1.3%
161200	**7.7%**	**5.4%**	**5.1%**	4.1%	5.0%	**5.4%**	4.2%
163110	1.2%	1.4%	1.4%	1.4%	1.4%	1.4%	1.1%
163300	2.3%	2.1%	2.1%	0.5%	0.2%	0.0%	0.0%
163400	0.1%	0.2%	0.3%	0.1%	0.2%	0.3%	0.0%

Table 5. Pseudo-bias in net turnover.

	Threshold for plausibility indicator						
Publication cell	1	2	3	4	5	6	7
151200B	0.2%	0.2%	0.0%	0.0%	0.0%	0.0%	0.0%
151300C	0.2%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%
151600C	0.2%	0.2%	0.0%	0.0%	0.0%	0.0%	0.0%
152110	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
152121C	0.3%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%
152121E	0.6%	0.2%	0.2%	0.1%	0.0%	0.0%	0.0%
160220	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
161100	**5.1%**	0.2%	0.0%	0.1%	0.1%	0.1%	0.1%
161200	0.8%	0.4%	0.4%	0.4%	0.4%	0.5%	0.0%
163110	0.6%	0.3%	0.3%	0.3%	0.2%	0.2%	0.2%
163300	**34.6%**	**21.6%**	**18.0%**	**16.8%**	**13.1%**	**8.7%**	**7.7%**
163400	0.4%	0.4%	0.4%	0.4%	0.2%	0.1%	0.1%

Table 6. Pseudo-bias in the operating result.

	Threshold for plausibility indicator						
Publication cell	1	2	3	4	5	6	7
151200B	0.8%	0.2%	1.3%	1.1%	0.7%	0.6%	0.2%
151300C	0.4%	0.5%	0.5%	0.9%	0.1%	0.1%	0.1%
151600C	0.6%	0.2%	0.5%	0.3%	0.3%	0.2%	0.2%
152110	1.5%	1.7%	1.4%	0.3%	0.2%	0.2%	0.0%
152121C	**10.1%**	1.0%	2.4%	1.9%	1.9%	1.7%	1.7%
152121E	**12.7%**	0.8%	0.8%	0.3%	0.3%	0.1%	0.1%
160220	1.4%	1.3%	1.3%	1.2%	1.2%	2.1%	2.0%
161100	**8.2%**	2.7%	0.0%	0.3%	0.2%	0.5%	0.5%
161200	0.5%	0.2%	1.0%	1.2%	1.2%	0.3%	0.5%
163110	4.9%	**5.3%**	**5.3%**	**5.3%**	1.0%	0.9%	0.6%
163300	**8.0%**	**7.8%**	**8.0%**	**8.2%**	**8.3%**	0.6%	0.6%
163400	**39.5%**	0.8%	1.1%	1.1%	0.2%	1.1%	0.2%