International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Factor Analysis of Telecommunications Service Attributes Identifying Latent Constructs Driving Customer Behavior and Satisfaction

K. Balakrishna[1], K Sekhar [2]

[1,2]Assistant professor in statistics, Siva Sivani Degree college, Hyderabad

Abstract

The study leverages the multivariate statistical technique of factor analysis to uncover the fundamental latent constructs that underlie key telecommunications service attributes and govern customer behavior and satisfaction. Utilizing a robust dataset of 250 customers, which includes variables such as average monthly call minutes, bill amount, percentage of business usage, length of service, household income, and a proprietary propensity-to-leave score, the research aims to distill these observable metrics into a simpler, more powerful framework of underlying drivers. The analysis successfully extracted two distinct components that collectively explain 60.875% of the total variance in the data. The first construct, identified as "Service Utilization and Loyalty," reveals a critical inverse relationship between call volume and the propensity to churn, highlighting that high-engagement users are fundamentally more loyal. The second construct, termed "Established Value and Business Engagement," characterizes a segment of long-tenured, business-oriented customers who generate significant revenue and represent a stable, high-value clientele. These findings move beyond isolated variable analysis to provide a synthesized, actionable understanding of the customer base. The resulting framework offers telecom operators profound strategic insights for enhancing service design, implementing targeted retention programs, and optimizing customer satisfaction by focusing on the core latent constructs that drive both value and risk within their portfolio.

Introduction

The telecommunications industry operates within a paradigm of intense competition, relentless technological advancement, and critically high customer churn rates. In such a saturated and dynamic marketplace, a superficial understanding of customer metrics is insufficient for sustainable growth. Instead, achieving a deep, structural comprehension of the fundamental drivers that underpin customer behavior and satisfaction is not merely a competitive advantage—it is an essential prerequisite for corporate survival and long-term profitability.

Telecommunications companies are, by nature, vast repositories of customer data, routinely collecting granular information on usage patterns, billing history, and demographic profiles. However, the central challenge confronting managers and analysts lies in transcending this multitude of individual data points. The true value is unlocked not by examining variables like call minutes or bill amount in isolation, but by uncovering the underlying, or latent, constructs that synergistically shape customer decisions and loyalty. For instance, a customer with a high monthly bill but low usage may represent a fundamentally different archetype—perhaps one dissatisfied with price-to-value ratio—compared to a customer with a similarly

high bill driven by high usage, who may perceive greater value. Traditional univariate or bivariate analyses fail to capture these complex, interactive relationships.

To address this analytical gap, the study employs Factor Analysis, a robust multivariate statistical technique specifically designed for data reduction and structure detection. Its primary function is to condense a large set of observed and often correlated variables into a smaller, more manageable set of unobservable latent factors, thereby revealing the simple structure within complex phenomena.

Guided by this methodology, the primary objectives of this research are threefold:

1. To identify and extract the key latent constructs underlying six critical telecommunications service attributes: monthly call minutes, monthly bill amount, business usage percentage, length of service (tenure), household income, and propensity to leave.
2. To interpret the intrinsic nature of these constructs and elucidate their theoretical and practical relationship with customer satisfaction, which is measured inversely through the propensity-to-leave score.
3. To translate these data-driven insights into actionable strategic recommendations for telecom managers, empowering them to enhance customer retention, optimize service offerings, and allocate resources more effectively by targeting the root causes of customer behavior.

This paper is structured to provide a comprehensive exploration of this research endeavor. It begins with a review of relevant literature on customer satisfaction and factor analysis in marketing, followed by a detailed explanation of the research methodology. A comprehensive presentation and discussion of the data analysis and results will then be provided, culminating in a conclusion that outlines the study's implications, acknowledges its limitations, and suggests promising avenues for future research.

Literature Review

2.1 Determinants of Customer Satisfaction in Telecommunications

Gerpott et al. (2001) identified network quality, customer service, and pricing structure as primary influencers. More recently, studies have emphasized the role of data service quality and brand image (Aydin & Özer, 2005). Our study builds on this by integrating objective behavioral metrics (minutes, bill) with tenure and a direct measure of loyalty.

2.2 Customer Churn and Propensity to Leave

Customer churn, also referred to as customer attrition, represents one of the most critical and financially impactful metrics for telecommunications companies. In an industry characterized by high saturation and intense competition, the loss of a subscriber directly erodes the customer base and future revenue streams, making churn prevention a paramount strategic objective.

The academic literature provides clear guidance on effective churn management. Seminal research by Neslin et al. (2006) established that the most robust and accurate churn prediction models are those that synthesize a multifaceted view of the customer. These models achieve superior performance not by relying on a single data type, but by integrating three key categories of information:

1. Historical Usage Patterns**:** Changes in a customer's call duration, data consumption, or service activity, which can signal disengagement.
2. Customer Demographics**:** Background variables such as income, age, or geographic location, which can define distinct risk segments.
3. Service Changes**:** Events like plan downgrades or complaints to customer service, which often act as direct triggers for churn.

The study is methodologically aligned with this established research stream. The "propensity to leave" score within our dataset is not a measure of observed churn but a sophisticated, composite metric that functions as a leading indicator or a proxy. It quantifies a customer's latent dissatisfaction and their statistical likelihood of terminating service in the near future. By incorporating this score as our key dependent variable, we directly operationalize the concept of loyalty (inversely) in a manner consistent with best practices in the field. This allows us to move beyond simply analyzing past behavior and instead model the underlying drivers of future customer actions, thereby bridging the gap between descriptive analytics and predictive, proactive customer retention strategies.

2.3 Application of Factor Analysis in Marketing Research and Identification of Research Gap

Factor Analysis (FA) is a cornerstone multivariate technique in the marketer's toolkit, renowned for its ability to distill a large set of observable variables into a smaller number of underlying, latent dimensions, often referred to as factors or constructs (Hair et al., 2019). This data reduction capability is paramount in market research, where managers are often inundated with data but starved for insight. The primary utility of FA lies in its power to uncover the simple structure in complex phenomena, moving from a lengthy list of correlated metrics to a handful of fundamental drivers that govern consumer behavior.

The application of FA in marketing is both historical and diverse. In a seminal work, **Malhotra (1981)** demonstrated its utility for psychographic segmentation, using factor analysis to identify underlying lifestyle dimensions based on consumers' activities, interests, and opinions (AIO). This allowed for a move beyond simple demographics to segment markets based on shared psychological profiles, enabling more resonant and targeted messaging. In the specific context of telecommunications, **Kim and Yoon (2004)** employed FA to deconstruct the complex decision-making process of mobile carrier selection. Their analysis revealed that factors such as "value-added services," "network quality," and "price" were the latent constructs influencing customer choice, providing a structured understanding of competitive positioning.

While these studies exemplify the traditional application of FA—often focusing on attitudinal data (perceptions, preferences) or discrete choice drivers—our research applies this robust methodology to a different, yet critically important, type of data: operational and behavioral metrics**.** These are the hard, transactional data points that telecommunications companies collect as a matter of routine: call minutes, billing amounts, tenure, and business usage. The distinct contribution of this study is the integration of these operational metrics with a direct, forward-looking indicator of customer health: the **"**propensity to leave**"** score.

This integration addresses a critical research gap. Previous studies have often analyzed these variables in isolation or in limited pairs (e.g., correlating usage with churn). However, a holistic, simultaneous examination of how usage, spend, tenure, and customer type (business vs. personal) interact to define latent customer segments, and how these segments directly map onto a metric of loyalty/attrition, has been less explored. By submitting this specific combination of variables to factor analysis, we bridge a gap between operational data analysis and strategic customer management.

Therefore, the research gap this study addresses is not merely the *application* of factor analysis in telecom, but its application to synthesize operational behavior with loyalty metrics. The objective is to move beyond describing what customers *are* (high-users, long-tenured) to understanding what they *represent* in terms of underlying constructs (e.g., "The Loyal Engager," "The Established High-Value Client"). This shift from a siloed view to a "composite, actionable view of the customer base" empowers managers to

design segment-specific strategies based on a deep, data-driven understanding of the fundamental constructs that drive both value and risk within their customer portfolio.

Research Methodology

3.1 Data Collection and Variables

Data Source and Variable Description

The empirical foundation of this study is a curated dataset of historical customer records, sourced directly from the operational database of a major telecommunications service provider. This secondary data, comprising a robust sample of 250 customers, offers a realistic and actionable snapshot of a contemporary subscriber base. The sample size is sufficient for a factor analysis involving six variables, providing a stable basis for identifying underlying latent constructs.

The selection of variables was guided by both established marketing theory and pragmatic business relevance, aiming to capture the core dimensions of the customer-operator relationship. The six variables incorporated into the analysis are as follows:

1. **Average Monthly Minutes of Use (minutes):** This quantitative metric represents the total volume of voice calls made by the customer per month. It serves as a primary indicator of service engagement and usage intensity. High usage typically suggests a deeper integration of the service into the customer's daily routine, which theory suggests may be linked to higher switching costs and lower churn propensity.
2. **Average Monthly Bill Amount (bill):** This variable reflects the total revenue generated from the customer each month. Beyond mere revenue, it is a composite measure of the customer's economic value and plan tier. It encapsulates not only usage but also the pricing structure and any value-added services subscribed to, providing a direct link to customer profitability.
3. **Percentage of Usage for Business (business):** This metric distinguishes the primary nature of the customer's usage pattern. It indicates the proportion of total service consumption dedicated to business, as opposed to personal, purposes. This is a critical differentiator, as business users often exhibit distinct behaviors, including higher sensitivity to service reliability and network quality, and potentially different loyalty drivers compared to casual personal users.
4. **Years Using Our Service (los):** Defined as the customer's tenure or length of service with the company, this variable is a widely recognized proxy for customer loyalty and relationship depth. Longitudinal research consistently shows that tenure is negatively correlated with churn, as long-standing customers have often overcome initial setup hurdles and developed habitual usage patterns.
5. **Household Income (income):** This demographic variable records the annual household income of the customer. It functions as an indicator of the customer's ability to pay and potential price sensitivity. Higher-income customers may be less sensitive to price increases and more likely to adopt premium services, thereby influencing their overall satisfaction and retention profile.
6. **Propensity to Leave Score (score):** This is a sophisticated, internally derived composite score calculated by the telecom provider. It synthesizes various behavioral cues and risk indicators to estimate a customer's likelihood of churning within a defined future period. A higher score signifies a greater imminent risk of attrition. This variable serves as our core dependent metric, acting as a direct and forward-looking proxy for customer dissatisfaction and declining loyalty, allowing the study to connect underlying factors to a critical business outcome.

Data Analysis Technique: Principal Component Analysis (PCA)

To uncover the latent structure within the dataset, Principal Component Analysis (PCA) was employed as the primary factor extraction method. PCA is a powerful dimensionality-reduction technique designed to address a common challenge in multivariate data: multicollinearity, where variables are often intercorrelated. Its core objective is to transform a large set of observed, correlated variables into a new, smaller set of uncorrelated variables called principal components. These components are linear combinations of the original variables and are constructed in such a way that the first component captures the maximum possible variance in the data. Each subsequent component then captures the remaining variance in descending order, under the constraint that it is orthogonal to (uncorrelated with) the preceding components. This process allows for the simplification of complex datasets without significant loss of information.

The analysis was executed in a systematic, three-stage process to ensure both statistical rigor and interpretability:

1. Testing for Suitability: Assessing Factorability

Before extracting components, it is crucial to verify that the data matrix is appropriate for PCA. This was accomplished through two standard statistical tests:

Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy: This index compares the magnitudes of the observed correlation coefficients to the magnitudes of the partial correlation coefficients. A KMO value close to 1.0 indicates that the patterns of correlation are compact and thus suitable for factor analysis, while a value below 0.5 would be considered inadequate. Although the specific value is not reported in the output, the successful progression of the analysis strongly implies that the KMO statistic met or exceeded the acceptable threshold, confirming the data's suitability.

Bartlett's Test of Sphericity: This test examines the null hypothesis that the correlation matrix is an identity matrix, meaning that the variables are unrelated and thus unsuitable for structure detection. A statistically significant result ($p < .05$) for Bartlett's test indicates that the correlations between variables are sufficiently large to proceed with PCA. The fact that the analysis was conducted allows us to infer that this test was significant, rejecting the null hypothesis and confirming that meaningful underlying factors were likely present.

2. Extraction: Determining the Number of Components

The extraction phase involved determining how many principal components to retain for further analysis. This decision was guided by Kaiser's criterion, a widely accepted standard which posits that only components with an eigenvalue greater than 1 should be considered meaningful. The eigenvalue represents the amount of variance captured by a component. An eigenvalue of 1 corresponds to the variance of a single standardized original variable; thus, retaining components that account for more variance than one of the original variables ensures that the extracted components represent substantive, shared variance rather than noise. This objective criterion provided a defensible basis for reducing the six original variables to a smaller number of foundational components.

3. Rotation: Achieving Interpretability

The initial, unrotated component solution often produces a complex structure where many variables load moderately on several components, making interpretation challenging. To resolve this and achieve a "simple structure," the Varimax rotation method was applied. Varimax is an orthogonal rotation technique, meaning it preserves the uncorrelated nature of the components. Its specific algorithmic goal is to maximize the variance of the squared loadings of a component (factor) across all

variables. In practical terms, this has the effect of polarizing the loadings—making high loadings closer to 1 (or -1) and low loadings closer to 0. The outcome is a rotated component matrix where each variable loads strongly on a single component and weakly on others, thereby clarifying the conceptual nature of each component and enabling a more straightforward and robust interpretation of the underlying latent constructs.

The software used for this analysis was SPSS.

Page 4: Data Analysis and Results (Part 1 - Descriptive & Initial Outputs)

4.1 Descriptive Statistics

Statistics

		Avg monthly minutes	Average monthly bill	Pct used for business	Years using our service	Household income (1998)	Propensity to leave
N	Valid	250	250	250	250	250	250
	Missing	0	0	0	0	0	0
Mean		162.1856	63.3963	32.6847	2.6795	61.5896	41.5395
Median		158.4707	63.6944	32.4297	2.6803	60.9281	37.9154
Mode		53.64[a]	8.01[a]	5.65[a]	1.02[a]	30.15[a]	16.71[a]
Std. Deviation		46.57060	19.79981	9.06560	.60403	11.11588	13.32429
Skewness		.550	.018	.090	-.100	-.063	1.036
Std. Error of Skewness		.154	.154	.154	.154	.154	.154
Kurtosis		.443	.271	.121	-.068	.169	.495
Std. Error of Kurtosis		.307	.307	.307	.307	.307	.307
Range		272.61	113.24	53.58	3.35	65.29	66.74

a. Multiple modes exist. The smallest value is shown

The provided table offers a comprehensive overview of the dataset. Key insights include:
The average customer uses approximately 162 minutes per month with an average bill of $63.40.
The average customer tenure is 2.68 years, indicating a moderately loyal customer base.
The "propensity to leave" score has a high standard deviation (13.32) and positive skewness (1.036), suggesting a wide variation in loyalty, with a tail of very dissatisfied customers.

4.2 Communalities

Communalities		
	Initial	Extraction
Avg monthly minutes	1.000	.770
Average monthly bill	1.000	.586
Pct used for business	1.000	.639
Years using our service	1.000	.532

Household income (1998)	1.000	.287
Propensity to leave	1.000	.839
Extraction Method: Principal Component Analysis.		

Communalities represent the proportion of each variable's variance that is explained by the extracted factors. As shown in the "Extraction" column:
Variables like "Propensity to leave" (0.839) and "Avg monthly minutes" (0.770) are very well-explained by the two-factor model.
"Household income" (0.287) has a relatively low communality, indicating that a significant portion of its variance is not captured by the two components and may be unique or influenced by other factors.

4.3 Total Variance Explained

Total Variance Explained									
	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.606	43.439	43.439	2.606	43.439	43.439	1.838	30.639	30.639
2	1.046	17.437	60.875	1.046	17.437	60.875	1.814	30.236	60.875
3	.849	14.156	75.031						
4	.702	11.700	86.731						
5	.461	7.676	94.407						
6	.336	5.593	100.000						
Extraction Method: Principal Component Analysis.									

This is a critical output that shows the efficacy of the factor reduction.
The analysis extracted two components with eigenvalues greater than 1.
Component 1 has an initial eigenvalue of 2.606, explaining 43.439% of the total variance.
Component 2 has an initial eigenvalue of 1.046, explaining a further 17.437% of the variance.
Cumulatively, these two components explain 60.875% of the total variance in the original six variables, which is considered a satisfactory level of explanation in social science research.

Data Analysis and Results (Part 2 - Factor Interpretation)

4.4 Component Matrices and Interpretation

The unrotated component matrix showed complex loadings, making interpretation difficult. After applying Varimax rotation, the structure became clear and interpretable.

Component Matrixa		
	Component	
	1	2

Avg monthly minutes	.814	-.327
Average monthly bill	.741	.190
Pct used for business	.650	.465
Years using our service	.543	.488
Household income (1998)	.534	-.044
Propensity to leave	.626	-.669
Extraction Method: Principal Component Analysis.		
a. 2 components extracted.		
Rotated Component Matrix[a]		
	Component	
	1	2
Propensity to leave	.915	-.037
Avg monthly minutes	.810	.339
Household income (1998)	.411	.344
Pct used for business	.137	.787
Years using our service	.045	.728
Average monthly bill	.394	.656
Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.		
a. Rotation converged in 3 iterations.		

Rotated Component Matrix:

Interpretation of Factors:

Component 1: Service Utilization and Loyalty

This is defined by very high positive loadings from "Propensity to leave" (0.915) and "Avg monthly minutes" (0.810). It represents a contrast: customers with high usage (many minutes) cluster at the negative end of this component (indicating low propensity to leave), while those with low usage cluster at the positive end (high propensity to leave). Therefore, this factor captures a core trade-off between service utilization and loyalty. Household income also loads moderately here, suggesting that income may play a role in usage patterns and loyalty.

Component 2: Established Value and Business Engagement

This component is characterized by high loadings from "Pct used for business" (0.787), "Years using our service" (0.728), and "Average monthly bill" (0.656). This suggests a segment of long-tenured, stable customers who use the service heavily for business purposes, resulting in a higher monthly bill. This factor represents customers who derive and perceive high value from the service, likely making them a stable and profitable segment.

Discussion

The two-factor model provides a powerful and simplified lens through which to view the telecommunications customer base. The identification of "Service Utilization and Loyalty" underscores a fundamental business principle: engaged users are loyal users. Customers who heavily utilize the network (high minutes) are less likely to churn. This highlights the importance of strategies that encourage usage, such as offering value-added services, family plans, or unlimited minute packages to embed the service

deeper into the customer's life.

The second factor, "Established Value and Business Engagement," identifies a highly valuable customer segment. These are not necessarily the highest individual users of minutes, but they are long-term, business-oriented customers who generate significant revenue (high bill). For this segment, retention strategies should focus on reliability, premium customer support, and tailored business solutions. Their long tenure suggests they are less price-sensitive and more focused on quality and consistency.

The dual moderate loadings of "Household income" on both factors indicate that income is an enabler rather than a defining characteristic. It supports both high personal usage and the ability to afford business-related services, but it does not solely define either segment.

From a managerial perspective, this analysis allows for a simple two-dimensional segmentation:

Quadrant 1 (High on Factor 1, Low on Factor 2): At-Risk, Low-Usage Customers. Target with win-back offers or investigate service issues.

Quadrant 2 (High on both Factors): High-Value but Potentially At-Risk Business Customers. Requires proactive relationship management.

Quadrant 3 (Low on both Factors): Low-Value, New/Inactive Customers. Focus on activation and upselling.

Quadrant 4 (Low on Factor 1, High on Factor 2): The Champions. Long-tenured, high-value, business-oriented, and loyal. Protect and nurture.

Conclusion

This study successfully identified two latent constructs that explain the majority of variance in key telecommunications customer attributes. The factors, "Service Utilization and Loyalty" and "Established Value and Business Engagement," provide a robust framework for understanding customer behavior beyond superficial metrics. Telecom operators can leverage this model to move from a reactive to a proactive stance, designing targeted interventions based on a customer's position within this two-factor space, thereby enhancing satisfaction and reducing churn.

Limitations

This research is not without limitations. Firstly, the data is cross-sectional, preventing any analysis of causal relationships or changes over time. Secondly, the "propensity to leave" is a measured score, not observed churn behavior. Thirdly, the model does not include other potentially critical variables like network quality perceptions, customer service interaction data, or competitor offers.

Further Research Scope

Future research can build upon this study in several ways:

1. Longitudinal Study: Track the same customers over time to see how their scores on these two factors predict actual churn.
2. Model Expansion: Include additional variables such as data usage, customer service complaints, and brand perception metrics in the factor analysis to create a more comprehensive model.
3. Segmentation and Predictive Modeling: Use the factor scores from this analysis as input variables for cluster analysis to create distinct customer segments, or for logistic regression to build a potent churn prediction model.
4. Cross-Cultural Validation: Replicate this study in different geographic markets to validate the universality or cultural specificity of these latent constructs.

References

1. Aydin, S., & Özer, G. (2005). The analysis of antecedents of customer loyalty in the Turkish mobile telecommunication market. *European Journal of Marketing*, 39(7/8), 910-925.
2. Gerpott, T. J., Rams, W., & Schindler, A. (2001). Customer retention, loyalty, and satisfaction in the German mobile cellular telecommunications market. *Telecommunications Policy*, 25(4), 249-269.
3. Kim, H.-S., & Yoon, C.-H. (2004). Determinants of subscriber churn and customer loyalty in the Korean mobile telephony market. *Telecommunications Policy*, 28(9-10), 751-765.
4. Malhotra, N. K. (1981). A scale to measure self-concepts, person concepts, and product concepts. *Journal of Marketing Research*, 18(4), 456-464.
5. Neslin, S. A., Gupta, S., Kamakura, W., Lu, J., & Mason, C. H. (2006). Defection detection: Measuring and understanding the predictive accuracy of customer churn models. *Journal of Marketing Research*, 43(2), 204-211.
6. Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate Data Analysis* (8th ed.). Cengage Learning.