Comparative Statistical Analysis of Supplier Performance for a Retailer

Esra Agca Aktunc
Assistant Professor of Industrial Engineering
Faculty of Engineering and Natural Sciences
Kadir Has University
Istanbul, Turkey
esra.agca@khas.edu.tr

Abstract

The critical role of supply chain management for the well-being of communities all around the world has been highlighted as a result of a multitude of disruptions, including natural or man-made disasters, pandemics, and geopolitical uncertainties, in recent years. Supply Chain Management (SCM) coordinates the flow of material, information, and services across functions and across companies. Supply chain disruptions have become more frequent, and the external and internal risks on supply chains should be considered more seriously to prepare for such shocks better. Supplier performance measurement is one of the most crucial processes to identify risks by quantifying the efficiency and effectiveness of supplier actions. In this study, the performances of three critical suppliers of a footwear retail company are statistically analyzed to identify the risks and to propose actions for improvement. Normality tests are applied to the supplier data to determine the appropriate hypothesis testing procedures. Due to the lack of normality in the data, Kruskal-Wallis test is performed to compare the performances of the three suppliers in terms of order delivery delay time measured in days and the number of split deliveries. The statistical analysis reveals important risks regarding over-reliance on suppliers with low service performance levels.

Keywords
Supplier performance measurement, Supply chain, Statistical analysis, Kruskal-Wallis test, Non-parametric hypothesis testing

1. Introduction
All the goods and services reach their consumers through a supply chain, since the products people want or need are not necessarily produced where and when they want to consume them. A supply chain encompasses all activities associated with the flow and transformation of raw materials into finished products, transporting and distributing these products to the end user, as well as the associated information flows. Supply chain management (SCM) is “the systematic, strategic coordination of the traditional business functions and the tactics across these business functions within a particular company and across businesses within the supply chain, for the purposes of improving the long-term performance of the individual companies and the supply chain as a whole” (Mentzer et al. 2001). The fundamental objectives of SCM are delivering the right goods and services to the right consumers at the right time and place in the right quantity and condition to achieve competitive advantage and profitability (Ballou 2004). However, there are various uncertainties, risks, and several potential disruptions involving supply chain operations and making the right decisions about all supply activities is an extremely difficult task in this complex business environment. One of the crucial SCM decisions involves choosing reliable suppliers to improve customer service level and quality. In this study, the focus is on measuring and comparing the performance of the major suppliers of a footwear retail company. The demand and supply data for a group of footwear products is statistically analyzed and performance measures for suppliers are compared to provide managerial insights.

SCM has gained global interest in the recent years as a result of the disruptions affecting the whole world population such as the COVID-19 pandemic. In addition to the pandemic, natural disasters, climate change, and political instabilities have exposed and exacerbated the issues in supply chain management systems and this is forcing businesses and countries to rethink their strategic and operational decisions. There has been a significant shift to e-commerce as more consumers prefer online shopping for its convenience and safety. The share of e-commerce in total retail in Turkey rose from 11.87% in the first quarter to 17.4% in the second quarter of 2020, and there was a 51.8% increase in the share of e-commerce in the gross domestic product compared to 2019 (Republic of Turkey Ministry of Trade 2021). Similarly, the share of e-commerce in retail rose from 11.8% in the first quarter to 16.1% in the second quarter of 2020 in the US and from 20.3% to 31.3% in the UK (OECD 2020). The retailer that is the subject of this study has also entered e-commerce in Turkey and would like to gain a competitive edge in the marketplace by managing its supply processes more efficiently. Therefore, a systematic analysis of the historical supplier data is necessary to draw useful conclusions for the retailer and this study aims to provide a statistical analysis of supplier performance to the retailer with the fundamental aim of supporting and improving their decision-making processes regarding their supply chain.

1.1 Objectives

The objectives of this research study are as follows:

- Comparing service performance metrics of suppliers based on historical data
- Identifying significant service performance differences between suppliers
- Providing managerial insights and proposals for improvement of supply operations

2. Literature Review

Supplier performance management is defined as "the process of evaluating, measuring, and monitoring supplier performance and suppliers' business processes and practices for the purposes of reducing costs, mitigating risk, and driving continuous improvement" (Gordon 2008). Although it has been reported by some studies that soft and non-quantifiable or non-measured criteria can have a great impact on buyer-supplier relationships, developing and monitoring quantifiable measures is indispensable for supplier performance monitoring (Kannan and Tan 2002). As more companies are dependent on their suppliers with more activities, supplier performance measurement systems (PMSs) using a set of metrics to quantify the efficiency and effectiveness of suppliers' actions have gained importance (Hald and Ellegaard 2011).

Since the interests and priorities of a firm should be considered to achieve a strategic fit between a firm's business model and its supply chain strategy, Huang and Keskar (2007) proposed a mechanism where a set of metrics to measure supplier performance are selected among a comprehensive and configurable set of metrics based on a firm's business strategy and an optimal supplier selection decision is made based on this selected set of metrics. Chai et al. (2013) provide a systematic literature review of supplier selection methods and report that multicriteria decision making methods are, indeed, the most popular methods due to their effectiveness in ranking in the presence of multiple supplier performance metrics. Supplier performance measurement should be based on clear definitions of the metrics used and the information shared in order to ensure that all parties are working towards the same goals. As an investigation of the dynamics that may alter performance information, Hald and Ellegaard (2011) examined how the performance measurement information is shaped and reshaped between the evaluating buyers and the evaluated suppliers during the evaluation process and provided insights to optimize supplier performance.

Supplier PMSs are based on quantitative and qualitative performance metrics that need to be monitored over time. Supplier or vendor selection criteria have been discussed since the early 1960s and the most commonly used supplier performance measures include lead times, service level, prices, quality, reliability, satisfaction levels, and responsiveness (Weber 1991, Schmitz and Platts 2004, Waters 2011, Gutierrez et al. 2015). Luzzini et al. (2014) provide a framework covering critical choices a firm must make in vendor evaluation system design and implementation. This framework includes supplier rating system based on commercial, logistics, and quality indicators where reliability of times and quantities is included in logistical performance indicators. Alikani et al. (2019) incorporated sustainability and suppliers' risk factors into supplier selection using fuzzy sets to quantify inputs of decision makers and Data Envelopment Analysis. They showed that considering sustainability criteria or risk factors separately may result in unfit supplier rankings. Maestrini et al. (2017) provided a comprehensive review of supply chain PMSs and showed that although supplier PMSs receive a lot of attention, there are still several gaps to be addressed, such as the impact of supplier PMSs on performance, the causes of the supplier PMS adoption, and the consequences of a fit or misfit between the supplier and the customer perception about the supply chain PMS. Based on this review, most supplier PMSs include metrics on the quality of the product or service exchanged, delivery performance (e.g. punctuality, timeliness and lead time), supplier capabilities (e.g., financial stability, innovative potential, sustainability effort), and relationship characteristics (e.g., collaboration, commitment, trust). Hamdi et al. (2018) reviewed sourcing strategies and supplier risk assessment in supply chains, and they emphasized that supplier selection must be based on multi-dimensional criteria, not just on the lowest cost. Relying on the review of supplier performance metrics in the literature and based on the available data obtained from the retailer, two critical measures of supplier reliability, namely, the number of split deliveries and delivery delay times, are considered in the statistical analysis in this study.

3. Methods

Initially, normality tests are carried out on the data to determine the appropriate hypothesis testing procedures that can be used. Due to the lack of normality in the supplier data, Kruskal-Wallis test, which is a non-parametric test adopted to compare more than two groups, is applied to analyze the differences between three suppliers in terms of the number of split deliveries and delay times. The effect sizes of the Kruskal-Wallis tests are computed and interpreted. Detailed numerical and graphical analysis of the data is provided. All statistical analyses in this study are performed using SPSS 27.

4. Data Collection

The data obtained from the footwear retailer is based on the procurement process report, warehouse entry report, and final product report. Data covers the years 2019 and 2020 and the reports include information such as season, variety code, deadline, demand quantities, warehouse entry dates, and warehouse entry quantities. After the data is pre-processed by eliminating the records with missing or inconsistent information, measures such as percent of satisfied demand, delivery lead times, early/late delivery rates, and quality control pass rates are computed for the three critical suppliers of the selected product group. The final data set includes 204 orders for Supplier 1, 198 orders for Supplier 2, and 92 orders for Supplier 3.

5. Results and Discussion

5.1 Numerical Results

The supplier performance measures selected are the number of split deliveries and the delay time measured in days. The descriptive statistics for these two measures are provided for each supplier in Table 1 below. A comparison of the mean number of split deliveries as well as the variances shows a slightly better performance by Supplier 3. Supplier 2 has the highest variance, but Supplier 1 also has a high variance of almost 5. In terms of delay time, Supplier 3 again has the lowest mean that is 40% less than the mean of Supplier 1 and 45% less than the mean of Supplier 2. However, in terms of the variance of delay time, Supplier 2 has the lowest variance and Supplier 1 has the highest variance.

Table 1. Descriptive statistics

	Statistic		Supplier 1	Supplier 2	Supplier 3
Number of Split Deliveries	Mean		3.00	2.90	2.71
	95% Confidence Interval for Mean	Lower Bound	2.69	2.58	2.31
		Upper Bound	3.31	3.21	3.10
	Median		2.00	2.00	2.00
	Variance		4.985	5.066	3.682
	Std. Deviation		2.233	2.251	1.919
	Minimum		1	1	1
	Maximum		11	11	10
Delay Time (days)	Mean		67.19	72.71	40.29
	95% Confidence Interval for Mean	Lower Bound	56.82	64.38	26.14
		Upper Bound	77.55	81.04	54.45
	Median		38.50	66.50	20.50
	Variance		5633.482	3531.485	4671.133

	Std. Deviation	75.057	59.426	68.346
	Minimum	-25	-50	-48
	Maximum	351.000	295.000	283.000

The supplier data on the number of split deliveries and delay time in days are tested for normality using the Kolmogorov-Smirnov and Shapiro-Wilk tests both of which revealed that the data set was not normally distributed. The results of normality tests for all three suppliers are presented in Table 2. All p-values (in the "Sig." columns of Table 2) except for the delay time of Supplier 2 (p = 0.044) are below 0.001.

Table 2. Tests of normality

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
Supplier		Statistic	df	Sig.	Statistic	df	Sig.
Number of Split Deliveries	1	0.246	204	0.000	0.819	204	0.000
	2	0.261	198	0.000	0.806	198	0.000
	3	0.252	92	0.000	0.810	92	0.000
Delay Time (days)	1	0.185	204	0.000	0.820	204	0.000
	2	0.065	198	0.044	0.967	198	0.000
	3	0.203	92	0.000	0.805	92	0.000

a. Lilliefors Significance Correction

Due to the lack of normality, as a non-parametric test suitable for comparison of the median values of data for three suppliers, Kruskal-Wallis test is performed and the results of the test are shown in Table 3. The p-value for delay time is below 0.001, indicating that there is a significant difference between the three suppliers in terms of delay time. However, the p-value for the number of split deliveries is 0.63, which means that there is no significant difference between the three suppliers in terms of the number of split deliveries.

Table 3. Results of Kruskal-Wallis test

Test Statistics[a,b]

	Delay Time (days)	Number of Split Deliveries
Kruskal-Wallis H	28.477	0.924
df	2	2
Asymp. Sig.	0.000	0.630

a. Kruskal Wallis Test

b. Grouping Variable: Supplier

Effect size estimates of hypothesis tests allow the assessment of the strength of the relationship between the investigated variables. The effect size of the Kruskal-Wallis test can be calculated as the eta-squared measure (η^2) and the epsilon-squared measure (E_R^2) (Tomczak and Tomczak 2014). These measures are based on the H-statistic value obtained in the Kruskal-Wallis test, the number of groups (k) and the total number of observations (n) as shown in the equations below.

$$\eta^2 = \frac{H - k + 1}{n - k}$$

$$E_R^2 = \frac{H}{(n^2 - 1)/(n + 1)}$$

The η^2 estimate assumes values between 0 to 1 and indicates the proportion of variance in the dependent variable explained by the independent variable. The widely accepted interpretation values are 0.01-0.06 for small effect, 0.06-0.14 for moderate effect, and above 0.14 for large effect. The E_R^2 coefficient also assumes values between 0, indicating no relationship, and 1, indicating a perfect relationship. The η^2 and E_R^2 measures of delay time test are 0.054 and 0.058, respectively. Therefore, there is a small effect of supplier type on delay time. The η^2 and E_R^2 effect sizes of Kruskal-Wallis tests on the number of split deliveries are 0 and 0.0019, respectively. These effect sizes are in accordance with the obtained p-values in that there is no significant relationship between supplier performance in terms of the number of split deliveries.

The frequency of values below and above the median for two measures are provided for each supplier in Table 4. The majority of the number of split deliveries are less than or equal to the median value for all suppliers, which shows that the majority of orders are delivered in low numbers of split deliveries. On the other hand, the majority of orders are delivered with delay times above the median by Supplier 2, whereas the other two suppliers delivered the majority of their orders with delay times below the median. Since Supplier 2 has a median delay time of 66.5 days as shown in Table 1 above, the fact that most of the orders are delayed for more than 66.5 days is an indicator of underperformance.

Table 4. Number of orders with number of split deliveries and delay time below and above the median values of each measure

		Supplier		
		1	2	3
Number of Split Deliveries	> Median	87	78	36
	<= Median	117	120	56
Delay Time (days)	> Median	90	129	27
	<= Median	114	69	65
	Total number	204	198	92

In addition to the frequencies of data points below and above the median values in Table 4, the mean ranks for each supplier, when the data for all the suppliers are sorted, are shown in Table 5. The mean ranks of the suppliers are close for the number of split deliveries. However, the mean rank of Supplier 3 is significantly less than that of Suppliers 1 and 2 for delay time. Therefore, if we rank the suppliers in terms of their delay time performance Supplier 3 would be the best and Supplier 2 would be the worst in terms of the median and the mean rank.

Table 5. Mean ranks for the number of split deliveries and delay time

	Supplier	N	Mean Rank
Number of Split Deliveries	1	204	254.64
	2	198	242.83
	3	92	241.73
	Total	494	
Delay Time (days)	1	204	244.57
	2	198	279.97
	3	92	184.11
	Total	494	

5.2 Graphical Results

The histograms of the number of split deliveries of three suppliers are presented in Figure 1. These histograms can be visually interpreted in terms of normality and it can be safely concluded that the number of split deliveries does not follow a normal distribution for any supplier as shown above with the numerical results of the normality tests.

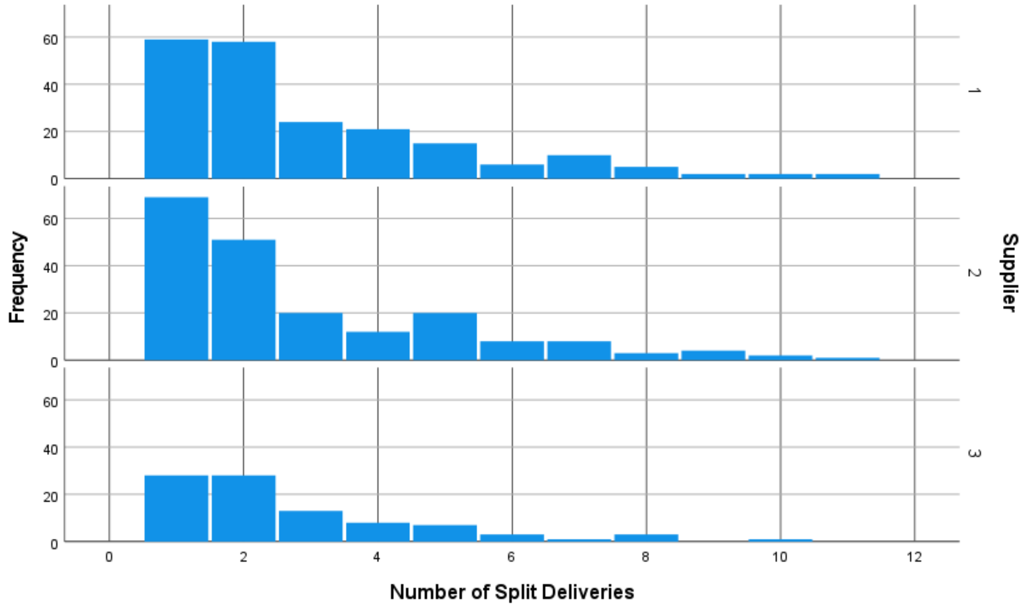

Figure 1. Histograms for the number of split deliveries of Suppliers 1, 2, and 3

To compare the performance of the three suppliers in terms of the number of split deliveries, the box plots are provided in Figure 2. All suppliers have equivalent distributions with the same median and similar number of outliers on the upper end of the distribution. This visual representation supports the Kruskal-Wallis test results reported above showing no significant difference between the three suppliers (p = 0.63) in terms of the number of split deliveries.

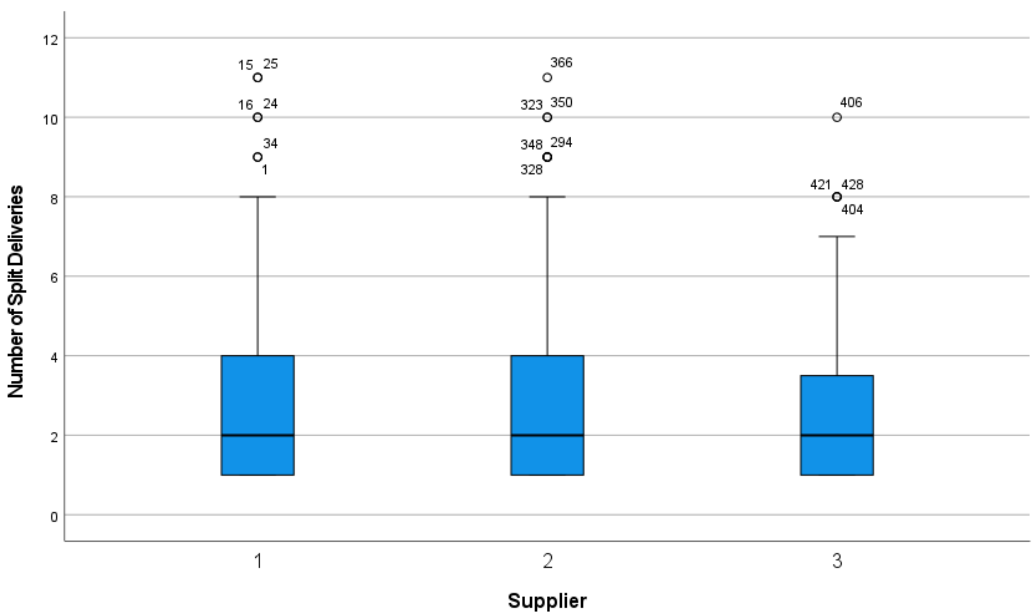

Figure 2. Boxplots for the number of split deliveries of Suppliers 1, 2, 3

The histograms of delay time data for three suppliers are presented in Figure 3. These histograms can be visually interpreted in terms of normality and it can be concluded that the delay time does not follow a normal distribution for any supplier as shown above with the numerical results of the normality tests.

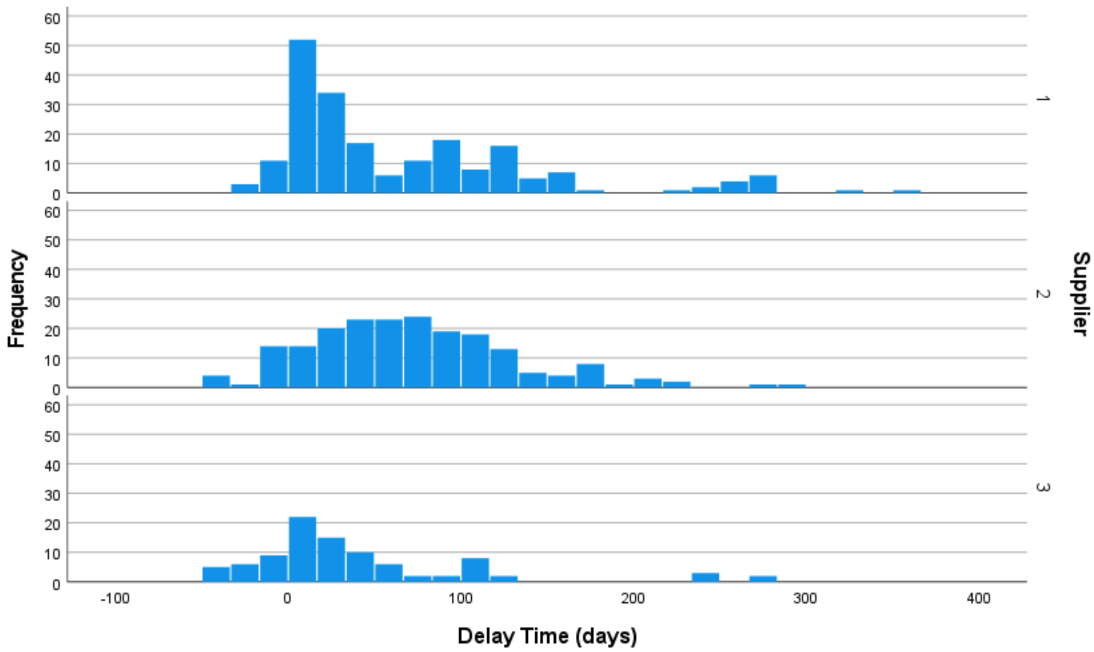

Figure 3. Histograms for delay time data of Suppliers 1, 2, and 3

The delay time performance of the three suppliers can be compared using the box plots displayed in Figure 4. Supplier 3 clearly has lower quartile values and less variance which supports the Kruskal-Wallis test results reported above.

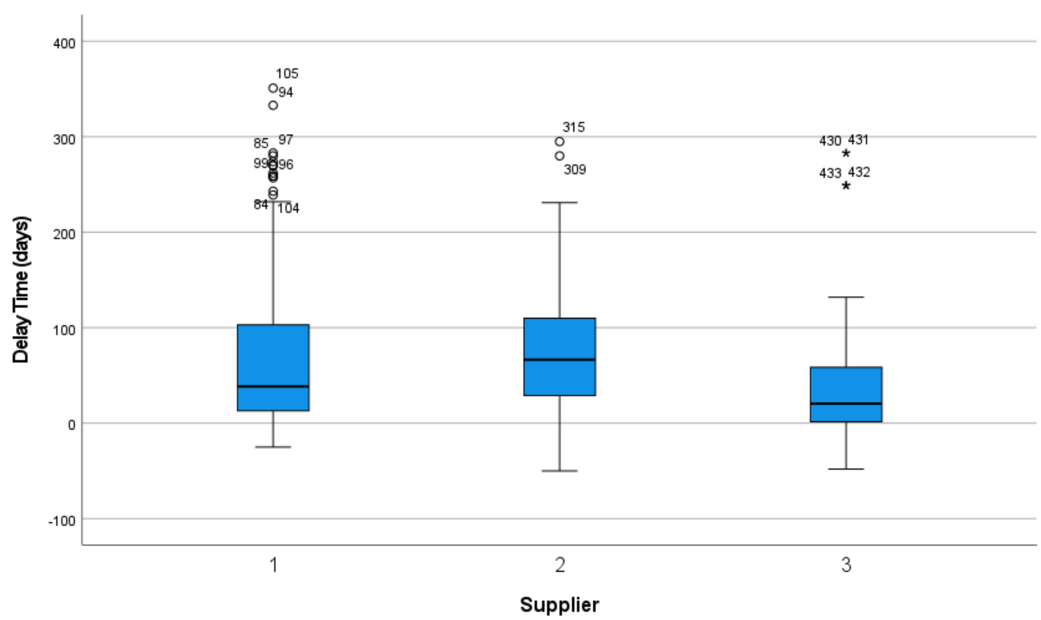

Figure 4. Boxplots for delay time data of Suppliers 1, 2, 3

5.3 Proposed Improvements
Currently, the retailer is relying on Supplier 1 for 41% of the orders, which constitute more than half of the products in the selected category; however, this supplier displays worse performance than Supplier 3 in terms of delivery delay time. Supplier 2 is also shown to underperform in terms of delay times having the highest mean and median, as well as the worst mean rank. Therefore, the retailer should consider reallocating its demand among its suppliers to improve reliability. A more reliable delivery process can be achieved by utilizing Supplier 3 more. If the existing contracts pose limitations on reallocating the demand, then the retailer should be aware of the expected performance level from each supplier and modify the timing and quantity of orders accordingly to avoid shortages or excess inventory holding costs.

6. Conclusion
Given all the uncertainties and risks involved in the current global marketplace, choosing reliable suppliers as partners, especially in the rapidly growing e-commerce market, is becoming more important. Supplier performance measurement systems are necessary for quantifying and monitoring the efficiency and effectiveness of suppliers' actions. Although several studies focus on supplier performance measurement, there are still research questions that can be explored and there is still room for improvement for many firms in practice. A statistical analysis of the supplier performance for a footwear retailer is presented in this study to provide insights about the shortcomings of the current state and to propose actions for the improvement of supply operations. The service performance metrics of three suppliers determined as the number of split deliveries and delay time are compared based on historical data. The differences between suppliers in service performance are identified, such as the significantly worse delay time performance of Supplier 2 and the significantly high performance of Supplier 3. Based on these results, reducing the number and size of orders made with Supplier 1 and 2 while utilizing the capacity of Supplier 3 more is suggested. This type of statistical analysis can be extended in the presence of additional data on supplier performance such as the logistical costs, lead times, distances, or responsiveness. More informed decisions can be made by considering several other criteria. Such analyses are valuable for retailers to understand the extent of service capacity of suppliers and to manage expectations accordingly. For retail companies to gain an edge over competitors in the marketplace, quantitative and statistical analyses are expected to remain essential for resilient supply chains in the future.

References
Alikani, R., Torabi, S.A., and Altay, N., Strategic supplier selection under sustainability and risk criteria. *International Journal of Production Economics*, vol. 208, pp. 69-82. https://doi.org/10.1016/j.ijpe.2018.11.018.

Ballou, R., *Business logistics management*, 5th ed., Upper Saddle River, New Jersey, Prentice-Hall, 2004.

Chai, J., Liu, J. N. K., and Ngai, E. W. T., Application of decision-making techniques in supplier selection: A systematic review of literature. *Expert Systems with Applications*, vol. 40, no. 10, pp. 3872–3885, 2013. https://doi.org/10.1016/j.eswa.2012.12.040.

Gordon, S., *Supplier evaluation and performance excellence: A guide to meaningful metrics and successful results*, Fort Lauderdale, Florida, J. Ross, 2008.

Gutierrez, D. M., Scavarda, L. F., Fiorencio, L., and Martins, R. A., Evolution of the performance measurement system in the logistics department of a broadcasting company: an action research. *International Journal of Production Economics*, vol. 160, pp. 1-12, 2015. https://doi.org/10.1016/j.ijpe.2014.08.012.

Hald, K. S. and Ellegaard, C., Supplier evaluation processes: the shaping and reshaping of supplier performance. *International Journal of Operations & Production Management*, vol. 31, no. 8, pp. 888-910, 2011. https://doi.org/10.1108/01443571111153085.

Hamdi, F., Ghorbel, A., Masmoudi, F., and Dupont, L., Optimization of a supply portfolio in the context of supply chain risk management: literature review. *Journal of Intelligent Manufacturing*, vol. 29, pp. 763-788, 2018. https://doi.org/10.1007/s10845-015-1128-3.

Huang, S. H. and Keskar, H., Comprehensive and configurable metrics for supplier selection, *International Journal of Production Economics*, vol. 105, no. 2, pp. 510-523, 2007. https://doi.org/10.1016/j.ijpe.2006.04.020.

Kannan, V.R. and Tan, K.C., Supplier selection and assessment: Their impact on business performance. *Journal of Supply Chain Management*, vol. 38, no. 3, pp. 11-21, 2002. https://doi.org/10.1111/j.1745-493X.2002.tb00139.x.

Luzzini, D., Caniato, F., and Spina, G., Designing vendor evaluation systems: An empirical analysis. *Journal of Purchasing & Supply Chain Management*, vol. 20, no. 2, pp. 113-129, 2014. https://doi.org/10.1016/j.pursup.2014.03.002.

Maestrini, V., Luzzini, D., Maccarrone, P., and Caniato, F., Supply chain performance measurement systems: A systematic review and research agenda. *International Journal of Production Economics*, vol. 183, pp. 299-315, 2017. https://doi.org/10.1016/j.ijpe.2016.11.005.

Mentzer, J. T., DeWitt, W., Keebler, J. S., Min, S., Nix, N. W., Smith, C. D., and Zacharia, Z. G., Defining supply chain management, *Journal of Business Logistics*, vol. 22, no. 2, pp. 1-25, 2001.

OECD, E-commerce in the time of COVID-19, Available: https://www.oecd.org/coronavirus/policy-responses/e-commerce-in-the-time-of-covid-19-3a2b78e8/, 2020.

Republic of Turkey Ministry of Trade, E-Commerce Information Platform, https://www.eticaret.gov.tr/istatistikler, 2021.

Schmitz, J. and Platts, K. W., Supplier logistics performance measurement: indications from a study in the automotive industry. *International Journal of Production Economics*, vol. 89, no. 2, pp. 231-243, 2004. https://doi.org/10.1016/S0925-5273(02)00469-3.

Tomczak, M. and Tomczak, E., The need to report effect size estimates revisited. An overview of some recommended measures of effect size, *Trends in Sport Sciences*, vol. 1, no. 21, pp. 19-25, 2014.

Waters, D., *Supply chain risk management: vulnerability and resilience in logistics*, Kogan Page Publishers, 2011.

Weber, C. A., Current, J. R., and Benton, W. C., Vendor selection criteria and methods, European Journal of Operational Research, vol. 50, no. 1, pp. 2-18, 1991. https://doi.org/10.1016/0377-2217(91)90033-R.

Acknowledgements

The author would like to acknowledge Simay Altintas, Bengisu Baytas, Nazli Dur, Asli Zulal Ozokten for providing the data used in this study.

Biography

Esra Agca Aktunc, Ph.D. is currently an Assistant Professor of Industrial Engineering in the Faculty of Engineering and Natural Sciences at Kadir Has University. She earned an undergraduate degree in Industrial Engineering from Bilkent University, a Master's and a Ph.D. degree in Industrial and Systems Engineering from Virginia Tech. Dr. Agca Aktunc's research areas include humanitarian logistics, disaster and emergency management, transportation, supply chain management, healthcare operations management, and scheduling. She has published in journals indexed by SCI® and SCI-Expanded®, such as Annals of Operations Research, International Journal of Physical Distribution & Logistics Management, Socio-Economic Planning Sciences, Journal of Natural Gas Science and Engineering, and International Journal of Energy Economics and Policy. She is a member of Institute for Operations Research and the Management Sciences (INFORMS), Institute of Industrial and Systems Engineers (IISE), and Operational Research Society of Turkey (YAD).