sciendo

Research Article

Identification of Quality of VAT Data: The Case of Albania[1]

Elsa Dhuli

Abstract

The increasing demand of users for quick and qualitative statistics, have increased the focus of statisticians from different countries to develop strategies, produce systems, compile statistical methods for the use of administrative sources as primary source. An important issue of official statistics is the reduction of the response burden from enterprises. Administrative sources are frequent in time, have wide scope coverage, are fast in terms of deadlines, but in content they carry suspicious values. In the case of Quarterly Survey these suspicious values affect the quarterly growth rate analysis. In Short Term Statistics cases where the value between the quarters is not accurate give a wrong picture of the development. In the Structural Survey, an annual net sale gives a wrong picture of the structure of the economy. Sources used for this paper are Value Added Tax and Quarterly Survey of Enterprises. The paper tries to identify the different patterns of the VAT in Albania by analysing two years VAT data, 2016-2017 and propose the editing method for correcting them. As the focus is the use of VAT in Short Term Statistics, the patterns are identified in the panel data created by linking with the same ID code the businesses on both sources. Various methods have been used from other countries' statistics for the editing of these suspicions values. The experience of these countries is studied in orded to be used in Albania as well. This process will reduce the time for processing VAT file, will be used in effective way for quarterly business survey or structure business survey purposes and will increase the quality of statistical procces and product by using effectively the VAT file.

Keywords: Identification suspicious value; Patterns; VAT file

1. Introduction

Information obtained from administrative sources that serves as an input for the production of official statistics is not compiled for statistical purposes, but for dealing with administrative purposes. Currently in Albania, the data from administrative sources are used as auxiliary information for imputation and treatment of missing values in statistical surveys, mainly in Quarterly Business Survey. To reduce the response burden, to reduce costs and increase the efficiency of statistical surveys the transition from the traditional survey-based model estimation to alternative source estimates sets the need for using some sources (Hoogland, 2009).

The process of using VAT file for statistical purposes requires some steps to be followed, from obtaining the data to the publication of results. An important step after the preparation work for obtaining data from the Tax Authority is the identification of suspicious values, error values, estimation of those values and their imputation and correction. The task of statisticians is to identify these errors and to find the estimation methods to reduce the error rate. Identifying suspicious values and editing them in the administrative source is one of the necessary steps before their use. Researchers like Granquist and Kovar (1997), Hoogland (2010), Lewis (2012) and others have developed and implemented strategies for editing suspicious values in administrative sources.

---

[1] *The study is an independent research and shall not be taken as an official institution opinion*

In the business statistics, the pressure to reduce the response burden from respondents and the increasing demand to use the VAT file has led to the need for studying and analyzing the possibilities of its use in several ways: reducing the number of surveys, reducing the burden of a survey in the direction of the partial replacement of enterprises by two dimensions: economic activity and size of enterprises, and estimating the missing values from the survey through imputation. In this paper the first section focuses in the statistical procces needed for controling VAT file; the second explains the suspicious value patterns and why the selective editing method used in VAT file and third analyses the identified patterns in VAT file in case of Albania. The conclusion describes that the imputation and editing process is neccessary in VAT file and the different editing methods have to apply for different patterns by the two dimensions: economic activity and size of enterprises for the purpose of quarterly survey in the short term statistics.

The main objective of this paper is to find out possible types of error values that have a suspicious impact on the net sales variable, using selective editing principles. Granquist and Kovar (1997) have explained in details the principles of selective editing in their research for errors that actually occur. Editing VAT values can be considered in two phases: input and output editing (Hughes et.al, 1990). Value control in an isolated area from information obtained from other resources such as, the VAT file, requires the use of selective editing methods. Input editing usually requires the implementation of a set of rules. Failure to identify suspicious errors or values and not apply the editing rules negatively affects the net sales result. For this reason, identifying potentially suspicious input values and implementing editing rules reduces errors in the final result, thus increasing the quality of the data. This does not necessarily mean locating and correcting all the mistakes. For more information see: Granquist (1995) and Granquist and Kovar (1997). Selective editing is performed for values that are really problematic and have an impact on the final net sales result. In the case when the declaration of VAT since 2014 is done online by businesses 14 days after the current month and the ability to re-contact the enterprises for further clarification of suspicious, null or negative values are impossible, methods used significantly reduce the amount of error or doubt in the net sales values.

By Hoogland (2010), the models of suspicious values vary; each of them has a special editing method. For example: the sales variable may have negative values. The negative value according to the concept by the tax authority may be a deduction of more than one declaration at a previous stage, but it may also be a correction of previously evaluated statements. This is a problem in interpreting the growth rate of net sales for a related period since this value can result negative in the end, which is difficult to interpret and compare with positive values. For example, if an enterprise has a value of 1500 ALL in quarter t-4 and -500 ALL in the first quarter. The growth rate is negative (1500 / -500 = -3) and negative value is difficult to interpret or to compare with positive growth rate. However, the negative value is a matter of doubt. In this case the method used is the transformation of negative values into positive values. This method is better to use when two years are analysed and the negative values is faced only in one year.

The patterns in the form {(0,0,0, x), (0,0,0, y)}, mean that in the first three quarters the net sales value is "0" and in the fourth quarter is positive value, and if this pattern continues to be the same in the coming years, this shows that the value of net sales in the fourth quarter represents the year. The method set out in this case is through the imputation of the quarterly net sales values, estimating the three previous quarters. The other patterns as {(0, 0, x, 0), (0, 0, x, 0)}, mean that the seasonal activities have to consider if the values are in their normality in time series.

2. Methods and Sources Used

2.1 Methods and steps followed

The important issue in using VAT file is the production of reliable statistics based on identified administrative sources such as VAT file or combined different administrative sources. Hoogland (2009) explain that the statistical procceses required to control VAT file are identified, from Preparation and Receipt data to the Publication of the results. For the purpose of this paper only three procceses are treated, where two firsts are mentioned and third is treated in more details.

For the editing of VAT file errors, according to Hoogland (2009), the implemented strategy consists of three preliminary stages:

- Preparation and receipt of data;
- Technical standardization of the data received;
- Identification of systematic errors generated by the data obtained from the VAT file.

2.1.1 Preparation and receipt of data

Prior to receiving the data file, periodic review of the legal framework is carried out (see[2] for the legal framework) to see the inclusion or elimination of the Enterprise Value Added Tax (VAT) scheme and the change of applicable regimes on them, for special economic activities. In Albania the Tax regime is in three levels 0 percent, 6 percent and 20 percent. This process is used to identify user needs and to analyze the extent to which these needs are met at the time of recieving the file.

The cost-effectiveness assessment process or the restrictions and the quality of the use of an administrative source at the moment of receipt by the tax authority are elements that help to determine the extent to which this file is usable and to what extent it will be used. In the receiving phase, based on the quality of the VAT file, it is decided if the VAT file will be used directly or indirectly as an administrative source. VAT is indirectly used to validate quarterly Short Term Statistics (STS) or Annual Structure Business Statistics (SBS) data estimates; for imputation and classification of enterprises by size (small, medium, large); for comparative macro level analysis (VAT and SBS) and as input data on the imputation models (STS or SBS).

A necessary element is the definition of key components and variables that will be used depending on the purpose of using the VAT file, either directly or indirectly. The list of variables is determined by the study of the fields defined in the VAT declaration form compiled by the tax authority, mandatory to fill each month by businesses (Annex 1, Table 1 List of variables). Data from the Tax Authority are transmited to statistical office 25 days after the current month.

Taxpayers are divided into three categories: (1) those who realize 8 million ALL net sales value in a year declare and pay VAT each month and report to Tax authority 14 days after the current month; (2) medium taxpayers who make 2-8 million ALL net sales value in a year, declare and pay every three months on a monthly basis 14 days after the current quarter and (3) micro and small taxpayers who realize less than 2 million ALL net sales value per year and declare once a year. This division helps statistics to understand and analyze the constraints created from a tax file that is being used by the statistician. The study is limited to the VAT file.

2.1.2 Technical standardization of the data received

After exporting to the database, the file is subject to control for possible errors occurring during data transmission due to the selected delimiter or property of the variables as well as the consistency of the variables. In the data collection instrument, basic rules for quality control have been built up. Controlling the taxable value under regimes 20 percent, 6 percent and 0 percent or special regimes for particular economic activities such as travel agent regimes and the declared value in the declaration form affects the analysis by economic activity. Other controls include periodicity, dublicate values and identification of negative sales values or cases where the economic activity NACE Code (even if not used from Statistics) is missing in the file is done at this stage of the standardization process.

The VAT file is integrated with the Business Register on the basis of the common variable (NIPT-Tax ID Number) and populated with the demographic variables of the enterprises: Activity in 4-digit of NACE, City, Legal Form, Status and Ownership (see Annex, Table 1). VAT file data are aggregated at Enterprise level on a monthly, quarterly and annual basis. The first two stages of the

---

[2] https://www.tatime.gov.al/eng/c/6/71/value-added-tax

statistical process relate to the identification of needs, confirmation of needs and preparatory phases of the VAT file from the recipient to the user. Those two steps (first and second) are not in the focus of this study.

2.1.3 Identification of suspicious value patterns in VAT file

The paper focuses on identifying and editing errors in the data obtained from the VAT file. One of the elements that are taken from the tax file is the Net sales value, which may be in future relevant to produce directly the statistics on Net sales from business. At this stage is important the VAT declaration for a specific period related to that period. For the tax authority this approach is not important for the administrative procedure, but takes particular importance in the statistical context. Even the delivering deadline of VAT file which is 25 days does not happen in regular way. VAT data is reviewed, evaluated and edited in relation to the values of Net sales, as it happens that they can declare negative, empty or "0" values. In the case, for example, where the value "0" is for a full quarter it may happen that the enterprise was not active for that period. But there are also cases where they declare "0" for a monthly period and continue declaring very high value at the end of the quarter, which shows that the enterprise reports once every three months. But there are also cases where VAT file has very low or very high extreme values in the interim periods of a quarter or between quarters. The negative value could be considered as reembursment of the highest value declared before.

To identify the units with suspicious or incorrect Net sales values, the detection rule of value have been applied (Ouwehand, 2010). The experience of other countries implemented in the conditions of Albania has its advantages and disadvantages. Some of the quarterly models are only corrected when they occur for two consecutive years (Hoogland, 2010). For the purposes of this study, this rule has been implemented in a two-year time series, 2016-2017 of the VAT file. For example, a single VAT unit is detected when sales for 8 quarters are {(0, 0, 0, x), (0, 0, 0, y)}, where $x > 0$ and $y > 0$. In such cases, the sales value can not be used for quarterly periods, but only for annual periods. At the moment when the VAT unit does not follow this model, but continues in the third quarter with quarterly value, it is taken into account for each quarter {(0, 0, 0, x), (0, 0, 0, y), (z, z, z, z)}, where $z > 0$. But this does not happen. From the control of the VAT file, several different combinations of suspicious values or errors that were considered as models were discovered. Each model is individually analyzed for the specified and the correction method.

The main reasons for having suspicious or wrong values in the VAT file may arise as a result of: i) erroneous statement by the VAT unit; ii) erroneous completion of the standard VAT declaration form; iii) mistakes in the declaration of VAT value by fulfilling online standard form for a specified period more than adequate; iv) declaration depending on the established VAT regime; vi) frequent changes in the VAT regimes; vi) declaration depending on the activity in which is part of; vii) reimbursement of the VAT. Important in these reasons is the identification of the most important ones that could effect in the erroneous VAT value.

2.2 Sources used

The main source used to conduct this study is VAT data file linked with Quarterly Business Survey (QBS). As we mentioned before the NACE code in VAT file it is not as qualitative as while this code is received from QBS. Also to analyse are taken the panel data by linking the sample of QBS with VAT. In the taxation system, one transaction is part of the application field if the activity is good supply or service supply, done against payment, within the territory of the Republic of Albania, by a taxable person acting as such.

2.2.1 Types of errors in the VAT file

The types of errors associated with the VAT file are on their continuity: reference period, data errors, data completeness, errors in economic activity, net sales, errors in the statement, coding errors. The VAT file contains the errors as below mentioned as from the Internal report (INSTAT):

Reference period – If the reporting month is 14 days after the end of the reference month for all units subject to VAT and complete the monthly online statement. Other units complete the simplified tax declaration and / or income tax and report every three months or once a year.

Data Errors - overlapping of data occurs when the taxpayer corrects his tax declaration.

Data Completeness - Files close to the statistics come in different time periods, as such the reported values carry changes including declarations carried out in the past by units or corrections of values previously reported. As a result, in different periods it turns out that we have complete and updated data of the file.

Net sales errors - related to the main variable Net sales, this variable is not found as such in the VAT file, as sales are listed under the different sections as explained in the paragraph above. If errors occur, they are a result of human errors, and do not pose a concern for product quality.

Declaration errors - are more acceptable, although large units are few in number, they still exist. Errors are of different models, for example, the values are reported in LEK or thousand LEK (absence of three "0"); one or several months are not reported; reported in the third quarter or fourth quarter the values of the previous quarters; reported negative values or may be excessive declarations in the previous period and reimbursed in the following periods.

Codification - coding errors are acceptable since many times are derived from the moment of registration at the National Business Registration (NBR), where the economic activity code is not set correctly and the next one carries an error in the tax file. From the statistical point of view, the economic activity code established by the Tax Authority is not taken into account for the very problems that this placement entails.

3. Analysis of Suspicious Value Patterns in VAT File in Albania

For identifying the patterns two years are studied, 2016-2017, in the whole VAT dataset where 17 models of suspicious values have been identified (see Hoogland, 2009). The model analysis is conducted in two variants: (1) the models were built for each particular year and (2) for the two years together, and share of number of units with suspicious values and share of net sales indicator, as a ratio to sales annual value quarterly declarations of VAT units are analysed. Furthermore, some tests are done to understand the problems faced.

In the first case, the results show that 15.903 percent of VAT enterprises carry suspicious or wrong values, while net sales are identified as suspicious at 8.6 percent of cases in 2016, compared to 2017, the proportion of suspected value companies decreased to 9.0 percent, while errors in net sales have dropped to 1.6 percent (see Annex, Table 2). This means that the quality of VAT file is improved from year to year.

In the version (2), it was noted that the presence of suspicious values are reduced where in to 1.44 percent of units the suspicious values are present, and for the respective years, sales of suspicious values amounted to 2.57 percent (2016) and 3.63 percent (2017). This analysis is valid when the data will be used to perform short-term analysis on the panel data set based on VAT file.

The first 1-4 patterns (Anex, Table 2) have a "0" declaration in three different quarters. For models where the value of net sales in three different quarters is "0" regardless of the meaning of "0", error statement or passive enterprises, the enterprise performance could not be understood. To see the performance of the switchboard and understand the contents of the "0" value, the units were continuously analyzed for two years. The editing methods are difficult to choose as the economic activity has to take into consideration. Another consideration is the changes on declaration rules by tax authorities for specific activities. From the analysis it was noted that the number of enterprises with the same declaration model was reduced in number, while the economic activity analysis was difficult to do so.

The presence of null value in NACE code or wrong code complicates the analysis of different patterns by economic activity. However, based on the nature of the economic activities, it can be judged that during the construction activity there are periods of breaks where net sales are "0", or units declare at the end of the year after the completion of construction works. In order to see if their declaration continues in the following years in this way, then the inclusion process will need to be included in the quarterly surveys.

This patterns were identifyed as well in the Accomodation activity, where model 2 (0, 0, x, 0) is most acceptable in these units in the coastal areas. In these cases it is difficult to judge whether the statement is wrong, when the activity is seasonal.

In patterns 5-8 where (Anex, Table 2) the presence of "0" value is only in one quarter within the year. The analysis consist in the presence of wrong value or double declaration in the continuing quarter in case were higher than previous quarter with '0" value. The edit method is a simple one by imputation method. The problematic cases here come from the extreme value (highest or lowest) that gives the highest/lowest impact in the growth rate. One of the reasons for highest value identified could be the multiplication by 1000 ALL (from 1500 thousand ALL it is declared 1 500 000 ALL). The lowest values identifyed were for example "1" or "3" where the human errors could be present.

Another analysis was done combining models with each other on the built-in two-year panel. One of the combinations was {(x, 0, 0, 0) (0, 0, 0, x)}. Such a combination was difficult to understand first, but when the analysis focused on economic activity that unit helped to understand the behaviour of the enterprise. For example: this pattern could be present to the large unit in the Wholesale or retail trade of vehicles or in Project design activity in construction sector. These values are to be used yearly in annual structure survey.

Other models from 9 to 13 (Anex, Table 2) are units that contain suspicious values that relate not only to declarations but also to the economic activity where the unit is part of, within and between quarters. The other issue that could be considerd is the number of employees. By combining the VAT data with Pay roll records by unique ID of enterprises we could identify the VAT units where the increased number of employees in related periods has given impact in the increase of net sales value. Linear regression between net sales value and number of employees is tested. The best editing method is the calculation of the interquartile distance of median, quartile 1 and 3 as explained in Hoogland and Van Haren (2007) by two dimensions: economic activity and size of enterprise in the high or low suspicious value. Correction of these values will be made for units that have a high impact in total net sales and affect the final result for annual surveys and the quarterly growth rate. This set of correction models will be made for quarters within the year.

The patterns from 14 to 17 (Anex, Table 2), show the presence of negative values. It was noted that negative value VAT units were only in one year, and in the second year they had positive values, which means that their impact, especially in the case of significant values and the weight in calculating the quarterly growth rate was not accurate, in particular branches of the economy. As the reason for being negative could be the high declaration in the previous period (t-1), these values will be corrected by transforming them to the positive value. The main purpose of the paper was identification of suspicious or wrong values and proposal for editing methods based on selective editing principles and methods used by best experiences of the countries.

4. Conclusions

Using the VAT file is a necessity in nowadays challenges when the demand for accurate and rapid statistics is dominant. After the identification of the type of patterns the further study on the implementation of the right editing methods by using VAT file in order to be used as replacement of partially or fully quarterly survey or evaluate the results of the annual survey is needed. In correcting values with significant records, the chosen method is well-studied so that values that do not influence the calculation of the final result or the growth rate are not removed, as while the growth rate is calculated continuously. The suspicious influencing values are corrected in testing phase. The use of median and quartiles 1 and 3 distances, according to economic activity and size are one of the methods that can be used to correct the highest values, where the "0" value will be exluded from the file. The test done with limited records where all patterns were present helps us to identify for each group of patterns the right editing method to be used.

As a conclusion based on patterns the editing methods are proposed to be used. However, to eliminate considerably the suspicious values in the VAT file based on right judgments and to reduce to a minimum the errors it is necessary to create a common database with variables from Statistical Business Register, VAT file and Pay roll records to analyse the suspicious or wrong values. The

script for each pattern identified has to be created to run in automatic way. This procces will reduce the time for processing VAT file for quarterly enterprise survey or structure business survey purposes and will increase the quality of each statistical procces resulted at the end of producing of the qualitative statistical procces or product by using effectively the VAT file.

References

De Jong, A. (2003). IMPECT: recent developments in harmonized processing and selective editing. *Proceedings of UNECE Work Session on Statistical Data Editing, Madrid*.

Granquist, L. (1991). "Macro-editing–A review of some methods for rationalizing the editing of survey data". *Statistical Journal of the United Nations Economic Commission for Europe*, *8*(2), 137-154.

Groen, J. A. (2012). "Sources of Error in Survey and Administrative Data: The Importance of Reporting Procedures". *Journal of Official Statistics (JOS)*, *28*(2).

Hidiroglou, M. A., & Berthelot, J. M. (1986). Statistical editing and imputation for periodic business surveys. *Survey methodology*, *12*(1), 73-83.

Hoogland, J., & van Haren, G. (2007, June). Editing and integrating VAT and SBS data. In *Proceedings of the third International Conference on Establishment Surveys (ICES-III), Montreal*

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*.

Hoogland, J., Van Bemmel, K. and De Wolf, P-P. "Detection of potential influential errors in VAT turnover data used for short term statistics", *Proceedings of UNECE Work Session on Statistical Data Editing*, *Neuchatel, October 2009*.

Hoogland, J. (2010, March). Editing Strategies for VAT Data. In *Seminar on Using Administrative Data in the Production of Business Statistics-Member States Experiences, Rome* (pp. 16-18).

Hoglund, D. E. (1988), "A Method for Editing Periodical Data". *Master of Science Thesis, University of Stockholm.*

Hughes, P. J., Linacre, S. J., & McDermid, I. M. (1990). *The use of graphical methods in editing*. Australian Bureau of Statistics.

Hooper, E., Lewis, D., & Dobbins, C. (2010). The application of selective editing to the ONS Monthly Business Survey. In *Q2010 Conference*.

Lewis, D., and Finselbach, H. (2011). Editing and imputation in the context of data integration from multiple sources and mixed modes. *Conference of European Statisticians, Oslo.*

Lewis, D., and Finselbach, H. (2012). *"*Editing and Imputing VAT Data for the Purpose of Producing Mixed-Source Turnover Estimates*". Proceeding of UNECE Conference of European Statistics, Work Session on Statistical Data Editing (Oslo, Norway, 24-26 September 2012.pp.2-5*

Lewis.D and Lewis.P, (2011), "Finding and resolving Data Quality Issues in Value Added Tax Data". *Final Report for WP2 of ESSNet on Admin Data*.

Seyb, A., Stewart, J., Chiang, G., Tinkler, I., Kupferman, L., Cox, V. and Allan, D. (2009), "Automated editing and imputation system for administrative financial data in New Zealand". *Proceedings of UNECE Work Session on Statistical Data Editing*, *Neuchatel, October 2009*.

Orchard, C., Moore, K., & Langford, A. (2010). Practices for using VAT turnover data within the UK to produce estimates of growth and monthly turnover. *Office for National Statistics, Newport*.

Vlag.P., (2012). "Imputing Missing Values When Using Administrative Data for Short-Term Enterprise Statistics". *Proceeding of UNECE, Conference Of European Statisticians Work Session on Statistical Data Editing, Oslo Norway*.

E-ISSN 2281-4612
ISSN 2281-3993

Academic Journal of
Interdisciplinary Studies

Vol 8 No 1
March 2019

Appendix

Table 1: List of variables required from Tax authority

1. NIPT – Business identification code
2. Period – 20-25 days after the reference month
3. Year – calendar year referring VAT data
4. Economic code 1- NACE Rev.2 code of primary activity
5. Economic code 2- NACE Rev.2 code of secondary activity
6. Description 1 - Description of primary activity
7. Description 2 - Description of secondary activity
8. Subject – Name of VAT unit
9. Exluded sale – Sales exluded from tax regime
10. Sales without VAT – Sales where VAT tax is not used
11. Export of goods – Value of export of goods
12. Supply with 0% scale – Selfservice with 0 % cost
13. Sales with 20 % regime – Value of sales where regime is applied in 20 % of the value
14. Sales with 6 % regime – Value of sales where regime is applied in 6 % of the value
15. Sales of travel agent with 20% - Value of sales where regime is applied in 20 %
Net sales - Calculated variable by summing up of 9 to 15

Table 2: Patterns used for identification of suspicious value in VAT

No of patterns	Patterns	Value	2016		2017	
			Share of unit	Share of Yearly Net Sales	Share of unit	Share of Yearly Net Sales
1	0,0,0,x	x≠0	3.84	5.18	1.08	0.09
2	0,0,x,0	x≠0	1.26	0.05	1.06	0.04
3	0,x,0,0	x≠0	0.64	0.04	0.55	0.02
4	x,0,0,0	x≠0	0.91	0.04	1.35	0.12
5	w,x,y,0	x,y,z≠0	1.06	0.19	1.78	0.42
6	w,x,0,z	x,y,z≠0	1.04	0.17	1.04	0.27
7	w,0,y,z	x,y,z≠0	0.65	0.16	1.04	0.13
8	0,x,y,z	x,y,z≠0	6.13	2.75	0.69	0.50
9	x,x,x,x	x≠0	0.09	0.00	0.13	0.01
10	x,x,x,y	x≠0,y≠0	0.08	0.01	0.08	0.01
11	x,x,y,x	x≠0,y≠0	0.02	0.00	0.02	0.00
12	x,y,x,x	x≠0,y≠0	0.02	0.00	0.01	0.00
13	y,x,x,x	x≠0,y≠0	0.06	0.02	0.07	0.00
14	w,x,y,z	z<0	0.03	0.00	0.03	-0.02
15	w,x,y,z	y<0	0.03	0.00	0.02	-0.01
16	w,x,y,z	x<0	0.03	0.00	0.03	0.00
17	w,x,y,z	w<0	0.01	0.00	0.03	0.00

Table 3 Patterns used for identification of suspicious value in VAT, 2016&2017

No of patterns	Patterns	Value	Share of Unit	Share of yearly Net sales	
				2016	2017
1	0,0,0,x	x≠0	0.18	0.020	0.018
2	0,0,x,0	x≠0	0.51	0.003	0.008
3	0,x,0,0	x≠0	0.02	0.001	0.001
4	x,0,0,0	x≠0	0.06	0.002	0.001
5	w,x,y,0	x,y,z≠0	0.03	0.002	0.003
6	w,x,0,z	x,y,z≠0	0.08	0.014	0.009
7	w,0,y,z	x,y,z≠0	0.02	0.002	0.003
8	0,x,y,z	x,y,z≠0	0.24	0.027	0.045
9	x,x,x,x	x≠0	0.04	0.001	0.002
10	x,x,x,y	x≠0,y≠0	0.00	0.000	0.000
11	x,x,y,x	x≠0,y≠0	0.00	0.000	0.000
12	x,y,x,x	x≠0,y≠0	0.00	2.500	3.600
13	y,x,x,x	x≠0,y≠0	0.06	0.001	0.001
14	w,x,y,z	z<0	0.07	0.002	-0.015
15	w,x,y,z	y<0	0.05	0.001	-0.013
16	w,x,y,z	x<0	0.05	-0.001	0.000
17	w,x,y,z	w<0	0.04	0.000	0.000