International Journal on Management Education & Emerging Technology
ISSN: 2583-9438

Is Net Promoter Score (NPS) Just a Myth? Exploring the Reality Through an Automotive Aftersales Case Study

Abdallah Amro [1], Dr. Azlan Ali [2], Dr. Asad Ur Rehman [3]

[1] Graduate School of Management, Management and Science University, [2] Graduate School of Management, Management and Science University, [3]Graduate School of Management, Management and Science University
1012023020688@gsm.msu.edu.my, azlan_ali@msu.edu.my, asad_ur@msu.edu.my

Abstract	Article Info
This study investigates the effectiveness of the traditional Net Promoter Score (NPS) framework in categorizing customers as Promoters, Passives, and Detractors within the context of automotive repair services and proposes alternative methods to potentially improve customer satisfaction measures. By utilizing various NPS calculation methods—NPS (Original), NPS (Top 1), NPS (Top 3), and NPS (Average)—we find that NPS (Average) offers the highest explanatory power for customer satisfaction, followed by NPS (Top 3), with significant differences in service aspects. The results demonstrate that modifying traditional NPS categories or considering average scores without categorization can potentially provide more accurate models of customer satisfaction. This research contributes to the field of customer satisfaction metrics by challenging conventional NPS categorization and offering insights specific to the automotive aftersales sector. Future work includes exploring other industry-specific applications and further refining satisfaction measurement techniques.	*Keywords:* Net Promoter Score (NPS), customer satisfaction, automotive repair services, categorization methods, satisfaction metrics

Date of Submission: 24-12-2024
Date of Review: 30-12-2024
Date of Acceptance: 13-02-2025
IJMEET / Volume 3, Issue 1, 2025
ISSN: 2583-9438

INTRODUCTION

Measuring customer satisfaction is vital for businesses striving to maintain a competitive edge and foster customer loyalty. A variety of methods exist to gauge satisfaction, ranging from straightforward surveys to complex models. Companies frequently rely on these surveys and feedback mechanisms to assess loyalty and satisfaction (Parasuraman, 2006). For example, tools like SERVQUAL evaluate service quality (Parasuraman, 2000; Parasuraman et al., 2005; Zeithaml et al., 2000), while SERVPERF focuses on service performance (Cronin & Taylor, 1994). Similarly, the Customer Satisfaction Score (CSAT) and Customer Effort Score (CES) are popular metrics, with CSAT evaluating satisfaction via questionnaires (Baquero, 2022) and CES assessing the ease of customer interactions (Budiman et al., 2023).

Among these methods, the Net Promoter Score (NPS) is a widely-used tool for measuring a customer's likelihood to recommend a service (Biesok & Wyród-Wróbel, 2021; F. F. Reichheld, 2003). NPS offers businesses a simplified mechanism for collecting feedback and benchmarking across industries (Lewis & Mehmet, 2020). However, it has been criticized for oversimplifying the complexities of customer loyalty and satisfaction (Keiningham et al., 2007). For instance, the categorization into promoters, passives, and detractors has been deemed arbitrary (Cazzaro & Chiodini, 2023), and the score's ability to predict customer sentiment is often questioned (Eskildsen & Kristensen, 2011).

In the automotive repair service industry, customer satisfaction heavily impacts success and relies on service characteristics such as repair quality and timeliness (Attarfi & Dachyar, 2022; Novacescu, 2019). Customers seek accurate and efficient service to swiftly and safely return their vehicles to the road (Wang et al., 2022). To ensure customer satisfaction and loyalty, repair services must excel in both quality and efficiency (Fang & Fang, 2013). Despite the widespread use of NPS, its application within this industry's nuances is not comprehensively understood, highlighting a significant research gap.

This study aims to fill this gap by investigating different NPS calculation methods—Top 1, Top 2, Top 3, and NPS Average—to determine their effectiveness in capturing customer satisfaction in the automotive aftersales sector. The primary research question guiding this study is: How credible are the three NPS categories (promoters, passives, and detractors) when using the NPS calculations of Top 1, Top 2, and Top 3, and which method most accurately captures the relationship between NPS scores—including Top 1, Top 2, Top 3, and NPS Average—and customer-perceived service aspects?

To explore this question, the study was conducted in Saudi Arabia's automotive aftersales sector, employing customer surveys collected via SMS or phone calls by CRM departments independent of service teams, thus minimizing bias. These surveys assessed NPS ratings and satisfaction across seven critical service aspects, including the time taken to be welcomed, staff's willingness to listen, information provision about repair work, availability of substitute transportation, explanation of charges, adherence to repair timelines, and quality of vehicle handover.

Analytically, the study employs techniques such as ANOVA and post hoc tests (Games and Howell) to assess differences in service aspect ratings among various NPS categories. Regression analysis is used to evaluate the predictive power of service aspects on NPS scores using differing methods. Moreover, bootstrapping is applied to ensure robustness of estimates and to address the uncertainty associated with non-normal data distribution.

The structure of this paper is as follows: The next section reviews the literature on customer satisfaction measurements and critiques of the NPS. Following this, the methodology section details the research design, data collection process, and analytical techniques applied. The results section presents the findings of the study, and the discussion interprets these results in the context of existing research. Finally, the conclusion highlights the implications, limitations, and potential directions for future research.

Through this comprehensive approach, the research aims to identify the most effective NPS method that aligns with actual customer satisfaction, providing valuable insights for enhancing service quality and business outcomes within the automotive repair industry. This study not only addresses theoretical and practical gaps but also aims to refine the application of NPS, thereby advancing understanding and practice in customer satisfaction measurement.

LITERATURE REVIEW

Fred Reichheld introduced the Net Promoter Score (NPS) in a 2003 Harvard Business Review article aiming to measure customer loyalty and predict business growth using a "likelihood to recommend" question 0. The NPS quickly gained acceptance across industries due to its effectiveness in indicating customer satisfaction and potential success (Osmanski-Zenk et al., 2023; Reichheld, 2003). Reichheld later expanded NPS into a transformative approach that emphasizes actionable insights from customer feedback, refining it into the Net Promoter System. This system includes closed-loop feedback and employee engagement to drive improvement and foster a customer-centric culture within organizations (Madsen, 2020; Reichheld, 2011).

he Net Promoter Score (NPS) centers on a question: "On a scale of 0 to 10, how likely are you to recommend our product/service to a friend or colleague?" Respondents are categorized as promoters (scoring 9-10), passives (scoring 7-8), or detractors (scoring 0-6), reflecting varying levels of customer sentiment. Promoters actively recommend the product, aiding organic business growth, while passives are neutral and might switch brands when alternatives arise. Detractors potentially harm the brand with their negative experiences. NPS is calculated by subtracting the percentage of detractors from promoters, yielding a score between -100 and +100, indicative of customer loyalty and business performance (Adams et al., 2022; Bettencourt & Houston, 2023; Fisher, 2018). A positive NPS signifies more promoters than detractors, suggesting growth potential and value generation. It provides a standardized metric for assessing customer loyalty, helping companies track changes over time, identify areas needing improvement, and evaluate initiative success. Additionally, NPS facilitates benchmarking against industry peers for enhanced customer satisfaction and loyalty strategies (Dawes, 2023). Understanding customer perceptions of the NPS question and categories is crucial for effective interpretation (Reichheld, 2011). If customers perceive it merely as satisfaction, using an average score might suffice. However, if they grasp the categorical distinctions and specific cutoff points, the standard NPS calculation method is more appropriate (Agag et al., 2023; Grisaffe, 2007).

NPS Criticism

A common critique of the Net Promoter Score (NPS) is its oversimplification of the complexities involved in customer loyalty and satisfaction. By reducing customer sentiment to a single question, NPS potentially overlooks key aspects of the customer experience, such as product quality, pricing, customer service, brand reputation, perceived value, trustworthiness, and personal preferences (Keiningham et al., 2007; Cazzaro & Chiodini, 2023). Because NPS focuses exclusively on the likelihood to recommend, it may not fully capture all elements influencing customer loyalty (Kristensen & Eskildsen, 2014). Without understanding these nuanced factors, identifying specific areas for improvement can be challenging (Keiningham et al., 2008; Salisbury & Peasley, 2018). Moreover, NPS alone does not provide insights into what triggers customer sentiments. To gain a deeper understanding of customer preferences and loyalty drivers, businesses might need to complement NPS with other research methods, such as interviews or comprehensive surveys (Daud, 2012; Fan et al., 2015; Ibrahim & Wang, 2019).

Another concern regarding NPS is its reliability in predicting business growth. Although widely considered a success indicator, some argue that evidence supporting NPS's effectiveness is inconsistent and may not apply uniformly across different industries (F. Reichheld, 2011; Keiningham et al., 2007). Research suggests that while there are some correlations between NPS and growth in specific sectors, these relationships are less strong when multiple industries are combined (Keiningham et al., 2007; van Doorn et al., 2013). In contrast, NPS tends to be more insightful in industries where customer recommendations are common, especially when customers share a strong affinity with the products they use (Shaw, 2008). Additionally, other metrics like customer satisfaction and loyalty measures can serve as predictive indicators of firm performance, particularly in industries with short repurchase cycles and low switching costs (Gruca & Rego, 2005; Baehre et al., 2022). A frequent critique of the Net Promoter Score (NPS) is its arbitrary division of respondents into promoters, passives, and detractors (Cazzaro & Chiodini, 2023). NPS classifies those scoring 9 or 10 as promoters, 7 or 8 as passives, and 0 to 6 as detractors. Critics argue that these thresholds may not accurately capture variations in customer behavior or attitudes, leading to potential misrepresentation of their sentiments (Keiningham et al., 2007; Sweeney et al., 2020).

Furthermore, doubts have been raised about the reliability of these categorizations. The assumption that promoters are brand advocates and detractors engage in negative word of mouth does not always align with actual customer behaviors (Eskildsen & Kristensen, 2011; Ho & Nguyen, 2022). Some researchers suggest that NPS might not be as effective as traditional rating scales in predicting customer sentiment, as it may not differentiate well between the categories. Indicators like customer satisfaction or liking are considered to offer better insights (Cazzaro & Chiodini, 2023; Kristensen & Eskildsen, 2014).

Moreover, critics argue against reducing the 11-point NPS scale to a simple binary measure, as this could oversimplify and lose valuable information (Cazzaro & Chiodini, 2023). There is a growing sentiment that questions NPS's reliability as an indicator of customer retention and loyalty (Cazzaro & Chiodini, 2023), suggesting that traditional satisfaction and loyalty questions might be more accurate in predicting business outcomes.

A notable critique of the Net Promoter Score (NPS) is its exclusion of moderately satisfied customers, or "passives," who score 7 or 8 (Cazzaro & Chiodini, 2023). NPS focuses on those highly likely to recommend (scores of 9 or 10) or unlikely to recommend (scores of 0 to 6), ignoring the impact of passives. Critics suggest this exclusion might skew interpretations, potentially misrepresenting true customer loyalty. Conversely, supporters argue that passives neither strongly support nor harm the brand, thus minimally affecting business growth predictions (Piris & Gay, 2021).

Cultural factors significantly influence customer behaviors, affecting how NPS is perceived and understood across different regions (Agag et al., 2024; Shavitt & Barnes, 2020). Cultural norms can impact scoring tendencies; for instance, in some regions, giving a high score might be seen as inappropriate, resulting in more moderate scores (passives). While a 7 or 8 might be seen as favorable in some places, it may be considered average elsewhere (Cazzaro & Chiodini, 2023).

Research by Dawes (2023) indicates that NPS seeks to understand the likelihood of recommendations, aiming to derive insights from customer experiences. However, respondents might overstate their recommendation likelihood, as their responses could reflect imagined scenarios rather than actual behaviors. While this poses a limitation, NPS still offers valuable high-level insights at an aggregate level.
Researchers have put forward methods, for calculating the Net Promoter Score (NPS) in order to gain insights, into customer sentiment. These alternative calculations aim to address the criticisms and shortcomings of the NPS calculation offering perspectives on customer satisfaction.

Alternative methods to calculate NPS:
Researchers such as Grisaffe (2007), (Keiningham, Cooil, Aksoy, et al., 2007)), and Morgan & Rego (2007) propose alternative methods to evaluate customer feedback by calculating the average score from recommendation scales, rather than relying solely on the traditional Net Promoter Score (NPS). This average score approach considers all responses, including those of passives, offering a broader understanding of customer sentiment and satisfaction levels (Baehre et al., 2022). However, interpreting average scores can be more complex than traditional NPS metrics and may not align with established NPS benchmarks, making cross-industry comparisons challenging (Baehre et al., 2022).

The Weighted Net Promoter Score is another variation that adjusts the NPS by assigning weights to responses based on their business impact. This method acknowledges that not all customer feedback equally affects business outcomes, thereby providing a nuanced measure of customer sentiment and its organizational impact (Baehre et al., 2022).

Critics of NPS argue against the arbitrary nature of cutoff points used to categorize respondents as promoters, passives, or detractors, noting these can vary significantly across industries and miss the subtleties of customer sentiments. To address this, researchers recommend refining cutoff points through customer sentiment or statistical analysis to better capture customer behavior and intentions (Baehre, O'Dwyer, O'Malley, & Story, 2022). This optimization aims to offer a more accurate representation of overall satisfaction and customer sentiment through NPS.

The "NPS Top 3" method, discussed by van Doorn et al. (2013) and used by Baehre et al. (2022), adjusts the traditional NPS calculation by altering cutoff points for categorizing respondents. In this approach, individuals who rate a 6 are moved from the "Detractor" to the "Passive" category, and those rating an 8 are upgraded from "Passive" to "Promoter." Thus, scores of 8, 9, or 10 are considered promoters, highlighting the prominence of these three ratings in the "NPS Top 3" calculation.

Conversely, the "NPS Top 1" approach, also explored by Baehre et al. (2022), further narrows the promoter category to only those scoring a perfect 10. Ratings of 8 and 9 become passives, while scores from 0 to 7 remain detractors, aligning with the stricter traditional NPS model.

Research by Baehre et al. (2022) indicates no significant correlation between sales growth and NPS scores from either the original NPS model or the modified "NPS Top 3" and "NPS Top 1" methods. These findings highlight the need to acknowledge the limitations of using NPS scores as sole indicators of business growth or success, suggesting a more nuanced or comprehensive approach might be necessary for accurate prediction and insight.

Overview of NPS Measures

NPS Measure	Cut-off Points	Calculation	Past Studies
NPS (Original) (Top 2)	Promoters: 9 to 10, Passives: 7 to 8, Detractors: 0 to 6.	%Promoters - %Detractors	(Reichheld, 2003, 2011)

NPS (Top 1)	Promoters: 10, Passives: 8 to 9, Detractors: 0 to 7.	%Promoters - %Detractors	(Baehre, O'Dwyer, O'Malley, & Lee, 2022)
NPS (Top 3)	Promoters: 8 to 10, Passives: 6 to 7, Detractors: 0 to 5.	%Promoters - %Detractors	(van Doorn et al., 2013) (Baehre, O'Dwyer, O'Malley, & Lee, 2022)
Arithmetic Mean	N/A	$X = \frac{\sum x}{n}$	(Grisaffe, 2007; Keiningham, Cooil, Andreassen, et al., 2007; Morgan & Rego, 2007)

METHODOLOGY

This study aims to assess the effectiveness of several Net Promoter Score (NPS) calculation methods—NPS (Top 2) (Original), NPS (Top 1), NPS (Top 3), and NPS (Average)—to identify which best reflects customer satisfaction within the automotive aftersales sector. The focus is to address the challenge of accurately measuring customer satisfaction by examining how various NPS calculations influence feedback categorization into "Promoters," "Passives," and "Detractors."
The study was conducted under an industrial research agreement between Management and Science University and Gulf Motors Company, the study adheres to requisite ethical standards, ensuring participant confidentiality and consent.

Data Collection:

The research was conducted in 2024 across the Kingdom of Saudi Arabia. Gulf Motors Company and its subsidiaries pushed the survey to all their customers during the study period. Data was collected from a total of 7,633 respondents who provided NPS ratings. The surveys were carried out by independent Customer Relationship Management (CRM) departments. They used SMS and phone calls to reach customers within one week of service completion to ensure unbiased and timely feedback. Surveys were administered by independent CRM departments, with data collection occurring within one week after each service event via SMS and phone calls, ensuring methodical data reliability.

Instrumentation:

The survey incorporated the NPS question and asked customers to rate their satisfaction across seven service dimensions:
- The time taken to be welcomed
- The willingness of the staff to understand/listen to your problems
- The provision of information about the work required on the vehicle before starting the repair
- The substitute transportation service
- The explanation of charges
- Ability to keep to promised timings
- The quality of the vehicle at the handover" (cleanliness, condition, etc.)

Statistical Analysis:
Data analysis included ANOVA to compare mean differences across NPS categories using both traditional and alternative calculations (NPS Top 1 and NPS Top 3). The Games and Howell post hoc test was employed to discern significant differences, particularly given altered cutoff points (Friston & Penny, 2011; Gillett, 1994; Turbé et al., 2023).

Due to non-normal data distribution, identified via Kolmogorov-Smirnov and Shapiro-Wilk tests ($p < 0.05$) (Noughabi, 2018), bootstrapping with 1,000 iterations was used to derive reliable parameter estimates and confidence intervals (Boos et al., 2023; Rousselet et al., 2023; Yzerbyt et al., 2018).
Regression analyses were conducted on the NPS (Original), NPS (Top 1), and NPS (Top 3) methodologies. The goal was to determine the optimal model fit for customer satisfaction interpretation, establishing criteria to evaluate each method's effectiveness (Friston & Penny, 2011; Turbé et al., 2023).
Implementation Specifics:

RESULTS & DISCUSSION

In the study, 7,633 respondents provided Net Promoter Score (NPS) ratings, though responses to specific service aspects varied in number. Key areas evaluated included "The time taken to be welcomed" (2,158 responses), "The willingness of staff to understand and listen to customer problems" (2,166), "The provision of information about work required on the vehicle before repair" (2,112), "The substitute transportation service" (1,638), "The explanation of charges" (1,963), "Ability to keep to promised timings" (2,045), and "The quality of the vehicle at handover" (2,120).

Upon examining mean values for these service aspects across different NPS calculation methods—NPS (Original), NPS (Top 1), and NPS (Top 3)—distinct trends emerged. Typically, Promoters are expected to show the highest mean satisfaction scores, followed by Passives, with Detractors scoring the lowest (F. Reichheld, 2011). However, only the NPS (Top 3) method consistently followed this pattern across all service dimensions. For instance, under the NPS (Top 3) methodology, Promoters reported significantly higher satisfaction compared to Passives and Detractors across all service aspects. Conversely, the Original and Top 1 methods did not show this consistent gradient of satisfaction across categories, highlighting differences in customer perception and satisfaction depending on the NPS calculation method used.
These varying mean scores underscore the importance of selecting an appropriate NPS methodology to accurately gauge customer satisfaction and interpret feedback effectively across different service dimensions.

		TW		SW		PW		AT		EC		KP		QF	
		Mean	Count	Mean	Count	Mean	Count	Mean	Count	Mean	Count	Mean	Count	Mean	Count
NPS (Original)	Detractor	6.01	1259	6.04	1259	5.63	1259	3.03	1259	5.78	1259	5.65	1259	6.06	1259
	Passive	7.98	1197	8.27	1197	7.94	1197	5.05	1197	8.01	1197	8.15	1197	8.07	1197
	Promoter	7.30	5177	6.85	5177	6.37	5177	3.04	5177	6.84	5177	7.53	5177	7.76	5177
NPS (Top 1)	Detractor	6.42	1704	6.51	1704	6.11	1704	3.45	1704	6.26	1704	6.19	1704	6.48	1704
	Passive	8.17	1557	8.47	1557	8.16	1557	5.18	1557	8.20	1557	8.36	1557	8.29	1557
	Promoter	7.24	4372	7.10	4372	6.46	4372	2.75	4372	6.85	4372	7.79	4372	7.83	4372
NPS (Top 3)	Detractor	5.28	710	5.06	710	4.78	710	2.74	710	4.87	710	4.71	710	5.17	710
	Passive	7.25	994	7.55	994	7.08	994	4.05	994	7.29	994	7.25	994	7.41	994
	Promoter	8.10	5929	8.37	5929	8.04	5929	5.01	5929	8.10	5929	8.32	5929	8.26	5929

TW: The time taken to be welcomed, SW: The willingness of staff to understand and listen to customer problems, PW: The provision of information about the work required on the vehicle before starting the repair, AT: The availability of substitute transportation services. EC: The explanation of charges incurred. KP: The ability to adhere to the promised timeframe for completion of work, and QH: The quality of the vehicle at the time of handover (cleanliness, condition, etc).

Post Hoc Analysis: Detailed Assessment of Customer Satisfaction Using Different NPS Methodologies
The post hoc analysis aimed to discern the effectiveness of varied Net Promoter Score (NPS) methodologies—NPS (Original), NPS (Top 1), and NPS (Top 3)—in reflecting differences in customer satisfaction across multiple service dimensions. This assessment facilitates understanding of how customer perceptions vary amongst Detractors, Passives, and Promoters, providing valuable insights into service quality in the automotive sector.

NPS (Top 2) (Original): The analysis focused on examining the mean values of various service aspects across the three NPS categories—Detractors, Passives, and Promoters—to identify any significant differences in customer perceptions. The aim was to understand how various service dimensions were evaluated by different customer groups.
The analysis of "The Time Taken to be Welcomed" revealed no significant difference between Passives and Promoters, with a p-value of .494 (mean difference = 0.560, standard error = .491). This indicates that both groups shared a similar view on the promptness of the welcoming process. Similarly, for "The Willingness of Staff to Understand/Listen," there was no statistically significant difference between Promoters and Detractors (p = .181, mean difference = 0.958, standard error = .532), suggesting that both groups rated staff willingness similarly.
Regarding "Provision of Information About Work," the absence of significant differences between Promoters and both Detractors (p = .288, mean difference = 0.845, standard error = .554) and Passives (p = .023, mean difference = -1.508, standard error = .549) suggests uniform perceptions about staff communication on vehicle work. For "Substitute Transportation Service," no significant difference was found between Promoters and

Detractors (p = 1.000, mean difference = 0.002, standard error = .540), indicating a shared view on the availability of substitute transportation.

The "Explanation of Charges" also showed no significant differences between Promoters and Detractors (p = .354, mean difference = 0.846, standard error = .609), as well as between Promoters and Passives (p = .068, mean difference = 1.380, standard error = .603), indicating similar perceptions across these groups regarding clarity of charges. Additionally, the analysis of "Ability to Keep to Promised Timings" reflected no significant difference between Passives and Promoters (p = .426, mean difference = 0.650, standard error = .516), highlighting consistent views on adherence to timelines.

Lastly, the "Quality of the Vehicle at Handover" exhibited a non-significant difference between Promoters and Passives (p = .459, mean difference = 0.649, standard error = .541), suggesting a common evaluation of vehicle condition at handover. Bias analysis revealed minimal bias, such as a bias of -0.018 in the "Quality of Vehicle at Handover," which did not significantly affect the outcomes, as p-values exceeded 0.05. Thus, while some comparisons indicated significant differences, the majority did not, demonstrating consistency in perceptions across different NPS categorizations and suggesting that the NPS (Original) method might not effectively differentiate nuanced customer satisfaction for these service dimensions.

NPS (Top 1): The NPS (Top 1) methodology provides an alternative approach to understanding customer feedback by adjusting the classification into three categories: Promoters (score of 10), Passives (scores of 8 and 9), and Detractors (scores 0 to 7). This refinement aims to gain deeper insights into customer sentiments by performing post hoc tests to explore potential differences across these groups using the Games-Howell post hoc comparison analysis.

The analysis sought to determine whether significant differences exist in the means of various service aspects between these newly defined categories. For the service aspect regarding "The time taken to be welcomed," no significant mean differences emerged between Detractors and Promoters (Mean Difference = -.876, p = .307) or between Passives and Promoters (Mean Difference = .969, p = .241). This consistency suggests similar opinions among these groups concerning the promptness of their reception.

In evaluating "The willingness of the staff to understand/listen to your problems," the analysis similarly found no significant difference between Detractors and Promoters (Mean Difference = -.740, p = .427). A marginally non-significant difference appeared between Passives and Promoters (Mean Difference = 1.408, p = .056), hinting at subtle perception variances but lacking strong statistical backing.

For "The provision of information about the work required on the vehicle," there was no significant difference between Detractors and Promoters (Mean Difference = -.558, p = .657), nor between Promoters and Passives (Mean Difference = -1.539, p = .054), suggesting uniform satisfaction with communication on vehicle repairs.

Concerning "The substitute transportation service," the analysis revealed no significant differences between Detractors and Promoters (Mean Difference = .651, p = .530), although a significant difference was noted between Promoters and Passives (Mean Difference = -2.415, p = .001).

Similarly, for "The explanation of charges," there was no significant difference between Detractors and Promoters (Mean Difference = -.423, p = .819) or between Promoters and Passives (Mean Difference = -1.562, p = .081), indicating consistent perceptions of charge clarity. When examining "Ability to keep to promised timings," no significant differences were present between Passives and Promoters (Mean Difference = .724, p = .458) or between Detractors and Promoters (Mean Difference = 1.488, p = .046). Lastly, the "Quality of the vehicle at handover" showed no significant differences between Detractors and Promoters (Mean Difference = -1.190, p = .129) and between Promoters and Passives (Mean Difference = .752, p = .431).

The biases observed in these comparisons were minor, ranging from .000 to .028, pointing to the likelihood that original sample estimates reliably represent actual population mean differences. The standard error provided insights into the variability of bootstrap estimates, while the Bias-Corrected and Accelerated (BCa) 95% Confidence Intervals accounted for potential bias and skewness.

Overall, the analysis indicates that the NPS (Top 1) categories—Detractors, Passives, and Promoters—did not exhibit significant differences across the studied service aspects, suggesting that customer opinions within these groups remain relatively consistent.

NPS (Top 3): The Games-Howell post-hoc analysis of the NPS (Top 3) methodology revealed significant differences in customer satisfaction across all service aspects. This approach broadens the classification, identifying more distinct satisfaction gradients among Detractors, Passives, and Promoters.

Detractors consistently reported significantly lower satisfaction levels compared to other groups. For the service aspect "The time taken to be welcomed," Detractors were notably less satisfied than Passives (mean difference = -2.054, $p < .001$) and Promoters (mean difference = -2.893, $p < .001$). This dissatisfaction extended to "The willingness of staff to understand and listen," where Detractors rated their satisfaction lower than both Passives (mean difference = -2.553, $p < .001$) and Promoters (mean difference = -3.434, $p < .001$).

The trend continued for "the information provided about vehicle work required," with Detractors indicating less satisfaction than Passives (mean difference = -2.379, $p < .001$) and Promoters (mean difference = -3.280, $p < .001$). Similarly, in "The substitute transportation service" aspect, Detractors' experiences were less favorable compared to Passives (mean difference = -1.400, $p < .001$) and Promoters (mean difference = -2.349, $p < .001$).

Furthermore, Detractors reported lower satisfaction with "The explanation of charges" compared to Passives (mean difference = -2.398, $p < .001$) and Promoters (mean difference = -3.176, $p < .001$). The pattern persisted in "The ability to keep to promised timings," with Detractors showing less contentment than Passives (mean difference = -2.374, $p < .001$) and Promoters (mean difference = -3.452, $p < .001$). Lastly, Detractors were also less satisfied with "The quality of the vehicle at handover," being less content than Passives (mean difference = -2.233, $p < .001$) and Promoters (mean difference = -3.102, $p < .001$).

These findings highlight distinct levels of satisfaction among customer groups, with Promoters typically demonstrating the highest satisfaction, followed by Passives, while Detractors report the lowest. The statistical significance of the p-values across these dimensions robustly confirms these differences, illustrating the NPS (Top 3) methodology's efficacy in capturing nuanced customer satisfaction across a spectrum of service aspects.

Overall, the analyses reveal that NPS (Original) and NPS (Top 1) fail to distinctly differentiate satisfaction levels across customer categories, whereas NPS (Top 3) successfully highlights significant satisfaction contrasts. Using the NPS (Top 3) methodology showcases clear distinctions in customer experiences, evidencing the highest satisfaction among Promoters, followed by Passives, with Detractors expressing the least satisfaction across various service dimensions. This evidence supports the use of NPS (Top 3) for a comprehensive evaluation of customer satisfaction, providing both depth and clarity in understanding customer sentiments (Friston & Penny, 2011; Turbé et al., 2023).

Regression Analysis:

In the subsequent sections, we will conduct regression analyses for four variations of the Net Promoter Score (NPS) categorizations: Original NPS, NPS (Top 1), NPS (Top 3), and NPS (Average). The objective of these analyses is multi-fold and aims to provide a comprehensive understanding of how different NPS frameworks can influence the interpretation of customer satisfaction and satisfaction data among the seven service aspects.

NPS (Top 2) (Original): The regression analysis of the NPS (Original) methodology sheds light on how various service aspects impact customer satisfaction scores. The correlation coefficients between NPS (Original) and service dimensions, such as "The time taken to be welcomed" ($r = .280$, $p < .001$) and "The willingness of staff to understand/listen" ($r = .307$, $p < .001$), show weak positive relationships, with the coefficients ranging from .216 to .322, indicating only moderate association (Ratner, 2009).

The regression model achieved an overall R value of .393 and an adjusted R square of .150, signifying that 15.0% of the variance in NPS (Original) scores is accounted for by these predictors. The ANOVA results confirmed the model's significance, with a regression sum of squares of 70.547 and an F-value of 38.514 ($p < .001$), underscoring the predictors' meaningful impact. Key influences on NPS scores included "The provision of information about work required," "The substitute transportation service," "Ability to keep to promised timings," and "The quality of vehicle at handover," noted in the regression equation:

$$\begin{aligned} NPS\ (Original) = {} & 0.963 + 0.012 \times \text{"The provision of informantion about work required on the vehicle before it was done"} \\ & + 0.012 \times \text{"The susbtitute transportation service"} + 0.022 \times \text{"Ability to keep promised timing"} \\ & + 0.015 \times \text{"The quality of vehicle at handover (Cleanliness, still in pre} - \text{service condition, etc"} \end{aligned}$$

Collinearity diagnostics indicated no significant multicollinearity issues (VIFs under 5) (Kock, 2017), and the residuals' distribution demonstrated normality and homoscedasticity, reinforcing model validity (Espinheira et al., 2021).
This analysis highlights that while certain service aspects significantly affect NPS scores, others do not markedly contribute, offering targeted insights for potential service improvements to enhance customer satisfaction.

*NPS (Top 1):*The analysis for NPS (Top 1) demonstrated that all Pearson correlation coefficients between the service aspects and NPS (Top 1) scores were positive but weak, with values below the 0.3 threshold, indicating minimal effect sizes (Ratner, 2009). Specifically, the weakest correlation was between the "substitute transportation service" and NPS (Top 1) ($r = .156$, $p < .001$). Other weak correlations included "the explanation of charges" ($r = .212$, $p < .001$) and "the quality of the vehicle at handover" ($r = .233$, $p < .001$). Slightly stronger yet still weak correlations were seen for "the time taken to welcome customers" ($r = .228$, $p < .001$) and "providing information about required vehicle services" ($r = .242$, $p < .001$). The strongest among them, though still weak, were "the willingness of the staff to understand and listen" ($r = .263$, $p < .001$) and "ability to keep promised timings" ($r = .264$, $p < .001$).
In the multiple regression analysis, which used these seven service aspects as predictors to predict NPS (Top 1) scores, the model accounted for 10.3% of the variance (R Square = .103), adjusted to 9.9% when considering the number of predictors (Adjusted R Square = .099). The standard error of the estimate was .49208, indicating the variability in observed values not explained by the model. The overall model was statistically significant, as evidenced by $F(7, 1476) = 24.276$, $p < .001$. The Durbin-Watson statistic was 1.651, suggesting no autocorrelation in the sample (Kim, 2022).
Among significant relationships, "The willingness of the staff to understand/listen to your problems" ($B = .014$, $p = .020$), "The ability to keep to promised timings" ($B = .017$, $p < .000$), and "The quality of the vehicle at the handover" ($B = .013$, $p = .005$) were positively associated with NPS (Top 1) scores. Other variables, like "The time taken to be welcomed" and "The provision of information about the work required on the vehicle," showed positive coefficients but were not statistically significant ($p = .268$ and $p = .116$, respectively). Moreover, "The substitute transportation service" and "The explanation of charges" also had non-significant p-values ($p = .083$ and $p = .812$, respectively). "Ability to keep to promised timings" had the highest Beta value (.115), indicating its relatively strong importance among factors in predicting NPS (Top 1). The model derived from the regression analysis depicting NPS (Top 1) as a function of key service factors is as follows:

$$NPS\ (Top\ 1) = 0.911 + 0.014 \times "Willingness\ of\ staff\ to\ understand/listen" + 0.017 \times "Ability\ to\ keep\ promised\ timing" \\ + 0.013 \times "The\ quality\ of\ vehicle\ at\ handover\ (Cleanliness, still\ in\ pre-service\ condition, etc"$$

Collinearity diagnostics showed that most service aspects exhibited variance proportions below concerning levels, suggesting no significant multicollinearity. Although the highest condition index was slightly above 10, indicating potential multicollinearity, it remained below the problematic threshold of 30. Overall, the analysis suggests that certain service aspects significantly affect NPS (Top 1) scores, offering insights into areas where service improvements can enhance customer satisfaction.

NPS (Top 3): The analysis of NPS (Top 3) via Pearson correlation coefficients assessed the relationship between NPS scores and various service aspects, based on data from 1,484 respondents. The highest correlation was noted with "The willingness of staff to understand and listen to customer problems" ($r = .420$, $p < .001$), indicating a moderate relationship (Ratner, 2009). Similarly, "The ability to keep to promised timings" and "The provision of information about required work" also demonstrated moderate correlations with $r = .388$ and $r = .383$, respectively (both $p < .001$). Moderate correlations were additionally observed with "The quality of vehicle at handover" ($r = .354$, $p < .001$), "The time taken to be welcomed" ($r = .348$, $p < .001$), and "The explanation of charges" ($r = .347$, $p < .001$). However, "The substitute transportation service" showed the weakest correlation at $r = .239$ ($p < .001$).
In the multiple regression analysis intended to predict NPS (Top 3) scores using these seven predictors, the model explained 24.4% of the variance (R Square = .244), which was slightly adjusted to 24.1% (Adjusted R Square = .241) after accounting for predictor complexity. The standard error of the estimate was .68921, indicating the variability in NPS predictions not captured by the model. Change statistics highlighted a

significant increase in model explanatory power with an R Square Change of .244, an F Value of 68.230, and a significance level of $p < .001$, signifying considerable improvement over a null model. The Durbin-Watson statistic of 1.965 suggested no autocorrelation in the residuals (Kim, 2022).

ANOVA confirmed that the regression model was significant with an F value of 68.230 ($p < .001$). The regression sum of squares was 226.873, indicating the variance explained by the model with a mean square of 32.410. The residual sum of squares was 701.125, reflecting unexplained variance in NPS (Top 3) scores. The relatively large proportion of unexplained variance (approximately 75.4%) points to the existence of other influential factors beyond the current predictors.

Significant predictors identified in the regression included "The willingness of the staff to understand/listen" (B = .039, $p < .001$, Beta = .158), "Ability to keep to promised timings" (B = .030, $p < .001$, Beta = .134), "The quality of the vehicle at handover" (B = .027, $p < .001$, Beta = .120), "Provision of information about required work" (B = .021, p = .004, Beta = .091), and "Substitute transportation service" (B = .014, p = .007, Beta = .067). "The time taken to be welcomed" and "The explanation of charges" did not show statistically significant relationships (p = .103 and p = .135, respectively). The regression equation expressed the relationship as:

$$NPS\ (Top\ 3) = 0.997 + 0.039 \times "Willingness\ of\ staff\ to\ understand/listen" + 0.030 \times "Ability\ to\ keep\ promised\ timing" + 0.027 \times \text{The quality of vehicle at handover (Cleanliness, still in pre-service condition, etc)} + 0.021 \times "Provision\ of\ information\ about\ work\ required" + 0.014 \times "Substitute\ transportation\ service"$$

Collinearity diagnostics indicated no significant multicollinearity, with all Variance Inflation Factor (VIF) values well beneath the threshold of 10, confirming model robustness in terms of predictor independence.

In summary, the findings underscore the importance of specific service aspects, such as staff willingness and adherence to promised timings, in influencing NPS (Top 3) scores. However, significant unexplained variance suggests room for further research to uncover additional factors impacting customer satisfaction.

NPS (Average): The analysis of NPS (Average) focused on the relationships between customer recommendation likelihood and various service aspects, revealing several significant findings. Key elements like "willingness of staff to understand/listen," "provision of information about the work required on the vehicle," and "ability to keep to promised timings" had the highest correlations with recommendation likelihood, having coefficients of 0.468, 0.427, and 0.426, respectively. These values suggest a moderate correlation (Ratner, 2009). Conversely, "substitute transportation service," with the lowest correlation of 0.256, highlights that while alternate transportation is valued, it is less critical to recommendations compared to other service aspects. All p-values registered as .000, emphasizing the statistical significance of these findings, which the bootstrap analysis further supports by showing minimal bias and precise confidence intervals.

The regression model indicated that 29.9% of the variance in recommending Service Centers could be attributed to the included predictors (R Square = 0.299), with a slight adjustment to 29.6% for model complexity (Adjusted R Square = 0.296). This implies a modest degree of predictability, leaving some variance unexplained and pointing to other influential factors. The standard error of the estimate was 2.509, which estimates the typical prediction error for recommendation likelihood.

An F Change statistic of 90.056 and a Statistically significant F Change ($p < .001$) confirmed the model's relevance and suggested substantial model improvement over a baseline model. The Durbin-Watson statistic of 1.900 indicated no autocorrelation in residuals, reinforcing model robustness (Kim, 2022). An ANOVA analysis affirmed the regression model's significance with an F value of 90.056 ($p < .001$), suggesting that service factors collectively impact customer recommendations. The model's sum of squares was 3969.028, illustrating the explained variance, while the residual sum of squares was 9293.125, highlighting remaining unexplained variance.

In the regression analysis, the intercept was significant (B = 1.052, $p < .001$). Among service aspects, "willingness of the staff to understand/listen" had a strong positive impact on recommendation likelihood (B = .163, $p < .001$, Beta = .176), as did "Ability to keep to promised timings" (B = .119, $p < .001$, Beta = .140), and "quality of the vehicle at handover" (B = .108, $p < .001$, Beta = .125). Other notable predictors included "provision of information about work required" (B = .094, p = .001, Beta = .107) and "substitute transportation service" (B = .049, p = .008, Beta = .063). Even "time taken to be welcomed" was significant, though less impactful (B = .063, p = .017, Beta = .069), whereas "explanation of charges" was not statistically significant (p = .078). The model equation developed from this analysis was:

$$NPS\ (Average) = 1.052 + 0.163 \times \text{"}Willingness\ of\ staff\ to\ understand/listen\text{"} + 0.119 \times \text{"}Ability\ to\ keep\ promised\ timing\text{"}$$
$$+ 0.108 \times \text{The quality of vehicle at handover (Cleanliness, still in pre-service condition, etc)}$$
$$+ 0.094 \times \text{"}Provision\ of\ information\ about\ work\ required\text{"} + 0.049 \times \text{"}Substitute\ transportation\ service\text{"}$$
$$+ 0.063 \times \text{"}Time\ ten\ to\ be\ welcomed\text{"}$$

Collinearity diagnostics identified no significant multicollinearity concerns among predictors, with condition indices and variance proportions within acceptable ranges, indicating each predictor's unique contribution. Residual statistics showed predicted values distributed from 1.69 to 7.43, with a residual range from -7.427 to 8.311 and standard deviations consistent across expected outcomes. Bootstrap analyses confirmed model reliability, validating predictions with minimal bias and consistent standard errors, thereby reinforcing the model's robustness and effectiveness in capturing correlations between service aspects and customer recommendations.

CONCLUSION

The post hoc analyses of the original NPS, NPS Top 1, and NPS Top 3 methodologies offer insights into customer satisfaction across various service aspects. The results from the original NPS and NPS Top 1 methods indicated no significant differences in customer perceptions among Detractors, Passives, and Promoters for the service aspects evaluated. In contrast, the NPS Top 3 method revealed notable differences, with Detractors consistently exhibiting lower satisfaction than Passives and Promoters, as reflected in statistically significant mean differences. These findings underscore that the Top 3 method effectively distinguishes variations in customer satisfaction, with Detractors reporting significantly lower satisfaction across service dimensions compared to other groups.

When assessing the overall model fit for different NPS calculation methods in relation to service aspects, the NPS Average model demonstrated the highest fit at 29.6%, suggesting superior capacity to explain NPS score variations through the associated service aspects. This was followed by the Top 3 method with a 24.1% fit, the original NPS at 15%, and the Top 1 approach at 9.9%. Although none of the models exhibit strong explanatory power for NPS scores, the Average model provides relatively richer insights into customer satisfaction dynamics compared to the Top 3, Original, and Top 1 models. This hierarchy highlights the varied impact of service aspects on customer satisfaction and promoter scores across different methodologies.

This research emphasizes the relative effectiveness of the NPS Average approach over other NPS calculations. Nonetheless, the study does not conclusively determine whether customers perceive the NPS inquiry as a comprehensive assessment of their overall experience or specific satisfaction elements. Given this distinction, further research is essential to explore this perspective. In conclusion, the findings suggest that the NPS Average may offer a more genuine reflection of customer sentiment towards a company compared to the NPS Top 3, Original, or Top 1 methods. The study advocates for a broader approach to customer feedback analysis, one that transcends traditional NPS categories and captures the complex spectrum of customer satisfaction.

The observed tendency for customers to use the NPS scale as a continuum, rather than discrete categories, supports a more holistic evaluation of customer feedback that aligns with their overall experience and intentions to repurchase. Cultural factors significantly influence customer behaviors and NPS perceptions, varying across regions (Agag et al., 2024; Shavitt & Barnes, 2020). Cultural norms can affect scoring tendencies; in some areas, high scores may appear inappropriate, leading to more moderate scores, whereas elsewhere, such scores might indicate relative satisfaction (Cazzaro & Chiodini, 2023)(. The average score approach, considering all responses, offers a comprehensive view of customer sentiment and satisfaction (Baehre, O'Dwyer, O'Malley, & Lee, 2022), though it introduces complexities versus traditional NPS metrics, complicating cross-industry comparisons. The "NPS Top 3" method refines NPS by adjusting cutoff points for respondent categorization, drawing attention to the importance of these adjustments in clarifying customer satisfaction nuances (van Doorn et al., 2013; Baehre et al., 2022).

In essence, this study concludes with a recommendation for businesses to look beyond narrow NPS categorizations, considering holistic satisfaction metrics that more accurately reflect customer loyalty and satisfaction.

RECOMMENDATIONS

Future studies should aim to further explore the complex dynamics of customer satisfaction and the effectiveness of different Net Promoter Score (NPS) methodologies. One area of focus could be investigating cultural influences on NPS perceptions and scoring behaviors across diverse regions, as cultural norms can significantly shape how customers respond to satisfaction surveys (Agag et al., 2024; Shavitt & Barnes, 2020). Understanding these cultural impacts could enhance the global applicability of NPS by tailoring assessment tools that account for regional differences (Cazzaro & Chiodini, 2023).

Additionally, research could delve into identifying and incorporating other unaccounted factors that influence customer satisfaction and recommendation likelihood, beyond the current service aspects considered. This could involve examining elements such as brand reputation, emotional engagement, and social influence, which may offer more comprehensive insights into customer experiences and behaviors.

Further investigation into the NPS Average approach is warranted to assess its potential as a more holistic measure of customer sentiment. Researchers should analyze how average scores correlate with actual customer behavior and loyalty over time, providing deeper insights into the long-term reliability of this method compared to traditional NPS metrics. Exploring advanced statistical models or machine learning techniques may also enhance the predictive capabilities of satisfaction assessments, allowing for the identification of complex patterns and interactions among multiple service dimensions.

Moreover, aligning NPS studies with longitudinal analyses can help identify changes in customer satisfaction and promoter scores over time, offering valuable insights into the effectiveness of interventions aimed at improving service quality. By incorporating these recommendations, future research could significantly contribute to understanding customer satisfaction, refining NPS methodologies, and optimizing feedback systems to drive business success.

REFERENCES

[1] Agag, G., Ali Durrani, B., Hassan Abdelmoety, Z., Mostafa Daher, M., & Eid, R. (2024). Understanding the link between net promoter score and e-WOM behaviour on social media: The role of national culture. *Journal of Business Research*, *170*, 114303. https://doi.org/10.1016/j.jbusres.2023.114303

[2] Baehre, S., O'Dwyer, M., O'Malley, L., & Lee, N. (2022). The use of Net Promoter Score (NPS) to predict sales growth: insights from an empirical investigation. *Journal of the Academy of Marketing Science*, *50*(1), 67–84. https://doi.org/10.1007/s11747-021-00790-2

[3] Baehre, S., O'Dwyer, M., O'Malley, L., & Story, V. M. (2022). Customer mindset metrics: A systematic evaluation of the net promoter score (NPS) vs. alternative calculation methods. *Journal of Business Research*, *149*, 353–362. https://doi.org/10.1016/j.jbusres.2022.04.048

[4] Cazzaro, M., & Chiodini, P. M. (2023). Statistical validation of critical aspects of the Net Promoter Score. *The TQM Journal*, *35*(9), 191–209. https://doi.org/10.1108/TQM-05-2022-0170

[5] Grisaffe, D. B. (2007). *QUESTIONS ABOUT THE ULTIMATE QUESTION: CONCEPTUAL CONSIDERATIONS IN EVALUATING REICHHELD'S NET PROMOTER SCORE (NPS)*. www.netpromoters.com

[6] Keiningham, T. L., Cooil, B., Andreassen, T. W., & Aksoy, L. (2007). A Longitudinal Examination of Net Promoter and Firm Revenue Growth. *Journal of Marketing*, *71*(3), 39–51. https://doi.org/10.1509/jmkg.71.3.039

[7] Madsen, D. Ø. (2020). One Marketing Metric To Rule Them All? An Examination of the Emergence and Rise of Net Promoter Score as a Marketing Fashion. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3738866

Morgan, N. A., & Rego, L. L. (2007). *The single-question trap*. https://www.researchgate.net/publication/291841695

[8] Osmanski-Zenk, K., Ellenrieder, M., Mittelmeier, W., & Klinder, A. (2023). Net Promoter Score: a prospective, single-centre observational study assessing if a single question determined treatment success after primary or revision hip arthroplasty. *BMC Musculoskeletal Disorders*, *24*(1). https://doi.org/10.1186/s12891-023-06981-y

[9] Reichheld, F. (2003). *The One Number You Need to Grow*. www.hbr.org

Reichheld, F. (2011). *The ultimate question 2.0 (revised and expanded edition): How net promoter companies thrive in a customer-driven world*. Harvard Business Review Press.

[10] Shavitt, S., & Barnes, A. J. (2020). Culture and the Consumer Journey. *Journal of Retailing*, *96*(1), 40–54. https://doi.org/10.1016/j.jretai.2019.11.009

[11] van Doorn, J., Leeflang, P. S. H., & Tijs, M. (2013). Satisfaction as a predictor of future performance: A replication. *International Journal of Research in Marketing*, *30*(4), 434. https://doi.org/10.1016/j.ijresmar.2013.10.003