

XBRL UTILIZATION AS AN AUTOMATED INDUSTRY ANALYSIS

ALEX SUTA*¹ AND ÁRPÁD TÓTH¹

¹Research Center of Vehicle Industry, Széchenyi István University, Egyetem tér 1, Győr, 9026, HUNGARY

In the last two decades, electronic financial reporting went through a significant evolution, where to date, eXtensible Business Reporting Language (XBRL) has become the leading platform that is already obligatory for listed entities in the United States and was also legislated in the European Union from January 1, 2020. The primary objective of this research was to review the US-listed companies' 2018 quarterly reports. The study generated an automated industry analysis for the automotive industry from the aspect of four main financial item categories as an alternative to statistics-based, manually prepared industry analyses. Statistical tests were carried out between two industrial classification methodologies, the securities' industry identification marks and the reported Standard Industrial Classification (SIC) codes. The results showed a significant difference between the industry classification methodologies. Automated reporting was more precise with regard to the identification of the listed and reporting entities, however, the data fields of SIC codes within the XBRL data set provided an inaccurate classification, which is a potential area of improvement along with additional recommendations outlined in the Conclusion.

Keywords: XBRL, US-listed entities, ACL, Automated data analytics, Industry analysis, SIC codes

1. Introduction

The electronic reporting and automated fundamental reviews in the field of financial reporting is becoming increasingly important considering the difficulties and error-prone procedure of manual analysis from the available source of information. The eXtensible Business Reporting Language (XBRL) provides a standardized platform for this activity, which supports automated and digitalized reviews compared to the paper-based reports from the previous manual. This electronic reporting platform is already used as the official reporting form in the United States for listed entities, therefore, the application of a proper industry classification is essential. Even though XBRL reporting is required by the U.S. Securities and Exchange Commission (SEC), research institutions can choose from various generally accepted industry classifications. Despite the lack of regulation, it is a primary interest of research institutions to protect their reputations by adequately representing companies from the various industries. The two different approaches might provide different results, which can lead to inaccurate trend projections or unreliable industry comparisons. The validation of the XBRL classification and reports by marketing research firms can only be reconciled and validated to statistical industry reports which identify discrepancies. To date, Standard Industrial Classification (SIC) codes are used in the SEC's Electronic Data Gathering, Analysis,

and Retrieval (EDGAR) system to define the type of business of companies. Based on its primary activity, each company assigns a four-digit code to itself when registering an Initial Public Offering (IPO) with the SEC [1]. The four digits indicate levels of description of the industry classification, e.g. the location hierarchy for car manufacturers is Division D - Manufacturing (codes 20-39), code 37: Transportation equipment, and code 3711: Motor vehicles and Passenger Car Bodies [1]. The objective of this research was to review the US-listed companies, where XBRL reports are already required and implemented. Subsequently, through an automated review of the automotive industry, to then identify how these reports can be compared to the European listed entities. This information is crucial for stakeholders and regional policymakers to gain a clear view of the conditions of the target industry. According to the European Securities and Markets Authority (ESMA), from January 1, 2020 onwards, new requirements on the stock exchange-listed companies in the European Union came into effect to provide respective financial statements in a new European Single Electronic Format (ESEF). This is a significant change to the application of XBRL as companies now have to provide reports in this specific reporting language. The data sets include structured information; for this reason, a new wave of research initiatives is expected in this academic area that could follow on from inconsistent industry classifications, further hindering comparability.

*Correspondence: suta.alex@ga.sze.hu

2. Literature review

2.1 XBRL utilization in industry-specific data analysis

Prior literature has documented uses of XBRL in a variety of data analysis environments, generally in the research areas of accounting and financial reporting. Systematic financial data provides data analysts and investors with the ability to measure performance and risks, as well as create comparisons, ratings and other value-added products [2]. Connected to comparability aspects, several sources have been reviewed that are related to the semantic issue of industrial classification. Being a driver of electronic data interchange, XBRL data sets are constructed from multiple identifying tags and numerical data that can be processed by computer software [3]. While the technical background on data-centric analysis is available [4] 2013 [5] 2014, [6], it is uncommon in the industry-specific research literature that XBRL databases are used as the primary source of data.

Chychyła-Leone-Meza [7] measured financial reporting complexity by comparing the quantity of text in US Generally Accepted Accounting Principles (GAAP) and SEC regulations of textual data from XBRL filings. In this study, the variation with regard to the data content of different taxonomy versions (denominative tags, labels, documentation) is emphasized. For this reason, the annual changes in published taxonomy updates have to be taken into consideration [8]. Despite the existence of the XBRL Industry Resource Group established by the FASB [8], Standard Industrial Classification (SIC) codes are not part of taxonomy updates and their current 2007 form seems to be generally accepted for statistical use. Felo-Kim-Lim [9] observed changes in the information environment of analysts by the overuse of customized tags, creating assumptions based on industrial classification as a factor. Zhang-Guan-Kim [10] proposed an expected investor crash risk model based on financial information gathered from XBRL-based SEC databases. In terms of the estimation of the impacts, the industry median of customized tags is generated by 2-digit SIC codes as an adjustment tool in the regression model. Similarly, industrial classification was taken into account as a dummy variable during the analysis with regard to the XBRL adoption of reductions in audit fees as per Shan-Troshani-Richardson [11].

In other XBRL-based studies, industry-specific assumptions required solutions other than basic SIC codes. Liu-Luo-Wang [12] reviewed the effect of XBRL adoption on information asymmetry, where SIC was reclassified to identify high-technology industries.

2.2 Discrepancies between industrial classification systems

Since the emergence of the North American Industry Classification System (NAICS) in 1997 as a sound replacement of Standard Industrial Classification (SIC)

codes in U.S. industrial statistics, papers have reviewed the impacts of different frameworks in financial research. Effective comparative statistics require the use of a standardized classification system [13]. The U.S. Economic Census Bureau has made regulatory, business and academic purposes of performing economic research on historical data possible. In 1997, the existing framework, the SIC, was replaced by the NAICS [14]. Unlike the SIC's mixed production/market system, NAICS introduced a production-oriented economic concept that supports the examination of industry-specific indicators such as productivity, input-output relationships and capital intensity [15]. The specific rearrangements between industrial classes primarily affected manufacturing industries, where the SIC functions as a somewhat outdated alternative. U.S. government departments, namely the Bureau of Labor Statistics (BLS), Internal Revenue Service (IRS) and Social Security Administration (SSA), alongside the U.S. Securities and Exchange Commission (SEC), continue to use the most recent 2007 update of four-digit SIC codes. While maintaining a unified classification system is necessary for government departments, the lack of conceptual harmony between industrial classification systems creates a discrepancy with academic research [16]. Several papers have been collected that present empirical evidence of disharmonious schemes based on Financial Statement Data Sets. Kahle-Walking [17] observed differences in financial variables gathered from two statistical databases (CRSP and Compustat) using four-digit SIC codes to be substantial, moreover, showed that commonly used methods of industrial classification disagree due to frequent changes in the SIC codes of firms. Bhojraj-Lee-Oler reviewed the capital market applications of four broadly available industrial classification schemes and found that a significant degree of variance with regard to the number of companies represented in industry divisions exists. The study argues that the six-digit Global Industry Classification Standard (GICS), followed by NAICS, offers better comparability between firms concerning SIC in terms of the critical evaluation of financial ratios and that industrial classification is essential in instances of fundamental analysis. While GICS reflects the dynamic changes in industry sectors, being a privately available system mainly involved in investment processes, it is unlikely to be suitable in statistical research [18]. Kelton-Pasquale-Rebelein [19] referred to SIC codes as outdated in the field of industry cluster analysis and prepared an updated framework using NAICS. As opposed to classifying establishments according to similar products (SIC), the groups are formed from identical production processes (NAICS). Hrazdil-Zhang [20] and Hrazdil-Trottier-Zhang [21] published empirical results on the heterogeneity of industry concentration with the use of SIC and other classification schemes based on the market shares of sales and financial ratios of companies in the manufacturing sector (SIC 2000-3999). According to their findings, the SIC system remains inferior to GICS and NAICS in terms of industrial homogeneity.

Instead of ordinary company databases such as CompuStat or S&P 1500, Papagiannidis et al. proposed an exploratory big data method to gather regional research of industry clusters based in the UK. In this study, keywords connected to business operations were collected from official websites to enhance the level of detail provided by single SIC codes, supporting the formation of regional clusters. It is a common conclusion in the reviewed literature that the sole use of SIC codes in industry analysis could lead to the loss of information and false estimation of market forces; in this context the potential of XBRL as a primary data source of financial statements has been reviewed.

2.3 The multi-tier supply chain approach

One possible outcome of the barriers of traditional statistical classification systems is the addition of extra information to existing schemes. In an industrial analysis, especially in the automotive industry, it is essential to differentiate between operational properties, e.g. their position in the automotive supply chain. The contemporary position of an industry must be judged by the different weights of its market players. Assumptions about financial information are heavily affected by the final product, whether it is a part of the interorganizational supply chain, or sold to dealerships or directly to consumers in the form of passenger cars. In terms of a supply chain, manufacturers and suppliers can be classified into multi-tiered groups based on their position in the production chain, as well as the state of raw materials (tier 3 and additional sub-tiers) in addition to finished or semi-finished components (suppliers from tiers 1 and 2) compared to fully finished products (Original Equipment Manufacturers (OEMs)). Concerning the automotive industry, sources from both academia, business and governments [22–24] agree that market players from multi-tier supply chain structures can be ranked as follows:

1. OEMs: a concentrated group of companies accountable for the main manufacturing, assembly and design processes that possess a large market share and well-known brand names;
2. Suppliers from Tiers 1 & 2: potentially several hundred large or small companies, accountable for the supply of automotive parts and systems to OEMs. The range of sold goods is diverse and includes engine components, interior, exterior, transmission as well as cooling and electronic systems. Although their role in the supply chain is consistent, suppliers from Tiers 1 and 2 vary in their direct/indirect (through other participants) nature of interaction with OEMs, therefore, from a statistical viewpoint, can be aggregated;
3. Tier 3 and sub-tier suppliers: several thousand smaller companies are accountable for the supply of raw materials to suppliers from Tiers 1 & 2.

In the scientific literature, several utilizations of the multi-tier supply chain approach exist. Mena-Humphries-Choi [25] reviewed the existing literature at the time on structural arrangements (buyer-supplier-customer) and prepared three cases of theoretical linkage. According to the study, the most typical structure of the automotive industry is the “closed triad”, where the buyer (OEM) can insist on certain requirements (either assurance or training function) not only from Tier 1 but sub-tier suppliers as well. Masoud-Mason [26] used the multi-tier system in the automotive industry to simulate cost optimization on a supply-chain level. Thomé et al. [27] adopted a similar approach of representing many tiers and their interactions that affect selected flexibility measures (product, responsiveness, sourcing, delivery and postponement). Other popular fields of use are sustainability-related questions and green supply chains [28–30].

The available literature clarifies the widespread applicability and general acceptance of tiered levels of suppliers, which supports the methodology examined in the current study. Despite its academic use, the application of the well-established OEM / tiered system of suppliers in automotive business reports published by major consulting firms [31–33] is common practice.

3. Data collection and methods used

The SEC has published XBRL data sets containing raw aggregate financial statement data quarterly since 2009. At the same time, as a premium service, the SEC offers a professional version of its search engine [34] designed specifically to fit the goals of professional financial analytics. However, in line with tendencies identified from the literature review, even a discrepancy on the same platform exists between the Standard Industrial Classification codes current in XBRL data sets and the EDGAR search tool. To perform an automated industry analysis, a suitable classification is required. In this study, a possible classification using the software program ACL (Audit Command Language) Robotics Professional version 14.1.0.1581 is evaluated. From the listed U.S. entities, those operating in the automotive industry were selected to measure deviance in terms of crucial financial indicators between the two data sources. The specific choice of the automotive industry lies in its accurate definability, while the goal of the study was to provide an industry-independent methodology of data analysis that can be applied to several other fields. The two main platforms of data collection were EDGAR Pro Online (2019) operated by the SEC, which is equivalent to the quasi-manual download process of financial statements, and the obligatory quarterly reports of aggregate data sets in the XBRL format available on the SEC website. To avoid existing industrial classification issues, a multi-tier supply chain approach was introduced by grouping companies as OEMs and suppliers from Tiers 1 & 2 (T1&2 S).

4. Results

4.1 Data categorization: number of companies and industries

By using the EDGAR Pro Online search tool, market segments can be filtered, of which three categories connected to the automotive industry are available. At the same time, in the XBRL data set, companies are provided with much general information, including SIC codes that can be used for categorization. According to the list of codes provided by the SEC, six four-digit codes cover the automotive industry (and related services with the exception of retail) that were reviewed in the quarterly reports of 2018. A summary of publicly listed entities is presented in Table 1.

All entities listed on the New York Stock Exchange (NYSE), National Association for Securities Dealers Automated Quotations (NASDAQ) and Better Alternative Trading System (BATS) from the entire population are supposedly consistent data sources and regulated by the SEC. In addition to the variance in the number of listed entities in the automotive industry, the size of the entire population between the two sources is inconsistent and differs by over 24%. In terms of industrial classification, the taxonomy behind SIC codes in XBRL data sets is valid but incomparable to the customary EDGAR approach in the case of the identification of specific activities. Therefore, two additional categories were created to fit the measurement process; OEMs and suppliers from Tiers 1 & 2 (other automotive suppliers).

4.2 Errors in terms of the consistency and availability of samples

Listed entities from both data sources that are unmatched as a result of their supposedly consistent counterparts were found. Out of the sample sizes of 103 and 74, 50 companies are common in both which raises concerns over reliability.

Furthermore, data availability raised concerns in terms of search results from the SEC EDGAR Pro Online system. Out of the strong sample size of 103, 13 annual reports concerning 2018 were unavailable in the electric filing system of the SEC, while an additional 10 required data collection from official websites. Four financial statement items concerning the wealth and profitability of companies were selected for analysis in order to evaluate the differences between the two industrial classification schemes. The values of Total assets, Total Equity, Net sales revenue and Profit after-tax are central financial factors of investor decision-making. When necessary, exchange rates of the Federal Reserve were used according to the ASC (Accounting Standards Codification) standards issued by the Financial Accounting Standards Board (FASB) [35].

4.3 Comparison of financial information on an industrial level

Based on the financial statement data, descriptive statistics were calculated on the selected reporting lines. Differences were summarized in terms of both absolute values between the two data sources and percent deviations as presented in the Tables 2 and 3. A general observation of the data source is that the intervals between the minimum and maximum values are substantial for all four financial statement items. It is likely that – when used as a statistical sample – a normal distribution cannot be assumed. The standard deviation exceeds the mean values in the case of Total assets, therefore, the set of values (especially for the financial data of suppliers from Tiers 1 & 2) is highly dispersed.

A pattern can be observed in the deviation between the two data sources. The total values of OEM financial statement items are higher in the XBRL data set, in contrast to data derived from the online SEC source, while the opposite is seen in the case of suppliers from Tiers 1 & 2, where the total values are dominated by online

Table 1: Industrial specification of data sources

SEC's EDGAR Online Pro		SEC's EDGAR XBRL data set		
Industrial specification	Number of companies	Industry (SIC)	Number of companies	
OEMs	{ Auto & Truck Manufacturers	26	3711. Motor Vehicles & Passenger Car Bodies	20
	{ Automobiles, Parts & Service Retailers	24	3713. Truck & Bus Bodies	2
Suppliers from Tiers 1 & 2	{ Auto, Truck & Motorcycle Parts	53	3714. Motor Vehicle Parts & Accessories	41
			3715. Truck Trailers	1
			3716. Motor Homes	2
			3751. Motorcycles, Bicycles, and Parts	8
Total	103	Total	74	
Entire population size	5,736	Entire population size	7,133	

sources. As an attempt to generalize the automotive industry, mean values were calculated where XBRL represents higher values except for the net sales revenues of suppliers. These deviations are partly validated by the amount of incompletely matched samples, but the 103:74 sample-size ratio is not represented by the results. The Table 4 summarizes the difference between the results of descriptive statistics in the form of percentages.

Despite former expectations, OEMs do not represent the majority of the financial item totals (between 45.1 and 55.9%), total equity (between 31.8 and 48.6%), net sales revenue (44.7 and 57.6%) and profit after-tax (35.4-52.5%), the differences between data sources can be measured on a scale of 6.7% to 31.9% as seen in Table 4. Suppliers from Tiers 1 & 2 match to an even lesser extent, so percent deviations are typically higher, especially in the case of net sales revenue (56.7%). Based on the matrix, the individual averages of companies cannot be used for industry generalization, both in terms of absolute mean values and standard deviations. The deviation “hotspots” are clearly centered around the suppliers from Tiers 1 & 2.

4.4 Chi-square statistical testing

To support our assumptions of statistically significant deviation between data sources, Pearson’s chi-squared test was implemented, a full description of the steps is available in Appendix A [36, 37]. Selected categories of OEMs and suppliers from Tiers 1 & 2 were differentiated along with expected (data derived from online SEC-based financial statements) vs. observed (data derived from XBRL data sets) values. Based on the per-

formed Chi-square test, the results highlighted that the differences between the expected and observed values of financial statement items (Total assets, Total Equity, Net sales revenue and Profit after-tax) were significant. With a 95% confidence interval ($\alpha = 0.05$), OEMs and suppliers from Tiers 1 & 2 both exceeded the critical value of 16.92 with 7 degrees of freedom ($df = 7$). It is important to note the very significant (almost 10 times higher) impact of suppliers from Tiers 1 & 2 in terms of the total level of deviance.

5. Conclusions

XBRL preparation is obligatory, however, the content can include differences from the reported and published financial statements. Conclusions can be summarized in the following points:

- Potential duplication of lines in XBRL sources (e.g. 8 lines of certain financial statement items from China Automotive Systems, Inc.);
- Lack of standardization in tags: The XBRL platform manages to integrate more financial reporting taxonomy (different annual versions of IFRS and US GAAP). Due to the different (and potentially customized) tags, the definitions of some financial statement items converge; the structure of financial statements has yet to be fully harmonized between annual reports and XBRL statements;
- Errors in the reporting period (temporal differences): in some cases, outdated (1 or 2 years prior to

Table 2: Automotive market share of the entire population (% , number)

Representation % (n)	SEC’s EDGAR Pro Online	SEC’s EDGAR XBRL data set	
Entire population	100% (5,736)	100% (7,113)	
OEMs	0.45% (26)	0.31% (22)	Total assets Total equity Net sales revenue Profit after-tax
Suppliers from Tiers 1 & 2	1.34% (77)	0.73% (52)	
Total	1.80% (103)	1.04% (74)	
OEMs (matched)	0.28% (16)	0.20% (14)	
Suppliers from Tiers 1 & 2 (matched)	0.59% (34)	0.50% (36)	
Total (matched)	0.87% (50)	0.70% (50)	

Table 3: Absolute financial data from data sources (left – SEC’s EDGAR Pro Online / right – SEC’s XBRL data set) (USD in millions)

Statistics	Category	Total assets value		Total equity attributable to company owners		Net sales revenues		Profit after taxes	
Totals	OEM	↓ 783.346	↑ 874.139	↓ 153.668	↑ 221.111	↓ 606.225	↑ 649.422	↓ 23.941	↑ 35.125
	Tier 1&2 S.	↑ 953.149	↓ 688.812	↑ 329.224	↓ 234.365	↑ 749.694	↓ 478.399	↑ 43.717	↓ 31.834
Mean	OEM	↓ 27.012	↑ 39.734	↓ 5.299	↑ 10.050	↓ 20.904	↑ 29.519	↓ 826	↑ 1.597
	Tier 1&2 S.	↓ 12.880	↑ 13.246	↓ 4.449	↑ 4.507	↑ 10.131	↓ 9.200	↓ 591	↑ 612
St. dev	OEM	↓ 66.884	↑ 77.251	↓ 11.802	↑ 18.956	↓ 45.992	↑ 52.970	↓ 1.875	↑ 3.007
	Tier 1&2 S.	↓ 56.749	↑ 63.640	↓ 21.527	↑ 23.684	↓ 33.241	↑ 34.997	↓ 2.971	↑ 3.262

Table 4: Deviation matrix between data sources (%)

Statistic	Category	Total assets value	Total equity attributable to company owners	Net sales revenues	Profit after taxes
Totals	OEM	10.39%	30.50%	6.65%	31.84%
	T1&2 S.	38.38%	40.48%	56.71%	37.33%
Mean	OEM	32.02%	47.28%	29.18%	48.29%
	T1&2 S.	2.76%	1.29%	10.12%	3.50%
St.dev	OEM	13.42%	37.74%	13.18%	37.65%
	T1&2 S.	10.83%	9.11%	5.02%	8.93%

the current fiscal year) financial information is presented in current filings (e.g. an entity presents information from the 2017 fiscal year in the Q4 2018 filing as the most current);

- The inability to fully and feasibly automate data analysis in the case of automotive suppliers. Mean values are inconsistent between data sources due to the varying sample size of automotive suppliers. To perform a comprehensive industry analysis, error terms need to be defined clearly. Otherwise such an analysis would be performed with many predefined assumptions, leading to a decrease in the overall explanatory power and raising concerns about reliability/reproducibility.

Financial analysts should use XBRL datasets with concern, these points kept in mind. As a currently available best practice, the methodology of the U.S. Securities and Exchange Commission is a precedent for the building of inline XBRL statements into integrated datasets. An emerging challenge of regulatory bodies such as the European Securities and Markets Authorities is the supervision of companies uploading their data to a central system of a similar nature to produce well-structured databases for automated financial analytics.

Acknowledgements

The research presented in this paper was financed by the “Research Area Excellence Program – 2019 (TUDFO/51757/2019-ITM)” and University of Győr.

REFERENCES

- [1] NAICS Association (2019). Common SIC Questions, <https://www.naics.com/frequently-asked-questions/#NAICSfaq>
- [2] XBRL International (2019). An Introduction to XBRL, <https://www.xbrl.org/the-standard/what/an-introduction-to-xbrl/>
- [3] Yaghoobirafi, K.; Nazemi, E.: An Approach to XBRL Interoperability Based on Ant Colony Optimization Algorithm. *Knowledge-Based Systems*, 2019, **163**, 342–357 DOI: 10.1016/j.knosys.2018.08.038
- [4] Wenger, M. R.; Elam, R.; Williams, K. L.: A Tour of Five XBRL Tools: Products That Help Make Tagged Data Work for You and Your Clients. *Journal of Accountancy*, 2013, **215**(4), 48–55
- [5] Radzimski, M.; Sanchez-Cervantes, J. L.; Garcia-Crespo, A.; Temiño-Aguirre, I.: Intelligent Architecture for Comparative Analysis of Public Companies Using Semantics and XBRL Data. *International Journal of Software Engineering and Knowledge Engineering*, 2014, **24**(5), 801–823 DOI: 10.1142/S0218194014500314
- [6] Joyner, D.: How to Transform XBRL Data into Useful Information. *Review of Business and Technology Research*, 2017, **14**(1), 34–41
- [7] Chychyla, R.; Leone, A. J.; Minutti-Meza, M.: Complexity of Financial Reporting Standards and Accounting Expertise. *Journal of Accounting and Economics*, 2019, **67**(1), 226–253 DOI: 10.1016/j.jacceco.2018.09.005
- [8] SEC Reporting Taxonomy Technical Guide (version 2019), Financial Accounting Standards Board, Financial Accounting Foundation, 2019 https://www.fasb.org/jsp/FASB/Document_C/DocumentPage&cid=1176171806866
- [9] Felo, A. J.; Kim, J. W.; Lim, J. H.: Can XBRL Detailed Tagging of Footnotes Improve Financial Analysts’ Information Environment? *International Journal of Accounting Information Systems*, 2018, **28**, 45–58 DOI: 10.1016/S0740-624X(98)90003-X
- [10] Zhang, Y.; Guan, Y.; Kim, J. B.: XBRL Adoption and Expected Crash Risk. *Journal of Accounting and Public Policy*, 2019, **38**(1), 31–52 DOI: 10.1016/j.jaccpubpol.2019.01.003
- [11] Shan, Y. G.; Troshani, I.; Richardson, G.: An Empirical Comparison of the Effect of XBRL on Audit Fees in the US and Japan. *Journal of Contemporary Accounting & Economics*, 2015, **11**(2), 89–103 DOI: 10.1016/j.jcae.2015.01.001
- [12] Liu, C.; Luo, X. R.; Wang, F. L.: An Empirical Investigation on the Impact of XBRL Adoption on Information Asymmetry: Evidence from Europe. *Decision Support Systems*, 2017, **93**, 42–50 DOI: 10.1016/j.dss.2016.09.004
- [13] Kaya, D.; Pronobis, P.: The Benefits of Structured Data Across the Information Supply Chain: Initial Evidence on XBRL Adoption and Loan Contracting of Private Firms. *Journal of Accounting and Public Policy*, 2016, **35**(4), 417–436 DOI: 10.1016/j.jaccpubpol.2016.04.003

- [14] Comparing Economic Data, U.S. Economic Census Bureau, 2016 <https://www.census.gov/programs-surveys/economic-census/guidance/historical-data.html>
- [15] Chychyla, R.; Leone, A. J.; Minutti-Meza, M.: Complexity of Financial Reporting Standards and Accounting Expertise. *Journal of Accounting and Economics*, 2019, **67**(1), 226–253 DOI: 10.1016/j.jacceco.2018.09.005
- [16] Krishnan, J.; Press, E.: The North American Industry Classification System and its implications for accounting research. *Contemporary Accounting Research*, 2003, **20**(4), 685–717 DOI: 10.1506/N57L-0462-856V-7144
- [17] Kahle, K. M.; Walkling, R. A.: The Impact of Industry Classifications on Financial Research. *Journal of Financial and Quantitative Analysis*, 1996, **31**(3), 309–335
- [18] MSCI (2018). Global Industry Classification Standard (GICS®) <https://www.msci.com/gics>
- [19] Kelton, C. M.; Pasquale, M. K.; Rebelein, R. P.: Using the North American Industry Classification System (NAICS) to Identify National Industry Cluster Templates for Applied Regional Analysis. *Regional Studies*, 2008, **42**(3), 305–321 DOI: 10.1080/00343400701288316
- [20] Hrazdil, K.; Zhang, R.: The Importance of Industry Classification in Estimating Concentration Ratios. *Economics Letters*, 2012, **114**(2), 224–227 DOI: 10.1016/j.econlet.2011.10.001
- [21] Hrazdil, K.; Trottier, K.; Zhang, R.: A Comparison of Industry Classification Schemes: A Large Sample Study. *Economics Letters*, 2013, **118**(1), 77–80 DOI: 10.1016/j.econlet.2012.09.022
- [22] Tassej, G.; Brunnermeier, S. B.; Martin, S. A.: Interoperability Cost Analysis of the US Automotive Supply Chain. *Research Triangle Institute*, 1999, Report 7007-03
- [23] ILO (2005). Automotive Industry Trends Affecting Component Suppliers, Report for Discussion at the Tripartite Meeting on Employment, Social Dialogue, Rights at Work and Industrial Relations in Transport Equipment Manufacturing. International Labour Organization, pp. 26–42
- [24] Silver, D.: The Automotive Supply Chain, Explained. 2016
- [25] Mena, C.; Humphries, A.; Choi, T. Y.: Toward a Theory of Multi-Tier Supply Chain Management. *Journal of Supply Chain Management*, 2013, **49**(2), 58–77 DOI: 10.1111/jscm.12003
- [26] Masoud, S. A.; Mason, S. J.: Integrated Cost Optimization in a Two-Stage, Automotive Supply Chain. *Computers & Operations Research*, 2016, **67**, 1–11 DOI: 10.1016/j.cor.2015.08.012
- [27] Thomé, A. M. T.; Scavarda, L. F.; Pires, S. R.; Ceryno, P.; Klingebiel, K.: A Multi-Tier Study on Supply Chain Flexibility in the Automotive Industry. *International Journal of Production Economics*, 2014, **158**, 91–105 DOI: 10.1016/j.ijpe.2014.07.024
- [28] Jia, F.; Gong, Y.; Brown, S.: Multi-Tier Sustainable Supply Chain Management: The Role of Supply Chain Leadership. *International Journal of Production Economics*, 2019, **217**, 44–63 DOI: 10.1016/j.ijpe.2018.07.022
- [29] Sauer, P. C.; Seuring, S.: Extending the Reach of Multi-Tier Sustainable Supply Chain Management – Insights from Mineral Supply Chains. *International Journal of Production Economics*, 2019, **217**, 31–43 DOI: 10.1016/j.ijpe.2018.05.030
- [30] Sarkis, J.; Gonzalez, E. D. S.; Koh, S. L.: Effective Multi-Tier Supply Chain Management for Sustainability. *International Journal of Production Economics*, 2019, **217**, 1–10 DOI: 10.1016/j.ijpe.2019.09.014
- [31] The Future of the Automotive Value Chain - Supplier Industry Outlook 2025, Deloitte, 2017
- [32] Automotive News - North America top suppliers, Tenneco, 2019 <https://www.autonews.com/assets/PDF/CA116090622.PDF>
- [33] Global Automotive Supplier Study 2018, Roland Berger, 2017 <https://www.rolandberger.com/it/Publications/Global-Automotive-Supplier-Study-2018.html>
- [34] EDGAR Pro Online (2019) operated by Donnelley Financial, LLC, <https://pro.edgar-online.com/>
- [35] Foreign Exchange Rates - H.10, Federal Reserve, 2019. <https://www.federalreserve.gov/releases/H10/hist/>
- [36] Field, A.; Miles, J.; Field, Z.: Discovering Statistics Using R. Sage Publications, 2012, pp. 814–816 ISBN: 978-1-446-20046-9
- [37] Agresti, A.: An Introduction to Categorical Data Analysis. John Wiley & Sons. 2018, pp. 34–40 ISBN: 978-0-471-22618-5

Appendix A – Chi-square test steps

1) Contingency table

Financial statement items		OEMs	Suppliers from Tiers 1 & 2
Total assets	Expected	783,346	953,149
	Observed	874,139	688,812
Total equity attributable to the owners of companies	Expected	153,668	329,224
	Observed	221,111	234,365
Net sales revenue	Expected	606,225	749,694
	Observed	649,422	478,399
Profit after-tax	Expected	23,941	43,717
	Observed	35,125	31,834
Total		3,346,977	3,509,195

- 2) **H0** The financial values of the XBRL data source (observed values) are not significantly different from the values of the online SEC EDGAR Pro Online data source (expected values). Therefore, industry totals from the two sources are consistent.

3) Calculated marginal totals for the observed table

4) Expected value calculation - based on the specific financial statement item's proportion in the whole population

5) Degree of freedom

$$df = (r - 1)(c - 1)$$

$$df = (8 - 1)(2 - 1) = 7$$

6) Calculation of chi-square values

Financial statement items	$\chi^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i}$	OEMs	Suppliers from Tiers 1 & 2	
Total assets	E-O	-90,792	264,337	
	(E-O) ²	8,243,257,799	69,874,243,884	
Total equity attributable to the owners of companies	E-O	-67,442	94,860	
	(E-O) ²	4,548,480,637	8,998,332,516	
Net sales revenue	E-O	-43,197	271,295	
	(E-O) ²	1,865,960,356	73,600,784,685	
Profit after-tax	E-O	-11,184	11,883	
	(E-O) ²	125,073,800	141,206,159	
Chi-square		4,417	43,490	Chi-square total
				47,907

7) Determination of Significance level and Critical value:

significance level (alpha)	0.05
critical value	16.92

- 8) The chi square test result showed as the H0 hypothesis should be rejected with a 95% confidence interval (degree of freedom = 7).

The Chi-squared test results showed that the hypothesis H0 should be rejected.