

DO CUSTOMERS BEHAVE DIFFERENTLY AFTER JOINING A PAID LOYALTY PROGRAM? EVIDENCE FROM A LONGITUDINAL WITHIN-CUSTOMER ANALYSIS IN CZECH E-COMMERCE

PONAŠAJU LI SE KUPCI DRUKČIJE NAKON PRISTUPANJA PLAĆENOM PROGRAMU VJERNOSTI DOKAZI IZ LONGITUDINALNE ANALIZE POJEDINAČNIH KUPACA U ČEŠKOJ E-TRGOVINI

M Market-Tržište
Vol. 38, No. 1, 2026, pp. 9-26
UDK: 339.138:658.89:366.1
DOI <http://dx.doi.org/10.22598/mt/2026.38.1.9>
Original scientific paper

Filip Beneš^a, Mikuláš Gangur^b, Václav Stříteský^c

^a Prague University of Economics and Business, nám. W. Churchilla 1938/4, 130 67, Praha 3, CZECH REPUBLIC, ORCID: <https://orcid.org/0009-0000-6261-4145>, e-mail: filip.benes@vse.cz

^b University of West Bohemia, Univerzitní 22, 306 14, Plzeň, Czech Republic, ORCID: <https://orcid.org/0000-0002-7731-9146>, e-mail: gangur@kem.zcu.cz

^c Prague University of Economics and Business, nám. W. Churchilla 1938/4, 130 67, Praha 3, Czech Republic, ORCID: <https://orcid.org/0000-0002-1513-7937>, e-mail: vaclav.stritesky@vse.cz



Abstract

Purpose – While traditional loyalty programs are well-established in marketing, paid loyalty programs are gaining prominence across global e-commerce markets. This raises the question of whether paying members of such programs exhibit distinct purchasing patterns compared to their pre-enrolment behavior.

Design/Methodology/Approach – The present study addresses that research gap by conducting a quantitative analysis of 21,805 unique customers of a leading Czech

Sažetak

Svrha lako su tradicionalni programi vjernosti dobro etablirani u marketingu, plaćeni dobivaju sve veću važnost na globalnim tržištima e-trgovine. To nameće pitanje pokazuju li članovi koji plaćaju takve programe drukčije obrasce ponašanja u kupovini u usporedbi sa svojim ponašanjem prije učlanjenja.

Metodološki pristup Istraživanje rješava navedeni jaz provedbom kvantitativne analize na uzorku od 21 805 jedinstvenih kupaca vodećeg češkog e-trgovca na malo.

e-commerce retailer. Using paired *t*-tests, five behavioral hypotheses were tested across four sample variants.

Findings and Implications – Results indicated that customers' transaction frequency increased post-enrolment in a paid loyalty program compared to pre-enrolment, alongside a higher volume of items purchased. Contrary to expectations, members do not appear to systematically fragment their purchases and these behavioral changes are not reflected in their overall expenditure.

Limitations – This study is limited by its focus on a single e-commerce company in the Czech Republic within one industry. The use of secondary consumer data provided no control over the sampling process. The nature of the data excludes the identification of former members due to privacy laws in Europe.

Originality – The findings provide empirical support for paid loyalty programs and their effectiveness in enhancing behavioral loyalty and customer value by comparing members' behavior pre- and post-enrolment, rather than merely comparing members with non-members. This study offers practical insights for marketing practitioners when evaluating the design of their loyalty programs and determining the feasibility of a fee-based membership.

Keywords: paid loyalty programs, customer behavior, e-commerce, purchase frequency, loyalty marketing, Czech Republic

Pomoću uparenih *t*-testova testirano je na četiri varijante uzorka pet hipoteza o ponašanju kupaca.

Rezultati i implikacije Rezultati su pokazali da su se učestalost transakcija kupaca i broj kupljenih proizvoda povećali nakon u članjenja u plaćeni program vjernosti u usporedbi s prethodnim razdobljem. Suprotno očekivanjima, čini se da članovi ne usitnjavaju sustavno svoje kupovine, a navedene promjene u ponašanju ne odražavaju se na njihovu ukupnu potrošnju.

Ograničenja Istraživanje je ograničeno time što se temelji na podacima za jedno poduzeće za e-trgovinu u Češkoj Republici. Korištenje sekundarnih podataka o kupcima nije omogućilo kontrolu nad procesom uzorkovanja. Zbog zakona o zaštiti privatnosti u Europi priroda podataka onemogućila je identifikaciju bivših članova.

Doprinos Rezultati pružaju empirijsku potporu plaćenim programima vjernosti i njihovoj učinkovitosti u jačanju bihevioralne vjernosti i vrijednosti za kupce uspoređivanjem ponašanja članova prije i nakon u članjenja, umjesto pukog uspoređivanja članova s nečlanovima. Istraživanje nudi praktične uvide marketinškim stručnjacima pri procjeni dizajna njihovih programa vjernosti i utvrđivanju opravdanosti uvođenja članstva koje se naplaćuje.

Ključne riječi: plaćeni programi lojalnosti, ponašanje potrošača, e-trgovina, učestalost kupnje, marketing lojalnosti, Češka Republika

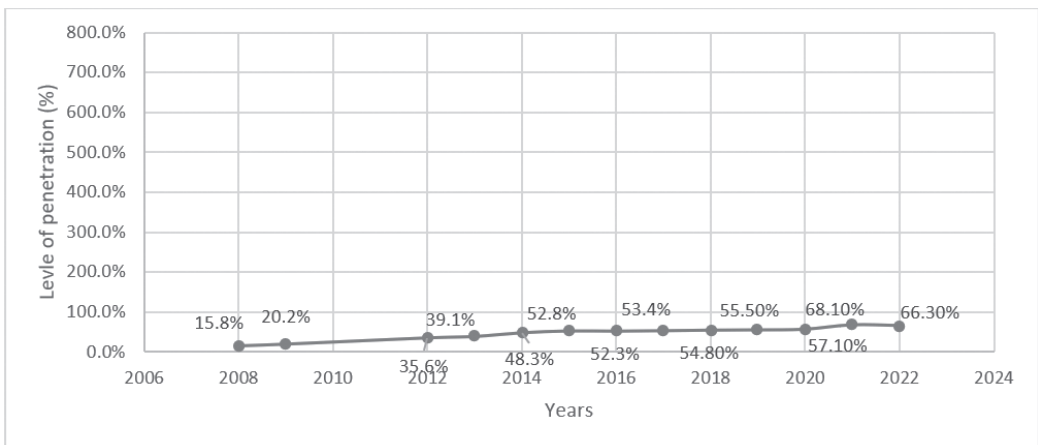
1. INTRODUCTION

A loyalty (or reward) program is a strategic marketing tool designed to nurture customer relationships in the hopes of consolidating their brand loyalty (Arbore & Estes, 2013; Leenheer & Bijmolt, 2008). Early definitions define loyalty programs as tools that reward members for their repeat patronage (Sharp & Sharp, 1997). The definition has evolved to combine marketing initiatives devised to positively influence consumers’ attitudes and behavior toward a company or brand (Henderson et al., 2011). To provide a more nuanced perspective on loyalty programs, Dorotic et al. (2012) identified five key characteristics of a well-designed loyalty program. First, it builds a relationship between members and loyalty program operators, ultimately leading to loyal customers and higher sales. Second, it is designed to enable communication between members and loyalty program operators. Third, it facilitates a long-term relationship between members and loyalty program operators, preventing par-

ticipation only during high-value deals. Fourth, it rewards members for loyal behavior. Lastly, it consistently deploys marketing efforts that drive targeted offers to its members.

Overall, loyalty programs are vastly popular. The Economist (2017) reports 3.3 billion U.S. loyalty program memberships on a population of 325 million in 2017 (United States Census Bureau, 2025). This accounts for about 29 memberships per household, ranging from free supermarket memberships to complex airline loyalty programs. However, only 46% of these memberships were active at that time, with an expected decline in usage (Colloquy, 2018). By 2024, the average number of total memberships in the U.S. was 19, with only 9.3 being active (Faria, 2025). In a post-COVID-19 market, the performance of loyalty programs is especially crucial as the majority of consumers reduced their spending, shopping frequency, and loyalty points accumulation (Jin et al., 2024).

GRAPH 1: Market penetration of loyalty programs in the Czech Republic



Source: Median, a.s. (2008–2022). MML-TGI Czech Republic dataset [Dataset]. Available from the Prague University of Economics and Business under institutional license.

In Central Europe, specifically in the Czech Republic, the market penetration of loyalty programs—the incidence of membership as a percentage of the total consumer base—has experienced rapid growth followed by near stagnation (see Graph 1) (Median, a.s., 2008–2022). This is based on data obtained from the Market & Media & Lifestyle (MML-TGI) project conducted by the Czech market research and public opinion agency Median a.s. Between 2008 and 2012, market penetration of loyalty programs increased by 19.8 percentage points (p.p.). From 2012 to 2015, growth slowed to an increase of only 17.2 p.p. only to decelerate even further from 2015 to 2018, recording an increase by no more than 2 p.p. Nonetheless, since the COVID epidemic, substantial visible growth has been reported, with marginal decreases between 2021 and 2022.

2. LITERATURE REVIEW

Current literature addresses a vast range of topics related to loyalty programs and adjacent fields. Nonetheless, many directions and questions remain unexplored (Breugelmans et al., 2015; Liu, 2007). Empirical research regarding European paid loyalty programs that incorporate selectivity through payment-based access to benefits is limited (Esmark et al., 2016; Ashley, et al., 2016). Rather than earning rewards through loyalty and repeat patronage, customers pre-pay for them.

A prominent example is Amazon Prime, a globally recognized paid membership that offers numerous benefits such as free delivery, streaming services, and e-books (Amazon, 2025). Some scholars classify Amazon Prime as a paid membership program rather than a loyalty program, contending that access is based solely on payment, without requiring ongoing demonstrations of loyalty (Du et al., 2024). Although it may have been a loyalty program initially, it has since evolved into a “revenue-generating business” (Walsman & Dixon, 2020).

However, this classification raises questions when applied to hybrid programs such as Miles & More. Traditionally, members who accumulat-

ed 100,000 miles (recently, M&M also incorporated points and shifted certain membership periods) within a calendar year were granted Senator status for two years. If they repeated this achievement within the following two years, their status was extended (Miles & More GmbH, n.d.). Notably, under specific circumstances, the Senator status could also be purchased for one year for a fee of EUR2,000 (Powell, 2022). This introduces a unique dynamic in which benefits can be earned through persistent engagement or acquired through direct payment, blurring the line between loyalty and membership programs.

While this distinction primarily concerns terminology, it is not the primary focus of this paper. A more pressing question is whether paid loyalty programs offer measurable benefits for their providers. Extensive research on loyalty programs indicates that, under reasonable conditions and terms, loyalty programs are beneficial for their providers, including increased customer retention, customer insight, and competitive differentiation (Sharp & Sharp, 1997; Dorotic et al., 2021; Breugelmans et al., 2015; Hwang et al., 2019). Nonetheless, research has also identified negative effects of such programs, including customer entitlement, frustration, or the program simply not performing as expected, resulting in a negligible or even negative effect on its members (Li et al., 2017; Strauss et al., 2005; Wang et al., 2016).

Although empirical research on paid loyalty programs is limited, existing studies suggest that paid loyalty program members exhibit characteristics that make them more valuable as customers (Ashley et al., 2016). These members display higher behavioral loyalty, focus more on provided rewards, and perceive the program as more valuable. Current research also indicates that membership fees are an effective mechanism for motivating customers to be more loyal while providing the firms with a certain revenue stream to operate the program (Walsman & Dixon, 2020). To expand the current published knowledge on paid loyalty programs, this paper

is aimed at investigating whether paid loyalty program members exhibit distinct purchasing behavior through a quantitative analysis of purchasing behavior in the Czech e-commerce market.

Furthermore, to avoid merely validating loyalty programs as effective marketing tools, this analysis also considers self-selection bias. A simple comparison between members and non-members does not account for self-selection bias—especially in the case of paid loyalty programs (Leenheer et al., 2007). As the program requires an enrollment fee, the decision to join can be a calculated move, where a customer realizes clear benefits—even with current purchasing habits—rather than a behavioral shift resulting from the program. Thus, the tested hypotheses focus on customer behavior pre- and post-enrolment. For example, the study focuses on whether consumer behavior changed positively post-enrolment or if customers were already purchasing more, indicating a predisposition to eventually join without significant behavioral differences.

In general, loyalty programs are designed to increase customer consumption by providing adequate rewards. Vroom (1964) claimed that loyalty programs create an expectancy of positive outcomes related to consumption. Latham and Locke (1991) further explained that, once consumers are aware of this correlation, they are likely to consume more. This is supported by empirical evidence indicating that companies utilizing loyalty programs experience an overall positive influence on sales and profitability (Bombajj & Dekimpe, 2020). However, whether this is valid when the benefits of a loyalty program are prepaid and are not “earned” remains unclear.

Prior research has indicated that loyalty programs can significantly increase the purchase frequency of members compared to non-members (Liu, 2007). These programs are designed to facilitate customer retention and promote sustained patronage (Sharp & Sharp, 1997). Ashley et al. (2016) further demonstrated that this effect

extends to fee-based loyalty programs, where members exhibit higher purchasing frequency than non-members. Based on these findings and the underlying logic of loyalty program design, the following hypothesis is proposed:

H1: Customer purchase frequency increases following their enrolment in a paid loyalty program.

In the context of rising shipping costs, consumers may reconsider their purchase decisions at checkout. However, where shipping prices are the main factor, this indecision often diminishes when shipping is free. Guo and Liu (2023) found that members of the loyalty programs offering free shipping often break up their purchases into smaller, more frequent orders shortly after enrolment, thus enjoying the benefit of free shipping without restraint. Consequently, their average spending per transaction does not increase, and the overall revenue from such customers may be offset by the increased operational costs. In the absence of shipping costs, members no longer feel the need to wait and consolidate items before purchasing. Based on this rationale, we posit the following:

H2: Customers purchase fewer items per transaction post-enrolment in a paid loyalty program.

Meyer-Waarden (2008) claims that loyalty program members exhibit significantly higher collective purchase volumes over a multi-year period compared to non-members, indicating a sustained behavioral shift. While being industry-specific, evidence from the airline sector demonstrates that firms implementing loyalty programs often attract and retain more active customers than firms that do not (Vilkaite-Vaitone & Papsiene, 2016). Together, these findings support the assumption that members of paid loyalty programs engage in a more frequent purchasing behavior that results in a greater number of items purchased. Therefore, the following hypothesis is proposed:

H3: Customers purchase a greater number of items post-enrolment in a paid loyalty program.

Expanding on Hypothesis 2, which is based on the findings of Guo and Liu (2023) that members purchase fewer items per transaction, Hypothesis 4 proposes to test whether this is concurrently reflected in the total amount of money spent per transaction:

H4: Customers spend less money per transaction post-enrolment in a paid loyalty program.

Loyalty programs, when effectively designed, can increase members' share of wallet by encouraging greater purchase frequency and volume (Leenheer et al., 2007; Meyer-Waarden, 2007). In the case of point-based programs, Dorotic et al. (2014) explain that the concept of redemption momentum—where customers increase their spending both pre- and post-point—redemption can significantly elevate purchase behavior. Paid loyalty programs, which typically require an upfront fee, may create a stronger psychological commitment to the brand. Supporting that claim, Ashley et al. (2016) found that paid program members record significantly higher monthly spending compared to non-members. Building on this evidence, the following hypothesis is proposed:

H5: Customers spend more post-enrolment in a paid loyalty program.

3. METHODOLOGY

3.1. Data source and sample

This study utilizes a comprehensive transactional dataset obtained from a leading Czech e-commerce company offering a paid loyalty program that grants its members free shipping, among other promotional rewards. The dataset spans a period from January 1, 2017 to November 9, 2021 and includes 1,026,701 unique transaction records from 23,492 end consumers. The acquisition of such longitudinal transaction data was crucial for appropriately assessing the behavioral impact of loyalty program membership (Meyer-Waarden, 2008). Of the total, 10,778 were enrolled in the company's paid loyalty program, while 12,714 were non-members. The original sample was generated through multiple rounds

of computer-assisted random sampling from the company's internal database. A final sample was selected for analysis after confirming consistency across preliminary descriptive statistics. The random sampling, preliminary analysis, and random choice of the sample were all conducted by the company providing the data. Although the authors contributed to the design of the variables of the data file, they maintained no control over the sampling and selection procedure. Subsequently, the sample was cleaned and aggregated to a customer level—from the original format, which displayed each transaction separately and was split per product category. The final sample comprised 21,805 unique customers.

All data utilized in this study was fully anonymized before being shared for analysis. The researchers were not granted access to any personally identifiable information. For extra precaution, the data file was secured with a high-entropy password. Both the data file and the password were handled separately.

3.2. Variable selection and testing

Each record before aggregation contained the transaction date, order ID, anonymized customer ID, anonymized customer email, product category, delivery method and location, loyalty program membership ID, loyalty program membership email, product price, and product quantity. By aggregation, five core behavioral dimensions were derived based on commonly used metrics in loyalty program research (Sharp & Sharp, 1997; Liu, 2007; Melnyk & Bijmolt, 2015). These are:

- Purchase frequency: number of purchases per customer;
- Items per transaction: average quantity of items per purchase;
- Total items purchased: total number of items purchased per customer;
- Spending per transaction: average monetary value per purchase;
- Total spending: total expenditure per customer.

As this study focuses on the pre- and post-enrolment purchase pattern of a customer, the purchasing history of each member was split accordingly using internal program variables (XY_ID, XY_EMAIL) and verified through long-term indicators (ID_INVOLVED, EMAIL_INVOLVED).

Parametric paired *t*-tests were used to compare pre- and post-enrolment customer purchasing behavior across the analyzed metrics, given their relevance, widespread use in economic research, and suitability for within-subject evaluations of mean differences (Park, 2009; Elliott & Woodward, 2007). The normality of paired differences was assessed using the Lilliefors test, which is commonly applied in empirical

research and is appropriate for use across different sample sizes (Abdi & Molin, 2007; Gel et al., 2007). The assumption of normality was rejected for all analyzed paired differences, with descriptive statistics presented in Table 1.

Nevertheless, paired *t*-tests were retained as the primary inferential procedure because the hypotheses concern changes in mean customer behavior and because, in large samples, the paired *t*-test is generally considered robust to violations of normality. Under the central limit theorem, the sampling distribution of the mean difference tends toward normality even when the underlying data is not normally distributed (Fagerland, 2012; Rasch et al., 2007; Lumley et al., 2002).

TABLE 1: Normality assessment of paired differences across sample variants

| Hypothesis | Sample variant | Mean diff. | Median diff. | SD diff. | Lilliefors test (normality decision) |
|------------|---------------------------------|------------|--------------|------------|--------------------------------------|
| H1 | Full sample | -0.04 | 0.1 | 1.02 | rejected |
| | Outliers excluded | 0 | 0.1 | 0.71 | rejected |
| | Seasonality excluded | -0.05 | 0 | 0.75 | rejected |
| | Outliers & seasonality excluded | -0.02 | 0 | 0.56 | rejected |
| H2 | Full sample | -0.1 | 0.1 | 2.24 | rejected |
| | Outliers excluded | -0.01 | 0.1 | 1.14 | rejected |
| | Seasonality excluded | -0.06 | 0.1 | 4.55 | rejected |
| | Outliers & seasonality excluded | -0.03 | 0.1 | 1.50 | rejected |
| H3 | Full sample | -0.15 | 0.2 | 4.82 | rejected |
| | Outliers excluded | -0.08 | 0.2 | 2.97 | rejected |
| | Seasonality excluded | -0.12 | 0.1 | 6.76 | rejected |
| | Outliers & seasonality excluded | -0.12 | 0.1 | 2.79 | rejected |
| H4 | Full sample | -2399.52 | 245.3 | 64365.01 | rejected |
| | Outliers excluded | -152.5 | 246.7 | 7395.15 | rejected |
| | Seasonality excluded | 7185.32 | 297 | 648340.94 | rejected |
| | Outliers & seasonality excluded | 914.36 | 297 | 38644.49 | rejected |
| H5 | Full sample | -2822.2 | 474 | 79377.70 | rejected |
| | Outliers excluded | -181.22 | 479.5 | 10074.13 | rejected |
| | Seasonality excluded | 15193.9 | 418.6 | 1184886.59 | rejected |
| | Outliers & seasonality excluded | 698.79 | 418.2 | 78190.96 | rejected |

Source: Data obtained from a proprietary database (name withheld due to confidentiality agreement). Table created by the authors.

All statistical tests were conducted at the conventional 5% significance level ($\alpha=0.05$). This threshold is standard across economic and marketing research, balancing the risks of Type I and Type II errors (Knapp, 2015; Maier & Lakens, 2022). To address the distributional concerns identified in the normality tests, Wilcoxon signed-rank tests were additionally performed as nonparametric robustness checks ("Wilcoxon signed ranks test," 2007; Davies & Kim, 2009). As a rank-based procedure evaluating the distribution of paired differences rather than mean differences, the Wilcoxon test does not necessarily lead to conclusions identical to those of the paired *t*-test, particularly in the presence of strongly skewed transactional data.

3.3. Robustness checks and sample adjustments

To ensure the robustness of the findings, all tests were conducted on four separate samples. The first was an unaltered whole sample, while the other three were carefully prepared. The second sample excluded transactions with extreme values to improve data robustness and mitigate the effects of outliers (Agrell & Niknazar, 2014; Wei & Yang, 2012). The third sample considered seasonality and removed transactions dating from November 1 to December 24 to control for holiday-related purchasing behavior and mitigate bias (Lu & White, 2014; Catherine et al., 2022). As seasonal effects are known to distort regular purchasing patterns (Basker, 2005), the final sample was the most conservative, excluding both outliers and pre-Christmas transactions to mitigate both effects and increase the sample's validity (Lu & White, 2014; Wei & Yang, 2012).

All variants were analyzed using the same statistical procedures to ensure the robustness of results across different data treatments. This methodological design is aimed at providing a rigorous and replicable framework for analyzing the purchase patterns of customers pre- and post-enrolment in a paid loyalty program.

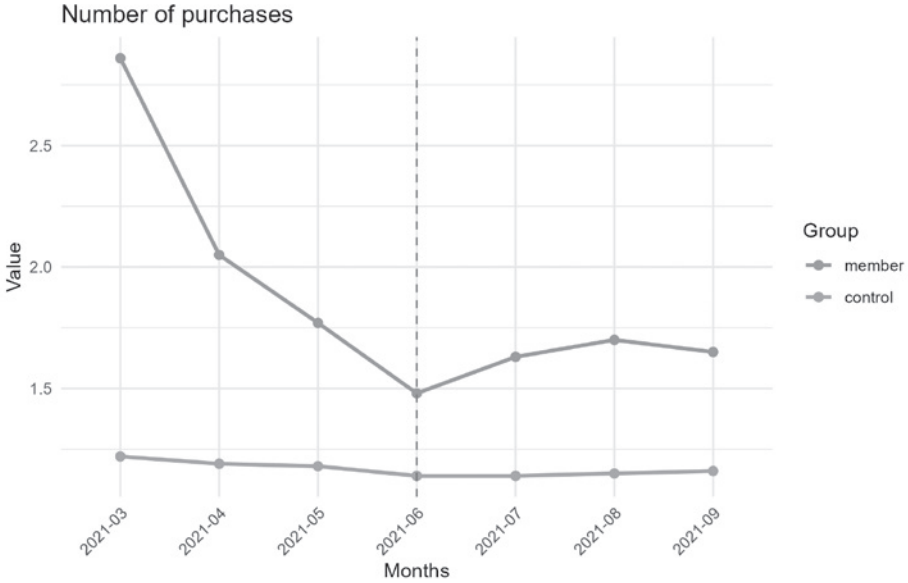
4. RESULTS

Prior to the within-customer analysis, a descriptive comparison between loyalty program members and non-members was conducted to illustrate general purchasing trends.

Specifically, customers active around June 2021 were divided into two groups: those who joined the paid loyalty program in that month and those who remained non-members throughout the observation period. Graphs 2-4 display the trajectories of purchase frequency, number of items purchased, and spending for both groups over a six-month period surrounding the enrolment month. This graphical comparison provides model-free descriptive evidence of purchasing trajectories for customers who joined the loyalty program and those who remained non-members.

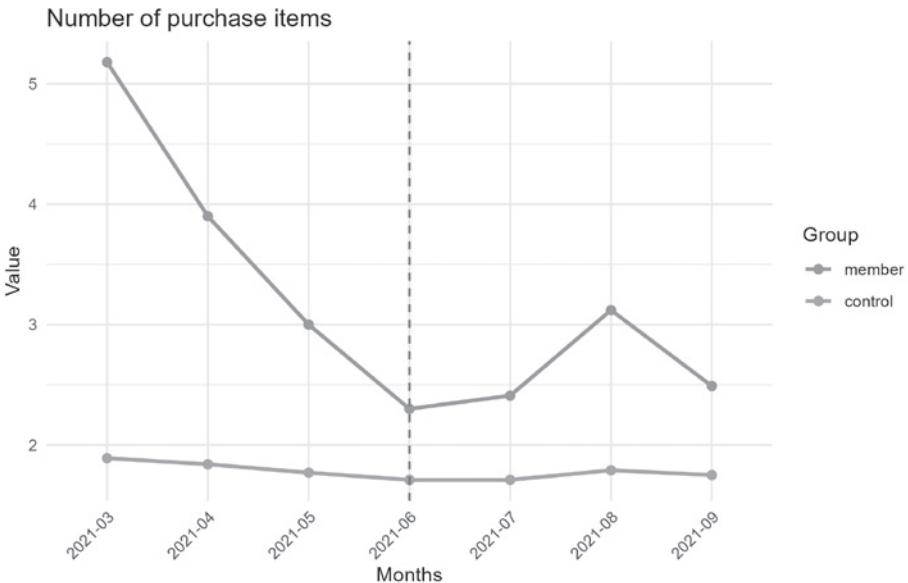
The descriptive comparison indicates that customers who joined the loyalty program exhibited consistently higher purchasing activity than non-members both before and after enrolment. At the same time, the trajectories of both groups remain relatively stable over time. The evidence suggests that differences between members and non-members may partly reflect self-selection bias.

GRAPH 2: Descriptive comparison of purchase frequency trajectories for paid loyalty program members and non-members



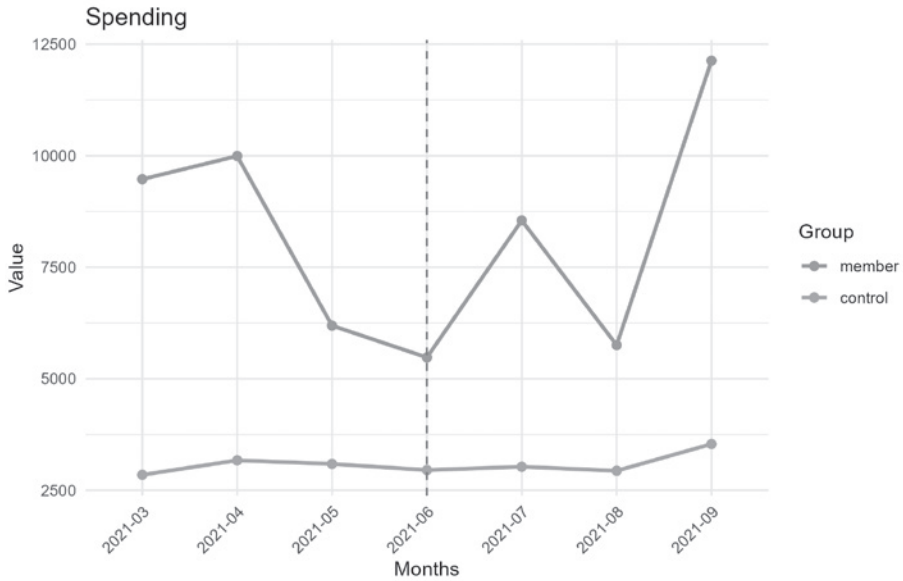
Source: Data obtained from a proprietary database (name withheld due to confidentiality agreement). Graph created by the authors.

GRAPH 3: Descriptive comparison of the trajectories of the number of items purchased by paid loyalty program members and non-members



Source: Data obtained from a proprietary database (name withheld due to confidentiality agreement). Graph created by the authors.

GRAPH 4: Descriptive comparison of the spending trajectories for paid loyalty program members and non-members



Source: Data obtained from a proprietary database (name withheld due to confidentiality agreement). Graph created by the authors.

A summary of the results for each hypothesis, analyzed by paired *t*-tests utilizing the most conservative sample, is presented in Table 2.

TABLE 2: Results of hypothesis tests using the sample with outliers and seasonal anomalies removed

| Hypothesis | Test statistic | <i>p</i> -value | Confidence interval | Central tendency (before) | Central tendency (after) | Supported |
|------------|----------------|-----------------|---------------------|---------------------------|--------------------------|-----------|
| H1 | -2.489 | 0.006 | (-Inf, -0.01) | 1.59 | 1.60 | Yes |
| H2 | -1.117 | 0.132 | (-Inf, 0.01) | 2.62 | 2.62 | No |
| H3 | -3.388 | 0.001 | (-Inf, -0.05) | 4.21 | 4.31 | Yes |
| H4 | 2.2 | 0.014 | (230.33, Inf) | 4703.73 | 3790.43 | Yes |
| H5 | 0.831 | 0.203 | (-683.52, Inf) | 7671.32 | 6973.19 | No |

Source: Data obtained from a proprietary database (name withheld due to confidentiality agreement). Table created by the authors.

Hypothesis H1 predicts a more frequent post-enrolment customer purchase compared to pre-enrolment in a paid loyalty program. While testing the whole sample, no statistically significant difference in purchasing frequency was observed between pre- and post-enrolment, indicating no statistically significant difference between the two periods. Testing the sample with outliers removed revealed a statistically significant difference, albeit in the opposite direction from what was expected. Members purchased less frequently after enrolling in the program than before. In contrast, the final sample variants, with the pre-Christmas period removed, yielded statistically significant differences between the two periods in the expected direction, suggesting that, after removing the pre-Christmas period, purchase frequency was higher compared to the pre-enrolment period.

Hypothesis H2 posits that customers purchase fewer items per transaction after joining a paid loyalty program than before. Analysis of the whole sample yielded a statistically significant difference between the number of items per transaction; however, the difference was in the opposite direction from that predicted by the hypothesis. In contrast, the other three samples did not indicate any statistically significant differences. Descriptively, in two instances, the means of both groups were identical. The other sample, with only the pre-Christmas period removed, indicated a negligible increase in the post-enrolment mean value. Overall, Hypothesis H2 is not supported.

Hypothesis H3 predicts that customers will purchase a greater number of items during the post-enrolment period in a paid loyalty program than before enrolment. Analysis of the whole sample and the sample with the pre-Christmas period removed did not provide a statistically significant difference in means, although descriptive statistics indicated a marginal increase

in mean values. Furthermore, testing the sample with only the outliers removed yielded no statistically significant difference either. However, in this case, the number of items purchased decreased marginally. Conversely, the most conservative sample showed a statistically significant difference in the expected direction, indicating that the number of items customers purchase after they enroll in the paid loyalty program increased.

Hypothesis H4 posits that customers spend less money per transaction after enrolling in a paid loyalty program than before. None of the results from the first three sample variants yielded statistically significant differences although, descriptively, two of the tests indicated a drop in post-enrolment mean values. Nevertheless, when it comes to the conservative sample, with outliers and the pre-Christmas period removed, there was a statistically significant difference between the two groups in the expected direction, suggesting that post-enrolment spending per transaction decreased.

The final hypothesis, H5, predicts that customers will spend more money after enrolling in a paid loyalty program than before. In this case, only one sample variant—without outliers—provided a statistically significant difference. However, this was in the opposite direction to what was predicted, suggesting that post-enrolment customer spending decreased marginally. The other three tests were not statistically significant. Descriptively, the means of the results from the whole sample indicated an increase in expenditure, while the two samples without the pre-Christmas period displayed a decrease in expenditure. Overall, hypothesis H5 is not supported.

For transparency, Table 3 summarizes the results of the five hypotheses for each tested sample variant.

TABLE 3: Complete overview of tests conducted on all data sample variants

| Hypothesis | Sample variant | Test statistic | p-value | Confidence interval | Central tendency (before) | Central tendency (after) | Supported |
|------------|------------------------------------|----------------|---------|---------------------|---------------------------|--------------------------|-----------|
| H1 | Whole | 0.326 | 0.372 | (-0.02, Inf) | 1.92 | 1.92 | No |
| | Without outliers | 2.621 | 0.004 | (0.01, Inf) | 1.88 | 1.85 | No |
| | Without pre-Christmas | -5.574 | 0.001 | (-Inf, -0.03) | 1.62 | 1.67 | Yes |
| | Without outliers and pre-Christmas | -2.489 | 0.006 | (-Inf, -0.01) | 1.59 | 1.60 | Yes |
| H2 | Whole | -1.962 | 0.025 | (-Inf, -0.01) | 2.69 | 2.73 | No |
| | Without outliers | 0.014 | 0.494 | (-0.03, Inf) | 2.65 | 2.65 | No |
| | Without pre-Christmas | -0.91 | 0.181 | (-Inf, 0.04) | 2.68 | 2.72 | No |
| | Without outliers and pre-Christmas | -1.117 | 0.132 | (-Inf, -0.01) | 2.62 | 2.62 | No |
| H3 | Whole | -0.126 | 0.449 | (-Inf, 0.09) | 5.19 | 5.20 | No |
| | Without outliers | 0.313 | 0.377 | (-0.05, Inf) | 5.01 | 5.00 | No |
| | Without pre-Christmas | -1.409 | 0.079 | (-Inf, 0.02) | 4.37 | 4.47 | No |
| | Without outliers and pre-Christmas | -3.388 | 0.001 | (-Inf, -0.05) | 4.21 | 4.31 | Yes |
| H4 | Whole | -0.108 | 0.457 | (-Inf, 321.8) | 3291.09 | 3313.75 | No |
| | Without outliers | 1.587 | 0.056 | (-2.51, Inf) | 2923.80 | 2854.68 | No |
| | Without pre-Christmas | 1.032 | 0.151 | (-4269.35, Inf) | 14244.20 | 7059.94 | No |
| | Without outliers and pre-Christmas | 2.2 | 0.014 | (230.33, Inf) | 4703.73 | 3790.43 | Yes |
| H5 | Whole | -0.621 | 0.267 | (-Inf, 1048.45) | 6616.92 | 7252.99 | No |
| | Without outliers | 1.764 | 0.038 | (13.17, Inf) | 5608.71 | 5413.46 | No |
| | Without pre-Christmas | 1.194 | 0.116 | (-5739, Inf) | 24835.96 | 9642.72 | No |
| | Without outliers and pre-Christmas | 0.831 | 0.203 | (-683.52, Inf) | 7671.32 | 6973.19 | No |

Source: Data obtained from a proprietary database (name withheld due to confidentiality agreement). Table created by the authors.

To address the distributional concerns identified in the normality tests, Wilcoxon signed-rank tests were conducted as a nonparametric robustness check. The results of these tests are presented in Table 4. While the paired t-tests based on the main analytical sample indicate statistically significant changes in several purchasing behavior metrics following enrolment in the paid loyalty program, the Wilcoxon signed-rank tests do not uniformly confirm these findings.

free-to-join and paid programs, which suggests that members exhibit higher levels of repeat patronage (Liu, 2007; Ashley et al., 2016). Chaudhuri et al. (2019) explain this effect by arguing that customers who pay a membership fee are psychologically motivated to make use of the program. This suggests that paid loyalty programs may effectively drive habitual purchasing behavior and increase the purchase frequency among members.

TABLE 4: Wilcoxon signed-rank robustness tests for pre- and post-enrolment purchasing Behavior conducted on the data sample variant with outliers and seasonal anomalies removed

| Hypothesis | Test statistic | p-value | Confidence interval | Median difference | Supported |
|------------|----------------|---------|---------------------|-------------------|-----------|
| H1 | 17172276 | 0.002 | (0.01, Inf) | 0.05 | No |
| H2 | 19117719.5 | 0.001 | (0.03, Inf) | 0.07 | Yes |
| H3 | 18398731 | 0.154 | (-0.01, Inf) | 0.12 | No |
| H4 | 22392672 | 0.001 | (249.7, Inf) | 296.95 | Yes |
| H5 | 21619212 | 0.001 | (326.5, Inf) | 418.2 | No |

Source: Data obtained from a proprietary database (name withheld due to confidentiality agreement). Table created by the authors.

5. DISCUSSION

This paper was aimed at investigating whether customers change their purchasing behavior after joining a paid loyalty program within the Czech e-commerce market. An analysis on a sample of over one million individual transaction records aggregated to 21,805 unique customers revealed significant differences across certain behavioral dimensions. The following section reviews the results for each hypothesis separately.

The results support Hypothesis H1, indicating that customers purchase more frequently after joining the paid loyalty program than before. The findings were consistent when tested on the two samples that excluded the pre-Christmas period. The observed increase in purchase frequency among paying loyalty program members aligns with existing literature on both

The results for Hypothesis H2 contradicted initial expectations. Prior research anticipated that customers who join a paid loyalty program offering free shipping would purchase fewer items per order because they would no longer be constrained by shipping costs (Guo & Liu, 2023). However, analyses across all four sample variants revealed the opposite trend. After joining the program, customers actually purchased the same number or slightly more items per order than before. This finding partially aligns with broader research, indicating that loyalty programs positively influence sales (Bombajj & Dekimpe, 2020). This is favorable for paid loyalty program providers as it suggests that customers, despite becoming paying members, may not inherently take advantage of the free shipping offers and may even increase their basket size after joining.

Hypothesis H3 was supported by the most conservative sample variant, indicating that customers purchase more items overall after joining the paid loyalty program than before. This suggests that paid loyalty program members not only purchase more frequently but also buy a higher total volume of products compared to their pre-membership period. These findings are consistent with existing research, indicating that long-term loyalty program engagement leads to higher cumulative purchases (Meyer-Waarden, 2008). A similar pattern has been observed in the airline industry, where long-term members demonstrate greater activity and are more strongly committed to the brand (Vilkaite-Vaitone & Papsiene, 2016). The membership fee may contribute to this outcome by strengthening psychological commitment to a single e-commerce brand. From a managerial perspective, the results highlight the strategic potential of paid loyalty programs to enhance customer value and strengthen behavioral loyalty.

The results for Hypothesis H4 were also supported only by the most conservative sample variant. It was hypothesized that free shipping would encourage members to place more frequent, lower-value orders rather than increase purchases to minimize delivery costs. However, the results do not fully support this assumption. The hypothesis proposed that members would spend less per transaction after joining the paid loyalty program. At the same time, members were found to order more frequently and purchase the same number or slightly more items per order. Taken in isolation, this combination of outcomes may reflect the influence of additional factors not directly captured in the dataset. Other factors, such as targeted promotions or higher-value item purchases, may also have influenced transaction values. From a business perspective, the findings imply that although transaction value may decrease after joining, the primary benefit of paid loyalty programs lies in stimulating purchase frequency and overall purchase volume rather than increasing spend per transaction.

The outcomes of Hypothesis H5 do not support the prediction that customers will spend more overall after joining a paid loyalty program. In three of the sample variants, the null hypothesis could not be rejected, suggesting no significant difference in spending after joining. One test suggests that customers may even spend slightly less. These findings do not align precisely with current research suggesting that joining a loyalty program, whether traditional or fee-based, correlates with a higher annual customer value (Ashley et al., 2016). However, the definition of customer value is important. While total expenditure does not increase, purchase frequency and the number of items purchased show a positive change after joining. For businesses, this demonstrates that membership fees in loyalty programs are unlikely to be the primary driver of increased total expenditure.

Furthermore, an additional consideration arises from the nonparametric robustness tests conducted using the Wilcoxon signed-rank procedure. While the primary paired *t*-tests indicate statistically significant changes in several purchasing behavior metrics following enrolment in the paid loyalty program, the Wilcoxon tests do not uniformly confirm these results. This difference likely reflects the strongly skewed nature of transactional purchasing data. Whereas the paired *t*-test evaluates mean changes in customer behavior, the Wilcoxon procedure relies on ranked paired differences and is therefore less sensitive to large behavioral shifts among a minority of customers. The results therefore suggest that the observed post-enrolment behavioral changes may not be consistently distributed across all customers but may instead reflect stronger changes among certain customer segments.

6. CONCLUSIONS, LIMITATIONS, AND FUTURE RESEARCH

The core focus of this paper was to examine whether customers behave differently after

joining a paid loyalty program compared with their behavior prior to membership. This was addressed through a longitudinal quantitative analysis of purchasing behavior in the Czech e-commerce market on data provided by a major e-commerce company operating at the top in its field. Results from the most conservative sample indicate that customers exhibit higher behavioral engagement after joining a paid loyalty program. Specifically, they shop more frequently and purchase more items overall than before becoming members. Although members benefit from free shipping, they do not appear to overtly exploit this advantage by deliberately fragmenting their orders, at least not at a scale detectable in the analysis. Furthermore, while members spend less per transaction, their overall expenditure remains relatively stable.

These findings contribute to the existing body of research on loyalty programs and extend empirical evidence on paid loyalty programs within European contexts and the Czech Republic. In addition to providing new empirical evidence, the results support current understanding that paid loyalty programs can enhance overall customer value. The study offers practical insights for both academic researchers and marketing practitioners. Importantly, it goes beyond simply demonstrating that members differ from non-members; rather, it shows that customer behavior changes after joining a paid loyalty program. The findings also expand on the notion of self-selection bias, suggesting that members may have exhibited some of these behaviors prior to joining.

The study is limited by its focus on a single company and industry. Although the data provider is a major e-commerce player in the Czech Republic with a significant market share, the consumer data reflects customer behavior from one business offering a wide variety of products (similar to Amazon). Additionally, the researchers had no involvement in the sample extraction process. Although the data provider stated that

computer-based random sampling was used, there is no independent verification of the data not being altered in ways that could enhance the company's performance. Lastly, due to EU data protection regulations, data on former members could not be tracked separately so it is mixed with non-members' data. The proportion of such customers is unknown.

A further limitation relates to the distributional characteristics of the transactional purchasing data. Although the primary analyses rely on paired *t*-tests to evaluate mean changes in customer behavior, the nonparametric Wilcoxon signed-rank tests used as robustness checks did not uniformly confirm these results. This suggests that behavioral changes following enrolment in the paid loyalty program may not be evenly distributed across the entire sample. Instead, the observed effects may be driven by stronger responses among particular customer segments.

Future research should examine whether these findings can be generalized across different industries and international contexts. Additionally, loyalty programs vary in design and reward mechanisms, which may influence behavioral outcomes. Furthermore, consumer behavior may differ across cultural contexts. As this study focused exclusively on paid loyalty programs in the Czech Republic, it does not allow for comparison with traditional programs that, for example, require purchase thresholds. Addressing this limitation would require the acquisition of additional data to enable comparison with a traditional program or to contrast the current findings with those of a paid loyalty program operating in a different market. Additionally, future research should consider the strongly skewed nature of transactional purchasing data and apply distribution-sensitive analytical approaches.

AI-ASSISTED STATEMENT – AI was used for grammar, spelling, and references finalization.

REFERENCES

1. Abdi, H., & Molin, P. (2007). Lilliefors test for normality. In: N. J. Salkind (Ed.), *Encyclopedia of measurement and statistics*. SAGE Publications. <https://doi.org/10.4135/9781412952644.n256>
2. Agrell, P. J., & Niknazar, P. (2014). Structural and Behavioral Robustness in Applied Best-Practice Regulation. *Socio-Economic Planning Sciences*, 48(1), 89–103. <https://doi.org/10.1016/j.seps.2013.12.004>
3. Amazon. (2025). Amazon Prime benefits & membership details. Retrieved November 16, 2024, from <https://www.amazon.com/gp/help/customer/display.html?nodeId=G6LDPN7YJHYKH2J6>
4. Arbore, A., & Estes, Z. (2013). Loyalty Program Structure and Consumers' Perceptions of Status: Feeling Special in a Grocery Store?. *Journal of Retailing and Consumer Services*, 20(5), 439–444. <https://doi.org/10.1016/j.jretconser.2013.03.002>
5. Ashley, C., Gillespie, E. A., & Noble, S. M. (2016). The Effect of Loyalty Program Fees on Program Perceptions and Engagement. *Journal of Business Research*, 69(2), 964–973. <https://doi.org/10.1016/j.jbusres.2015.09.001>
6. Basker, E. (2005). 'Twas Four Weeks Before Christmas: Retail Sales and the Length of the Christmas Shopping Season. *Economics Letters*, 89(3), 317–322. <https://doi.org/10.1016/j.econlet.2005.06.004>
7. Bombajj, N. J. F., & Dekimpe, M. G. (2020). When Do Loyalty Programs Work? The Moderating Role of Design, Retailer-Strategy, and Country Characteristics. *International Journal of Research in Marketing*, 37(1), 175–195. <https://doi.org/10.1016/j.ijresmar.2019.07.003>
8. Breugelmans, E., Bijmolt, T. H. A., Zhang, J., Basso, L. J., Dorotic, M., Kopalle, P., Minnema, A., Mijnlieff, W. J., & Wunderlich, N. V. (2015). Advancing Research on Loyalty Programs: A Future Research Agenda. *Marketing Letters*, 26, 127–139. <https://doi.org/10.1007/s11002-014-9311-4>
9. Catherine, S., Ebrahimian, M., Fereydounian, M., Sraer, D. A., & Thesmar, D. (2022). Robustness Checks in Structural Analysis [Working paper]. *Social Science Research Network*. <https://dx.doi.org/10.2139/ssrn.4206149>
10. Chaudhuri, M., Voorhees, C. M., & Beck, J. M. (2019). The Effects of Loyalty Program Introduction and Design on Short- and Long-Term Sales and Gross Profits. *Journal of the Academy of Marketing Science*, 47, 640–658. <https://doi.org/10.1007/s11747-019-00652-y>
11. Colloquy. (2018). *2017 Colloquy Loyalty Census* [Report]. LoyaltyOne. Retrieved December 8, 2020, from <https://www.colloquy.com/reports/>
12. Davies, R. J., & Kim, S. S. (2009). Using matched samples to test for differences in trade execution costs. *Journal of Financial Markets*, 12(2), 173–202. <https://doi.org/10.1016/j.finmar.2008.06.001>
13. Dorotic, M., Fok, D., Verhoef, P. C., & Bijmolt, T. H. A. (2021). Synergistic and Cannibalization Effects in a Partnership Loyalty Program. *Journal of the Academy of Marketing Science*, 49, 1021–1042. <https://doi.org/10.1007/s11747-020-00759-7>
14. Dorotic, M., Verhoef, P. C., Fok, D., & Bijmolt, T. H. A. (2014). Reward redemption effects in a loyalty program when customers choose how much and when to redeem. *International Journal of Research in Marketing*, 31(4), 339–355. <https://doi.org/10.1016/j.ijresmar.2014.06.001>
15. Dorotic, M., Bijmolt, T. H. A., & Verhoef, P. C. (2012). Loyalty Programmes: Current Knowledge and Research Directions. *International Journal of Management Reviews*, 14(3), 217–237. <https://doi.org/10.1111/j.1468-2370.2011.00314.x>
16. Du, S., Li, Z., Choi, T.-M., & Hu, L. (2024). *Optimal bundling cooperation strategies for online paid membership programs*. *Naval Research Logistics*, 72, 390–408. <https://doi.org/10.1002/nav.22232>
17. Elliott, A. C., & Woodward, W. A. (2007). Comparing one or two means using the t-test. In: *Statistical Analysis Quick Reference Guide-*

- book (pp. 3–14). SAGE Publications. <https://doi.org/10.4135/9781412985949.n3>
18. Esmark, C. L., Noble, S. M., & Bell, J. E. (2016). Open Versus Selective Customer Loyalty Programmes. *European Journal of Marketing*, 50(5/6), 770–795. <https://doi.org/10.1108/EJM-11-2014-0730>
 19. Fagerland, M. W. (2012). t-tests, non-parametric tests, and large studies - A paradox of statistical practice? *BMC Medical Research Methodology*, 12, Article: 78. <https://doi.org/10.1186/1471-2288-12-78>
 20. Faria, J. (2025, November 19). *Average number of loyalty programs U.S. consumers belong to* (Statista Statistic No. 618744). Statista. <https://www.statista.com/statistics/618744/average-number-of-loyalty-programs-us-consumers-belong-to/>
 21. Gel, Y. R., Miao, W., & Gastwirth, J. L. (2007). Robust Directed Tests of Normality Against Heavy-Tailed Alternatives. *Computational Statistics & Data Analysis*, 51(5), 2734–2746. <https://doi.org/10.1016/j.csda.2006.08.022>
 22. Guo, F., & Liu, Y. (2023). The Effectiveness of Membership-Based Free Shipping: A Field Study of an Online Retailer. *Journal of Marketing*, 87(6), 869–888. <https://doi.org/10.1177/00222429231158116>
 23. Henderson, C. M., Beck, J. T., & Palmatier, R. W. (2011). Review of the Theoretical Underpinnings of Loyalty Programs. *Journal of Consumer Psychology*, 21(3), 256–276. <https://doi.org/10.1016/j.jcps.2011.02.007>
 24. Hwang, E., Baloglu, S. & Tanford, S. (2019). Building Loyalty Through Reward Programs: The Influence of Perceptions of Fairness and Brand Attachment. *International Journal of Hospitality Management*, 76, 19–28. <https://doi.org/10.1016/j.ijhm.2018.03.009>
 25. Jin, X., Kim, T., & Lee, D. (2024). Spending Points During Crises: Adaptive Behavior on E-Loyalty Programs. *Electronic Commerce Research and Applications*, 65, Article: 101408. <https://doi.org/10.1016/j.elerap.2024.101408>
 26. Knapp, T. R. (2015). In (Partial) Defense of .05. *Journal of Modern Applied Statistical Methods*, 14(2). <https://doi.org/10.22237/jmasm/1446350700>
 27. Latham, G. P., & Locke, E. A. (1991). Self-regulation Through Goal Setting. *Organizational Behavior and Human Decision Processes*, 50(2), 212–247. [https://doi.org/10.1016/0749-5978\(91\)90021-K](https://doi.org/10.1016/0749-5978(91)90021-K)
 28. Leenheer, J., & Bijmolt, T. H. A. (2008). Which Retailers Adopt a Loyalty Program? An Empirical Study. *Journal of Retailing and Consumer Services*, 15(6), 429–442. <https://doi.org/10.1016/j.jretconser.2007.11.005>
 29. Leenheer, J., van Heerde, H. J., Bijmolt, T. H. A., & Smidts, A. (2007). Do Loyalty Programs Really Enhance Behavioral Loyalty? *International Journal of Research in Marketing*, 24(1), 31–47. <https://doi.org/10.1016/j.ijresmar.2006.10.005>
 30. Li, X., Ma, B., & Zhou, C. (2017). Effects of Customer Loyalty on Customer Entitlement and Voiced Complaints. *The Service Industries Journal*, 37(13–14), 858–874. <https://doi.org/10.1080/02642069.2017.1360290>
 31. Liu, Y. (2007). The Long-Term Impact of Loyalty Programs on Consumer Purchase Behavior and Loyalty. *Journal of Marketing*, 71(4), 19–35. <https://doi.org/10.1509/jmkg.71.4.019>
 32. Lu, X., & White, H. (2014). Robustness Checks and Robustness Tests in Applied Economics. *Journal of Econometrics*, 178(1), 194–206. <https://doi.org/10.1016/j.jeconom.2013.08.016>
 33. Lumley, T., Diehr, P., Emerson, S., & Chen, L. (2002). The importance of the normality assumption in large public health data sets. *Annual Review of Public Health*, 23, 151–169. <https://doi.org/10.1146/annurev.publhealth.23.100901.140546>
 34. Maier, M., & Lakens, D. (2022). Justify Your Alpha: A Primer on Two Practical Approaches. *Advances in Methods and Practices in Psychological Science*, 5(2), 1–14. <https://doi.org/10.1177/25152459221080396>
 35. Median, a.s. (2008–2022). *MML-TGI Czech Republic dataset* [Dataset]. Available from the Prague University of Economics and Business under institutional license.

36. Melnyk, V., & Bijmolt, T. (2015). The Effects of Introducing and Terminating Loyalty Programs. *European Journal of Marketing*, 49(3–4), 398–419. <https://doi.org/10.1108/EJM-12-2012-0694>
37. Meyer-Waarden, L. (2007). The Effects of Loyalty Programs on Customer Lifetime Duration and Share of Wallet. *Journal of Retailing*, 83(2), 223–236. <https://doi.org/10.1016/j.jretai.2007.01.002>
38. Meyer-Waarden, L. (2008). The Influence of Loyalty Program Membership on Customer Purchase Behavior. *European Journal of Marketing*, 42(1/2), 87–114. <https://doi.org/10.1108/03090560810840925>
39. Miles & More GmbH. (n.d.). *Programme. Miles & More Lufthansa*. Retrieved November 16, 2024, from <https://www.miles-and-more.com/row/en/program/at-a-glance.html>
40. Park, H. M. (2009). *Comparing group means: T-tests and one-way ANOVA using Stata, SAS, R, and SPSS*. Indiana University. Retrieved November 16, 2024 from <https://scholarworks.iu.edu/dspace/items/2acf01f0-0898-4250-b38a-1d1a77a25b5e>
41. Powell, S. (2022, July 25). *German members of parliament no longer receive complimentary Lufthansa Senator status as of this year*. *LoyaltyLobby*. <https://loyaltylobby.com/2022/07/25/german-members-of-parliament-no-longer-receive-complementary-lufthansa-senator-status-as-of-this-year/>
42. Rasch, D., Teuscher, F., & Guiard, V. (2007). How robust are tests for two independent samples?. *Journal of Statistical Planning and Inference*, 137(8), 2706–2720. <https://doi.org/10.1016/j.jspi.2006.04.011>
43. Sharp, B., & Sharp, A. (1997). Loyalty Programs and Their Impact on Repeat-Purchase Loyalty Patterns. *International Journal of Research in Marketing*, 14(5), 473–486. [https://doi.org/10.1016/S0167-8116\(97\)00022-0](https://doi.org/10.1016/S0167-8116(97)00022-0)
44. Stauss, B., Schmidt, M., & Schoeler, A. (2005). Customer Frustration in Loyalty Programs. *International Journal of Service Industry Management*, 16(3), 229–252. <https://doi.org/10.1108/09564230510601387>
45. The Economist. (2017). *Beyond the points: The next generation of customer loyalty*. *The Economist*. Retrieved December 4, 2020, from <https://expectexceptional.economist.com/beyond-the-points-customer-loyalty.html>
46. United States Census Bureau. (2025). *U.S. and world population clock*. *United States Census Bureau*. Retrieved May 15, 2025, from <https://www.census.gov/popclock/>
47. Vilkaite-Vaitone, N., & Papsiene, P. (2016). Influence of Customer Loyalty Program on Organizational Performance: A Case of the Airline Industry. *Inzinerine Ekonomika-Engineering Economics*, 27(1), 109–116. <https://doi.org/10.5755/j01.ee.27.1.10267>
48. Vroom, V. H. (1964). *Work and motivation*. John Wiley & Sons.
49. Walsman, M. C., & Dixon, M. J. (2020). Fee-Based Loyalty Programs: An Empirical Investigation of Benefit Redemption Behavior and its Effects on Loyalty. *Service Science*, 12(2–3), 100–118. <https://doi.org/10.1287/serv.2020.0260>
50. Wang, Y., Lewis, M., Cryder, C., & Sprigg, J. (2016). Enduring Effects of Goal Achievement and Failure Within Customer Loyalty Programs: A Large-Scale Field Experiment. *Marketing Science*, 35(4), 565–575. <https://doi.org/10.1287/mksc.2015.0966>
51. Wei, X., & Yang, Y. (2012). Robust Forecast Combinations. *Journal of Econometrics*, 166(2), 224–236. <https://doi.org/10.1016/j.jeconom.2011.09.035>
52. Wilcoxon signed ranks test. (2007). In N. J. Sal- kind (Ed.), *Encyclopedia of measurement and statistics* (pp. 1052–1053). Sage Publications. <https://doi.org/10.4135/9781412952644.n484>
- This is an open access article under the CC BY-NC license (<https://creativecommons.org/licenses/by-nc/4.0/>)