



YAŞAR UNIVERSITY

GRADUATE SCHOOL

PHD THESIS

SPATIAL MODELING
FOR
SERVICE QUALITY ASSESSMENT

GÖNÜL AYRANCI

THESIS ADVISOR: PROF.(PHD) LEVENT KANDİLLER

PHD IN BUSINESS ADMINISTRATION

BORNOVA / İZMİR
AUGUST 2024

JURY APPROVAL PAGE

We certify that, as the jury, we have read this thesis and that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of the Doctor of Philosophy.

Jury Members:

Signature:

Prof.(PhD) Levent KANDİLLER
Yaşar University

.....

Prof.(PhD) İge PIRNAR
Yaşar University

.....

Assoc.(PhD) Görkem ATAMAN
Yaşar University

.....

Prof.(PhD) Engin Deniz ERİŞ
Dokuz Eylül University

.....

Prof.(PhD) Zafer ÖTER
İzmir Kâtip Çelebi University

.....

Prof. (PhD) Yucel Ozturkoglu

ABSTRACT

SPATIAL MODELING FOR SERVICE QUALITY ASSESSMENT

Ayrancı, Gönül

PhD, Business Administration

Advisor: Prof.(PhD) Levent Kandiller

August 2024

In producing durable consumer goods, the tangible product cannot be separated from the intangible elements of the after-sales customer services provided. For manufacturers to increase their market share and create value for buyers, it is insufficient for the product to be technically competitive alone. It is also crucial that the after-sales services related to the product are provided at a quality level that satisfies customers. Therefore, manufacturers continuously endeavor to develop and improve their after-sales services. However, to ensure a high level of service quality, it is first necessary to accurately evaluate the provided services and identify the factors that shape, and influence customer satisfaction derived from these services.

This thesis investigates the factors affecting the after-sales service quality of manufacturing companies. The data are from after-sales service surveys conducted nationwide by a company producing durable consumer goods. Initially, the ability of the survey data to perform a multidimensional evaluation of service quality, as suggested in the literature, was examined using Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) methods. The priority weights of the identified dimensions were then determined through an Analytical Hierarchy Process (AHP) analysis based on expert opinions. Finally, the primary contribution of the thesis, the impact of geographical variables on customers' sensitivities to service dimensions, is examined using the Geographically Weighted Regression (GWR) method. This methodology allows for both inter-dimensional comparisons and regional comparisons in the evaluation of after-sales service quality. The importance

and impact of considering regional differences in service design for manufacturers providing after-sales services through a nationwide dealer network are particularly emphasized and demonstrated.

Keywords: after-sale service, service quality, confirmatory factor analysis, analytic hierarchy process, geographically weighted regression



ÖZ

HİZMET KALİTESİNİN ÖLÇÜMÜNDE COĞRAFI MODELLEME

Ayrancı, Gönül

Doktora Tezi, İşletme

Danışman: Prof. Dr. Levent Kandiller

August 2024

Dayanıklı tüketim mallarının üretimde, somut ürün, satış sonrası sağlanan müşteri hizmetlerinin soyut unsurlarından ayrılamayacak niteliktedir. Üreticilerin, pazar paylarını arttırabilmesi ve alıcılar için değer yaratabilmesi için ürünün yalnızca teknik açıdan rekabetçi olması yeterli değildir. Ürünle ilgili sunduğu satış sonrası hizmetlerin de müşterileri memnun edecek kalitede sunması oldukça önemlidir. Bu nedenle üreticiler satış sonrası hizmetlerini geliştirmek ve iyileştirmek için sürekli çaba göstermektedir. Ancak iyi seviyede hizmet kalitesi sağlamak için öncelikle sunulan hizmetin doğru bir şekilde değerlendirilmesi ve müşterilerin bu hizmetlerden kaynaklanan memnuniyet seviyesini şekillendiren ve etkileyen faktörlerin belirlenmesi gerekmektedir.

Tez çalışmasında, üretici firmaların satış sonrası hizmet kalitesini etkileyen faktörler incelenmektedir. Dayanıklı tüketim malları üreten bir firmanın tüm ülke çapında yaptığı satış sonrası hizmet anketleri verisi kullanılmıştır. Öncelikle anketten alınan verilerin literatürde önerildiği üzere servis kalitesi üzerinde çok boyutlu bir değerlendirme yapabilme yeteneği Keşfecidici Faktör Analizi (KFA) ve Doğrulayıcı Faktör Analizi (DFA) yöntemleri kullanılarak sorgulanmıştır. Belirlenen boyutların öncelikli ağırlıkları uzman görüşleri ile yapılan bir Analitik Hiyerarşi Prosesi (AHP) analizi ile ortaya çıkarılmıştır. Son olarak, tezin temel katkısı olan, coğrafi değişkenlerin müşterilerin servis boyutlarındaki hassasiyetlerine olan etkisi Coğrafi Ağırlıklı Regresyon (GAR) yöntemi ile incelenmiştir. Bu şekilde satış sonrası hizmet kalitesi değerlendirilirken boyutlar arası karşılaştırmanın yanı sıra, bölgesel karşılaştırmaların da yapılmasını sağlayan bir metodoloji ortaya konmuştur. Özellikle, ülke geneline yayılmış bayi ağı ile satış sonrası hizmeti sağlayan üreticilerin hizmet

tasarımında bölgesel farklılıkları deęerlendirmesinin önemi vurgulanmış ve etkisi gösterilmiştir.

Anahtar Kelimeler: satış sonrası hizmet, hizmet kalitesi, doğrulayıcı faktör analizi, analitik hiyerarşı prosesi, coęrafi aęırlıklı regresyon



ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude to my thesis advisor, Prof. Dr. Levent Kandiller. He guided me at every stage of this work, broadened my horizons with his vast knowledge, and became a great source of academic inspiration for me. I am grateful to him for patiently guiding me and encouraging me to overcome every challenge I faced.

I would also like to extend my sincere thanks to Prof. Dr. İge Pınar and Assoc. Prof. Dr. Gökem Ataman for their invaluable contributions and suggestions regarding my thesis. Their insights significantly enhanced the quality of this work. Additionally, I would like to thank the jury members, Prof. Dr. Engin Deniz Eriş and Prof. Dr. Zafer Öter, for their valuable constructive critiques and recommendations.

I cannot overlook the support of my dear friends Hüseyin Arslan, Saliha Ertan, Dr. Melisa Özbiltekin Pala, and Dr. Yeşim Deniz Özkan Özen, who provided me with strength throughout this journey.

To my parents, I extend my deepest thanks for their endless support, love, and understanding, which have been a constant source of strength.

Lastly, I am immensely grateful to Achilles, who shared this challenging journey with me, standing by my side every step of the way, and serving as my greatest source of motivation.

Gönül Ayrancı

İzmir, 2024

TEXT OF OATH

I declare and honestly confirm that my study, titled “SPATIAL MODELING FOR SERVICE QUALITY ASSESSMENT” and presented as a PhD Thesis, has been written without applying any assistance inconsistent with scientific ethics and traditions. I declare, to the best of my knowledge and belief, that all content and ideas drawn directly or indirectly from external sources are indicated in the text and listed in the list of references.

Gönül Ayrancı

16.08.2024



TABLE OF CONTENTS

JURY APPROVAL PAGE	iii
ABSTRACT	v
ÖZ	ix
ACKNOWLEDGEMENTS	xiii
TEXT OF OATH	xv
TABLE OF CONTENTS	xvii
LIST OF FIGURES	xxv
LIST OF TABLES	xxix
SYMBOLS AND ABBREVIATIONS	xxxv
1. CHAPTER: INTRODUCTION	1
2. CHAPTER: LITERATURE REVIEW	7
2.1. Service, Service Quality and After-Sale Services	7
2.1.1. Characteristics of Services	7
2.1.2. Service Quality	8
2.1.3. After-Sale Services and Service Quality	19
2.2. Service Quality Assessment	21
2.2.1. SERVQUAL	21
2.2.2. SERVPERF	26
2.2.3. Importance Weighting – Regression, Factor Analysis & AHP	27
2.3. Spatial Variability in Service Quality	31
3. CHAPTER: RESEARCH DESIGN AND METHODOLOGY	35
3.1. Data Collection	35
3.1.1. Survey Questions	35

3.1.2.	Determination of the Sample and Sample Size.....	36
3.1.3.	Descriptive Statistics and Data Cleaning	37
3.1.3.1	Repair/Maintenance Service	37
3.1.3.2	Installation/Setup Service.....	40
3.2.	Exploratory Factor Analysis	42
3.2.1.	Variables to Include	43
3.2.2.	Participants to Include.....	44
3.2.3.	Distributional Properties of Data	44
3.2.4.	Factor Extraction Methods.....	45
3.2.5.	Number of Factors.....	46
3.2.6.	Rotation of Factors.....	47
3.2.7.	Interpretation of the Results.....	47
3.3.	Confirmatory Factor Analysis.....	48
3.3.1.	Basic Definitions and Concepts	49
3.3.2.	Developing & Estimating Measurement Model	53
3.3.3.	Model Evaluation	54
3.3.3.1	Chi-square (χ^2) Fit Index	55
3.3.3.2	Root Mean Square Error of Approximation (RMSEA)	55
3.3.3.3	Standardized Root Mean Square Residual (SRMR)	56
3.3.3.4	Incremental Fit Indices (TLI & CFI)	57
3.3.3.5	Information Criterion-Based Fit Indices	58
3.3.4.	Modifying the Measurement Model.....	58
3.3.5.	Construct Validity	60
3.3.5.1	Convergent Validity	61
3.3.5.2	Discriminant Validity.....	62
3.4.	Analytic Hierarchy Process (AHP)	63

3.5.	Geographically Weighted Regression (GWR).....	66
4.	CHAPTER: FINDINGS & RESULTS	71
4.1.	Exploratory Factor Analysis Results.....	71
4.1.1.	Data Characteristics	71
4.1.2.	Factor Extraction	74
4.1.3.	Installation/Setup Dataset EFA Results	79
4.2.	Confirmatory Factor Analysis Results	80
4.2.1.	Estimating CFA.....	80
4.2.2.	Model Evaluation – Initial Model	82
4.2.3.	Modifying the Measurement Model.....	83
4.2.4.	Model Evaluation – Final Model	89
4.2.5.	Construct Validity Diagnostics	92
4.2.6.	Final CFA Model for Installation/Setup Dataset	97
4.3.	Interpretation of the Factors	98
4.4.	Analytic Hierarchy Process (AHP) Results	101
4.4.1.	Repair /Maintenance Service AHP Results.....	101
4.4.2.	Installation/Setup Service AHP Results.....	106
4.5.	Geographically Weighted Regression (GWR) Results	109
4.5.1.	Ordinary Least Squares for Repair/Maintenance Service	113
4.5.2.	Geographically Weighted Regression Repair/Maintenance Service	115
4.5.3.	Geographically Weighted Regression Installation/Setup Service....	126
5.	CHAPTER: DISCUSSION AND CONCLUSION	130
6.	REFERENCES	135
	APPENDICES.....	149
	APPENDIX 1. HOAM Questionnaire.....	149
	APPENDIX 2. EFA Results for Installation/Setup Service	153

APPENDIX 3. CFA Results for Installation/Setup Service.....	157
APPENDIX 4. AHP Results.....	167
APPENDIX 5. GWR Results for Installation/Setup Service	169



LIST OF FIGURES

Figure 2.1. GAP (American) Model of Service Quality	9
Figure 2.2. Initial Determinants of Service Quality	12
Figure 2.3. Three Ps of service quality	13
Figure 2.4. Synthesized Model of Service Quality	14
Figure 2.5. Satisfaction Service Quality Model	15
Figure 2.6. P-C-P Model for Service Quality	17
Figure 2.7. Antecedents and Mediator Model of Service Quality	18
Figure 2.8. The distribution of 22 items to SERVQUAL dimensions.	24
Figure 2.9. Conceptual model of a hypothetical SERVPERF model with CFA.....	30
Figure 3.1. Measurement model with one construct and four measured variables...	49
Figure 3.2. Measurement model with two constructs and four measured variables .	51
Figure 3.3. FC2 Measurement model with two constructs and four variables	52
Figure 3.4. One construct and four measured variables and correlated errors.....	53
Figure 4.1. The Pearson correlation matrix (a) and polychoric correlation (b)	72
Figure 4.2. Screeplot of eigenvalues against parallel simulation.....	74
Figure 4.3. Initial CFA model	81
Figure 4.4. 1-Modified CFA model	84
Figure 4.5. Final CFA model for Repair/Maintenance	97
Figure 4.6. Final CFA model for Installation/Setup	98
Figure 4.7. The hierarchical structure designed for Repair /Maintenance Service.	101

Figure 4.8. The hierarchical structure designed for Installation /Setup Service	106
Figure 4.9. Histograms, approximate distributions and rug marks of AHP.....	111
Figure 4.10. Frequency Distribution of Overall Satisfaction	113
Figure 4.11. Q-Q Plot Standardized Residuals	115
Figure 4.12. Geographic distribution of observations.....	116
Figure 4.13. Distribution of bandwidth values in km	117
Figure 4.14. Geographic distribution of local R^2 values.	119
Figure 4.15. Geographic distribution of values of Factor1 coefficient in GWR.....	121
Figure 4.16. Distribution of values of Factor 1 coefficients	122
Figure 4.17. Factor1 coefficients in the 95% confidence interval of OLS	122
Figure 4.18. Geographic distribution of values of Factor2 coefficient in GWR.....	123
Figure 4.19. Distribution of values of Factor 2 coefficients	124
Figure 4.20. Geographic distribution of values of Factor3 coefficient in GWR.....	125
Figure 4.21. Distribution of values of Factor 3 coefficients	126
Figure 4.22. Geographic distribution of values of Factor 1 coefficient in GWR....	128
Figure 4.23. Geographic distribution of values of Factor 2 coefficient in GWR....	128
Figure 4.24. Geographic distribution of values of Factor 3 coefficient in GWR....	129

LIST OF TABLES

Table 3.1. Descriptive Statistics for the Yes/No Questions	38
Table 3.2. Descriptive Statistics for Likert-Type Questions	39
Table 3.3. Descriptive Statistics for the Yes/No Questions	40
Table 3.4. Descriptive Statistics for Likert-Type Questions	41
Table 3.5. The typical scale for AHP	64
Table 3.6. The Random Index (RI)	66
Table 4.1. KMO values obtained from the Repair Dataset	73
Table 4.2. Unrotated factor solution.....	76
Table 4.3. 4-Factor solution after varimax rotation.....	76
Table 4.4. 3-Factor solution after varimax rotation.....	77
Table 4.5. Factor loadings table for 3-Factor solution	78
Table 4.6. Factor loadings table for 3-Factor solution	79
Table 4.7. Chi-square test for the initial (16 measured variable) model.	82
Table 4.8. Fit indices for the initial (16 measured variable) model.	82
Table 4.9. Modification indices for measured variables in the initial model.....	83
Table 4.10. Fit indices for 1-Modified model.	85
Table 4.11. Modification indices for measured variables in 1-Modified model.....	85
Table 4.12. Fit indices for 2-Modified model	86
Table 4.13. Modification indices for measured variables in 2-Modified model.....	86
Table 4.14. Fit indices for 3-Modified model	88
Table 4.15. Modification indices for measured variables in 3-Modified model.....	88

Table 4.16. Chi-square test for the initial (13 measured variable) model	89
Table 4.17. Fit indices for the Final model	90
Table 4.18. Modification indices for measured variables in the final model	90
Table 4.19. Residual covariance modification indices for measured variables	91
Table 4.20. Factor loadings of the Final Model	92
Table 4.21. Average Variance Extracted table of the Final Model	93
Table 4.22. R ² values of each measured variable of the Final Model	93
Table 4.23. McDonald's omega and Cronbach alpha values of the Final Model	94
Table 4.24. Fit index values of the Final Model and the Constrained Model	94
Table 4.25. Factor correlation of the Final model.	95
Table 4.26. Comparison of AVE values and squared correlations of Factors	96
Table 4.27. Repair/Maintenance Service Quality Dimension Correspondence	98
Table 4.28. Installation/Setup Service Quality Dimension Correspondence	100
Table 4.29. Pairwise comparisons for the first level of the hierarchical concept....	102
Table 4.30. Priority Weights	102
Table 4.31. Pairwise Comparisons for the Items under Factor 1	103
Table 4.32. Pairwise Comparisons for the Items under Factor 2	104
Table 4.33. Pairwise Comparisons for the Items under Factor 3	105
Table 4.34. Priority weights for Items under Factor 1	105
Table 4.35. Priority weights for Items under Factor 2	105
Table 4.36. Priority weights for Items under Factor 3	106
Table 4.37. Pairwise comparisons for the first level of the hierarchical concept....	107

Table 4.38. Priority Weights	107
Table 4.39. Priority weights for Items under Factor 1	108
Table 4.40. Priority weights for Items under Factor 2	108
Table 4.41. Priority weights for Items under Factor 3	109
Table 4.42. Descriptive statistics of factor scores	110
Table 4.43. Summary Statistics and Frequency Table	112
Table 4.44. ANOVA table of the OLS model	114
Table 4.45. Model coefficients and collinearity diagnostics table.	114
Table 4.46. Six-point summary table of the bandwidth values.	117
Table 4.47. Diagnostics of GWR compared to OLS.	118
Table 4.48. Basic distributional characteristics of GWR model coefficients	118
Table 4.49. Six-point summary table of the Factor 1 coefficients	121
Table 4.50. Six-point summary table of the Factor 2 coefficients	124
Table 4.51. Six-point summary table of the Factor 3 coefficients	125
Table 4.52. Diagnostics of GWR compared to OLS	126
Table 4.53. Basic distributional characteristics of GWR model coefficients	127

SYMBOLS AND ABBREVIATIONS

SYMBOLS:

λ_i	Loading of a measured variable i
ξ	Latent Variable
δ_i	Error of variable i in a CFA model
χ^2	Chi-square
α	Cronbach's Alpha
β	Coefficient of a Covariate in a Linar Model
ε	Error Term
R^2	Percent of the Variation Explained
ω	McDonald's Omega

ABBREVIATIONS:

AHP	Analytic Hierarchy Process
AIC	Akaike Information Criterion
AVE	Average Variance Extracted
BIC	Bayesian Information Criterion
CFA	Confirmatory Factor Analysis
CFI	Comparative Fit Index
CI	Consistency Index
CR	Consistency Ratio
df	Degrees of Freedom
DWLS	Diagonally Weighted Least Squares
EFA	Exploratory Factor Analysis
FCn	Factor Complexity n
GWR	Geographically Weighted Regression
INTSERVQUAL	Internal Service Quality
KMO	Kaiser-Meyer-Olkin
OLS	Ordinary Least Squares
PCA	Principal Component Analysis
RI	Random Index
RMSEA	Root Mean Square Error of Approximation
SBIC	Schwartz Modified Bayesian Information Criterion
SERPERF	Service Performance
SERVQUAL	Service Quality
SRMR	Standardized Root Mean Square Residual
TLI	Tucker-Lewis Index
ULS	Unweighted Least Squares

1. CHAPTER: INTRODUCTION

In recent years, with the competitive pressure created by digitalization across all sectors brought by Industry 4.0, manufacturers strive to offer product-related services or improve the services they provide to make their products preferable. For these firms to remain competitive in the market, it is essential that their service is of high quality and that customers are satisfied with the service provided. Of course, achieving this requires the continuous evaluation of the service quality. However, due to factors such as the multidimensional nature of service quality, the intangible nature of the service itself, and the formation of service quality satisfaction in the perceptions of customers, this evaluation process becomes quite challenging (Zeithaml et al., 1990). This difficulty and necessity have led to frequent examination of service quality in academic literature and efforts to provide suggestions and solutions.

Service quality, especially its measurement, is growing, having an impact on practitioners and researchers. The concept of perceived service quality has developed since the 1980s with numerous service quality models proposed by researchers (Ghotbabadi et al., 2015; Polyakova & Mirza, 2015; Seth et al., 2005; Yarimoglu, 2014). While some studies have focused on conceptual models, others have developed models for specific industries or tested the validity of these models across different industries. The most widely used model and dimensions for modeling service quality in the literature have been the SERVQUAL model, first proposed by Parasuraman et al. (1988). Despite the acceptance of the service quality dimensions based on this model by many researchers, the problem that customers may attribute different levels of importance to these dimensions has been highlighted (Zeithaml et al., 1990; Cronin & Taylor, 1992). In summary, when evaluating service quality, it is not sufficient to consider it only as a multidimensional concept; it is also essential to consider the differences in importance attributed to these dimensions by customers (Altuntas et al., 2012). Correctly identifying customers' different priorities will significantly impact any improvements firms make in their service activities.

The increasing importance given to customer satisfaction, the emerging idea that one way to stand out in competition and increase profitability is through improvements in customer service, and the perception that retaining existing customers is less costly than acquiring new ones have led manufacturers to become more service-oriented (Hussain et al., 2024). This shift necessitates that manufacturers focus not only on their products but also on the services provided alongside these products. These services, known as after-sales services (ASS), include tangible elements such as providing spare parts and are strategically important. Offering products along with services increases product sales, customer satisfaction, and loyalty and provides opportunities for market differentiation and growth (Murali et al., 2016). ASS is considered a crucial source of income and a primary differentiator for businesses that deal with durable items like cars, computers, and home appliances (Saccani et al., 2006). End consumers of many different product categories need ASS at some point. Therefore, ASS activities and strategic goals have become an essential part of the business for manufacturers producing durable items. Manufacturers maintain and manage customer relationships through the additional services they provide. ASS also holds strategic importance as it provides a feedback channel for customers regarding the design, quality, and sales activities of the products (Nasir et al., 2021).

Customers increasingly value not only the quality of the product but also its availability throughout its lifecycle and the services provided during this process (Verhoef & Lemon, 2013). Particularly in a world where the linear economy model is not sustainable, and the circular economy model is targeted, the lifespan, repairability, reuse, and recycling of products are crucial issues for the entire society (Franco, 2019). In the circular economy model, the efforts of manufacturers are critical. Proper usage of the product, extending its lifespan, being able to repair it when necessary, providing spare parts, and the reuse or transformation of the product's components, when it becomes unusable, can only be achieved through after-sales services provided by the manufacturer (Sauvé et al., 2016).

After-sales services (ASS) should not only be seen as the installation, maintenance, and repair of products but also as a set of high-quality services that include customer participation, establish long-term customer relationships, and attract new customers

through customer's recommendations. Additionally, after-sales services are highly effective in providing a good image and reputation, enabling more sales with less marketing effort (Rigopoulou et al., 2008).

Despite the significant impact of after-sales services provided with the product, there are few studies in the literature on ASS. The quality of after-sales services is a crucial factor for manufacturers to have a large market share and long-term success. However, providing the same quality of service across all service regions for banks, automobile manufacturers, and home electronics producers who serve over a large geographic area and developing strategies to improve service quality can be challenging (Mittal et al., 2004). The preferences and expectations of customers served by a nationwide dealer network may vary from region to region, and these variable regions may not be defined by legally defined geographic boundaries such as cities or districts (Chang & Zhang, 2016). Furthermore, the company's capacity and ability to provide services may vary according to the service region. It is emphasized that customers may have different sensitivity levels that show regional differences. Customers may not respond the same way to service improvements in all regions. This indicates that a regional approach to service strategy would be more effective. All these reasons suggest that geographic differences should also be considered when evaluating service quality. Evaluating customers' perceptions of service quality and service satisfaction based on regional differences will help the company define geographic customer segments and develop regional strategies accordingly.

In summary, the role of after-sales services provided by manufacturers and products in the company's competitive power is becoming increasingly important. Therefore, manufacturers must improve the perceived quality of the services they offer with the product. Although there is a basic consensus in the literature regarding the attributes of service quality, the dimensions in which it is evaluated, and the other factors it influences or is influenced by, there are deficiencies in evaluating the quality of after-sales services provided by manufacturers. Moreover, when conducting a multidimensional evaluation for after-sales services, it should be considered that customers may evaluate these dimensions with different priorities, and not all dimensions are equally important. The success of manufacturing firms providing after-sales services depends on their ability to evaluate these services accurately. Additionally, when these services are provided through a dealer network spread over

a large geographic area, regional factors affecting the perception of service quality make accurate evaluation challenging.

The main objective of this thesis is to propose a methodology that can be used by manufacturing firms to measure the quality of after-sales services. First, it is aimed to evaluate the data collected with a questionnaire prepared by the company and to determine the suitability of the questionnaire for the multidimensional service models suggested in the literature. Additionally, the study seeks to answer what weights should be assigned to the identified dimensions. Finally, a geographical model is proposed to measure the impact of regional differences on customers' evaluation of service quality.

The research questions of the thesis are as follows:

RQ 1: Is the questionnaire used by the company to measure service quality appropriate for this measurement?

RQ 2: According to the customer responses collected to measure the quality of after-sales services, how many dimensions do customers evaluate the service quality in?

RQ 3: Does the scale to measure after-sales service quality satisfy construct validity?

RQ 4: Is there a difference in importance among the dimensions evaluated in measuring service quality?

RQ 5: Are there significant geographical differences in the evaluation of service quality?

RQ 6: Does the consideration of geographical variability in the measurement of service quality increase the explanatory power of the model?

Although the evaluation of service quality and its dimensions have been frequently discussed in the literature for many years, studies evaluating the after-sales services provided by manufacturers in addition to the product have remained limited. Furthermore, different services such as installation, maintenance, and repair of the product offered in after-sales services have been considered as a whole, and the differences between these services have not been discussed, to our knowledge. While there are studies modeling geographical differences in evaluating after-sales service quality, there is no study discussing the geographical differences in service quality dimensions.

The thesis consists of five chapters, briefly summarized below:

The first chapter explains the importance, main objective, research questions, and originality of the thesis. The second chapter summarizes previous studies and proposed models related to the research problem of the thesis. The third chapter provides basic information and descriptive statistics about the data used and details the methodology used in the thesis, including the components, representations, and interpretations of the models employed. The fourth chapter presents the outputs of the proposed methodology using two different datasets. The fifth and final chapter discusses the important findings of the thesis, its contributions to the literature, and suggestions for further research.



2. CHAPTER: LITERATURE REVIEW

2.1. Service, Service Quality and After-Sale Services

2.1.1. Characteristics of Services

One of the most significant contributions of studies on services is the classification that includes the four characteristics of services, highlighting their distinction from the marketing of goods. As a result, it has been understood that services should be considered a separate component of companies' marketing strategies (Parry et al., 2011). A comprehensive understanding of service is defined by four well-documented characteristics.

First, most services are intangible. Due to their nature as performances rather than tangible items, it is never possible to develop precise production criteria for ensuring consistent quality. In order to maintain high standards, most services cannot be quantified, assessed, recorded, examined, or confirmed prior to being sold. Due to the intangibility of services, it becomes challenging to determine measurable aspects of service quality and for providers to convey what they offer to customers concretely. While customers can physically examine, test, and even try a product, this is not possible with services. Therefore, customers must rely on their personal experiences and perceptions when evaluating service quality (Zeithaml, 1981).

Second, services are heterogeneous: since services are performance-based, it is challenging to offer customers a consistent standard and to maintain this standard continuously (Chase, 1981). The service can exhibit variability in its type and quality, depending upon the proficiency, expertise, and behavior of the service workers. Additionally, customer experiences may differ, and the same service can lead to different customer perceptions.

Third, the production and consumption of services usually occur simultaneously and are, therefore, inseparable. Service quality necessitates the inevitable interaction between the consumer and the representative of the service provider (Habib & Victor,

1991). In fact, in some services, the consumer's input directly affects the quality of the service. The service firm may have less managerial control over service quality where consumer participation is intense.

Fourth, services are perishable. Services cease to exist immediately after they are performed. For this reason, perishability has frequently been linked to the inability to store or accumulate services (Moeller, 2010) while tangible goods can be manufactured, stored, and sold in advance, intangible services cannot be held for future use. The service is produced and delivered concurrently, with the consumer being a crucial component of the production process. Regarding intangible products, it implies that the consumer and the provider interact throughout the service delivery process. Consumers' degree of loyalty to the service provider is enhanced by the experience of the services provided.

The inherent differences in services cause significant challenges when it comes to evaluating and measuring service quality.

2.1.2. Service Quality

Due to the difficulty of establishing general quality criteria for services, the assessment of service quality has primarily centered around assessing if the service meets customer expectations. When evaluating service quality, two main aspects to consider: the consumers' perception of service quality and the elements that impact its quality. According to Grönroos (1984), companies should seek to find answers to these two inquiries to provide high-quality service and the best way to develop a model of service quality. To determine the factors affecting service quality, technical quality, functional quality, and image, perceived service quality concepts are identified and set as the dimensions of the Grönroos (1984) service quality model. Technical quality refers to a product or service's specific characteristics and attributes that determine what the consumer receives. On the other hand, functional quality refers to the processes and methods by which the customer obtains the product or service. The corporate image is determined by the perception of the firm by its customers, and customers' expectations are shaped according to the company's image.

Parasuraman et al. (1985) define the service quality based on the difference between customer expectations and experiences about the service and introduce a theoretical framework for evaluating the quality of service. The model was developed by analyzing qualitative data from substantial exploratory research conducted in various service organizations, including focus groups and in-depth executive interviews. Four specific gaps were discovered on the service provider's end that could potentially impact customers' perception of service quality.

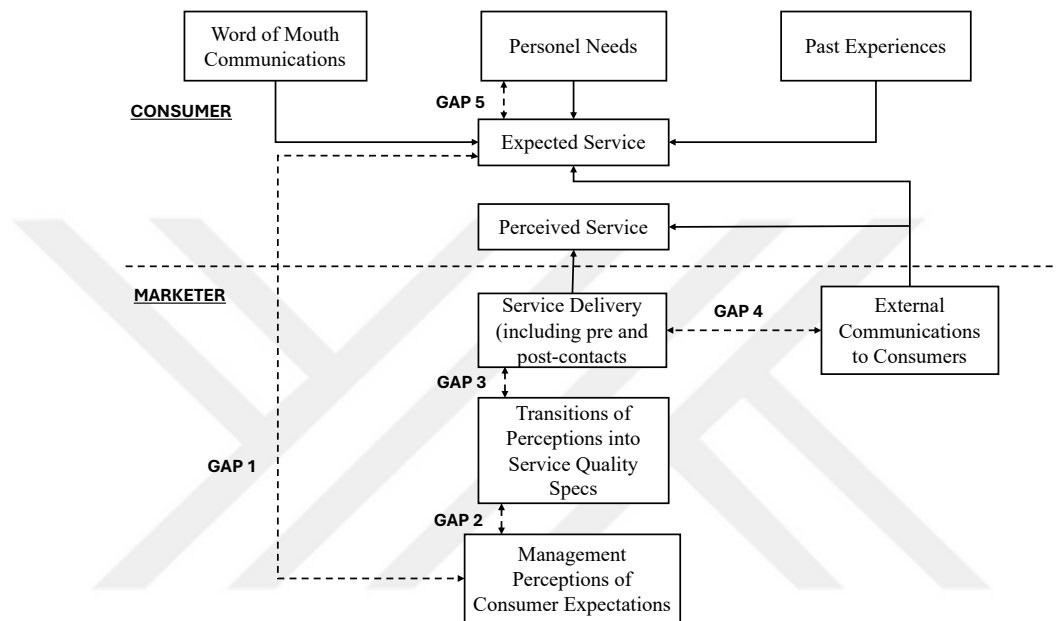


Figure 2.1. GAP (American) Model of Service Quality

Source: Parasuraman et al., 1985

GAP 1 (knowledge gap) defines the gap between company perceptions about customer expectations and actual customer expectations. There may be a discrepancy between what consumers truly expect and the management's impression of these expectations. The gap exists due to a deficiency in adequately targeting and prioritizing the market and customers. Merely having a marketing department does not inherently ensure a strong emphasis on the market. It necessitates using suitable management procedures, market analysis instruments, and mindset.

GAP 2 (policy gap) defines the gap between company perceptions of consumer expectations regarding service policy and actual service design. Also known as the service quality specification gap, this gap pertains to elements of service design and

refers to a situation where management cannot accurately convert customer expectations into service quality requirements.

GAP 3 (delivery gap) defines the gap between service design and service delivery. Despite the presence of criteria for effectively delivering services and treating customers appropriately, there is no guarantee of consistently high-quality service performance. A service organization's employees significantly impact consumers' perceived quality of service, and it is not always possible to standardize employee performance.

GAP 4 (communications gap) defines the gap between the delivery of a service and the way it is communicated to consumers. A company's promotional activities have the potential to influence consumers' expectations. Suppose expectations significantly impact how consumers perceive the quality of service. In that case, the firm needs to avoid making promises in its communications that it cannot fulfill. Overselling will initially build expectations but ultimately diminish impressions of quality when the promises are not met.

GAP 5 (consumer gap) forms the customer side of the model and defines the gap between the perceived service and the customer's expected service. Overall, the service quality model GAP 5 is defined as a function of the other four gaps. In other words, consumer perceptions of service quality depend on the gaps occurring on the company's side.

Furthermore, Parasuraman et al. (1985) investigated consumers' criteria for evaluating service quality. In their exploratory study, which laid the groundwork for their later work on SERVQUAL, Parasuraman et al. (1985) identified the critical determinants of service quality. The table below illustrates specific criteria for each determinant that emerged from the focus groups. These ten determinants are not asserted to be entirely distinct from one another. As the research was exploratory, future empirical studies will be necessary to measure potential overlaps among the criteria and to determine if any can be consolidated.

Reliability refers to the consistent and dependable delivery of services, ensuring that the firm meets customer expectations and fulfills its promises.

Responsiveness pertains to how quickly a firm responds to customer needs and requests. This dimension also encompasses the firm's ability to provide swift and practical solutions in unexpected situations.

Competence signifies the necessary knowledge, skills, and experience to perform the service effectively. This dimension highlights that the quality and nature of the service depend on the competence level of the personnel providing it.

Access denotes the ease customers can reach and communicate with the service provider. This dimension is concerned with the physical and communicative accessibility that facilitates customers in obtaining the service.

Courtesy involves the politeness, respect, consideration, and friendliness displayed by the service personnel towards customers. This dimension is crucial for creating a positive customer experience and plays a significant role in the perception of service quality.

Communication refers to the service provider's ability to convey information to and listen to customers. This dimension includes explaining the service, its operation, and cost in clear and understandable language, avoiding complex technical terms.

Credibility relates to the service provider's trustworthiness, honesty, and believability. This dimension reflects customers' trust in the service provider and the firm's sincerity in considering the interests of its customers.

Security ensures that customers receive services without danger, risk, worry, or doubt. It covers various aspects, including physical, financial, and confidentiality-related safety.

Understanding/Knowing the Customer involves understanding customers' needs, expectations, and preferences. This dimension emphasizes the importance of providing personalized service and meeting customer requirements. It also considers cultural and demographic differences, ensuring the service is appropriately tailored to different customer groups.

Tangibles refer to the physical evidence and visible aspects of the service. This dimension encompasses all material elements that reflect the quality of the service. It includes up-to-date and functional technology, equipment, and tools used in service delivery. The appearance of the service provider's personnel also falls under this

dimension. Professional, clean, and well-dressed staff instill confidence in customers and reflect the seriousness of the service.

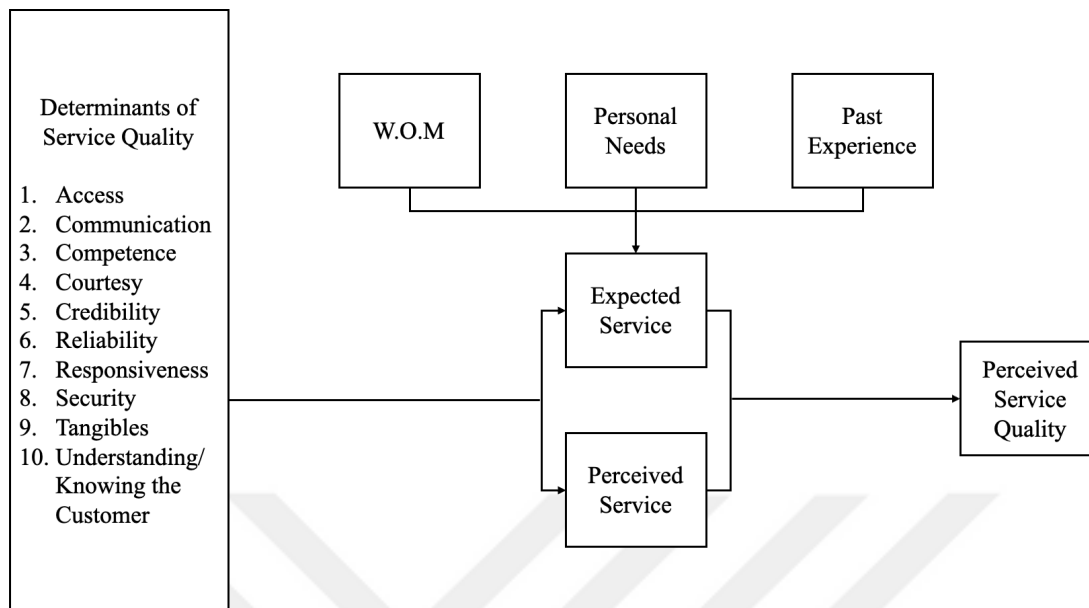


Figure 2.2. Initial Determinants of Service Quality

Source: Parasuraman et al., 1985

Later, Parasuraman et al. (1988) extended their initial study (Parasuraman et al., 1985) regarding the criteria consumers used to evaluate service quality and proposed a 22-item SERVQUAL scale for measuring customers' perceptions of service quality. The service quality construct, which initially had ten overlapping dimensions, was refined into a multiple-item scale with five dimensions. SERVQUAL is discussed in detail in Section 2.2. The refined service quality dimensions described in (Parasuraman et al., 1988) are Tangibles, Reliability, Responsiveness, Assurance, and Empathy. Three of these dimensions originally occur in Parasuraman (1985); Assurance and Empathy are defined as the dimensions obtained by combining the remaining seven dimensions.

As an alternative to Parasuraman et al. (1988), Haywood-Farmer (1988) defined the 'high quality' term for a service company as meeting customer expectations consistently and grouped the attributes of the service for reducing the complexity of service characteristics. From this perspective, service attributes can be separated into three basic groups called the three Ps of service quality: physical facilities, processes, and procedures. Second are people's behavior and friendliness. Finally, the third is professional judgment. These three Ps of service quality form a triangular shape, and

a balance of each P ensures the maintenance of ‘high’ service quality. In this way, Haywood-Farmer (1988) underscores the dimensionality of the service quality concept and prefers a radar chart-like structure that eliminates the necessity for convergence to one-dimensional criteria (Figure 2.3).

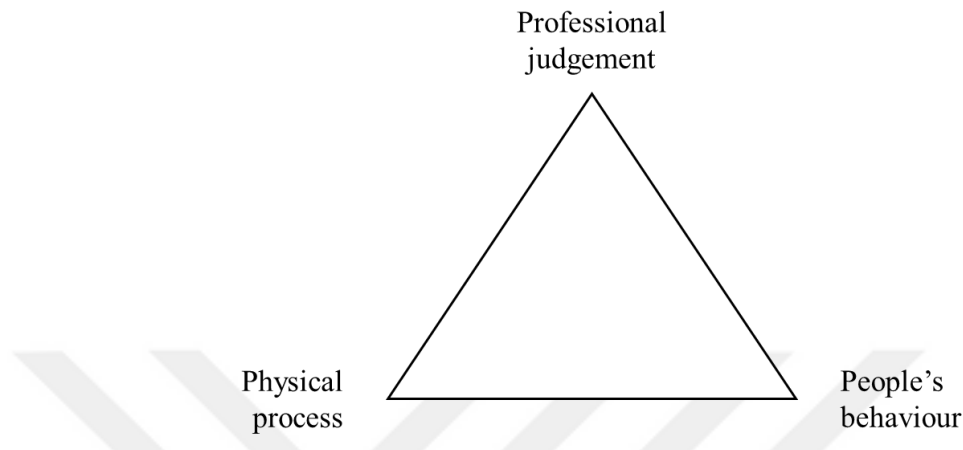


Figure 2.3. Three Ps of service quality

Source: Haywood-Farmer, 1988

Parasuraman et al. (1985) addressed perceptions that are formed via actual experiences. Nevertheless, Brogowitz et al. (1990) highlighted the importance of service quality as perceived by potential customers. Service companies must prioritize both customer attraction and customer retention. Potential clients often form opinions about the quality of service based on information they have acquired from word of mouth, advertising, or other forms of media. Therefore, a service quality gap might exist even if a client has not yet received the service. Suppose the customer believes that the service quality will not exceed their expectations. In that case, they are likely to reject the offer from a certain service provider, especially if there are better alternatives available. Hence, a Synthesized Model of Service Quality is offered to include the perceptions of potential consumers regarding the quality of service supplied, as well as the opinions of actual customers regarding the quality of service experienced (Brogowicz et al., 1990).

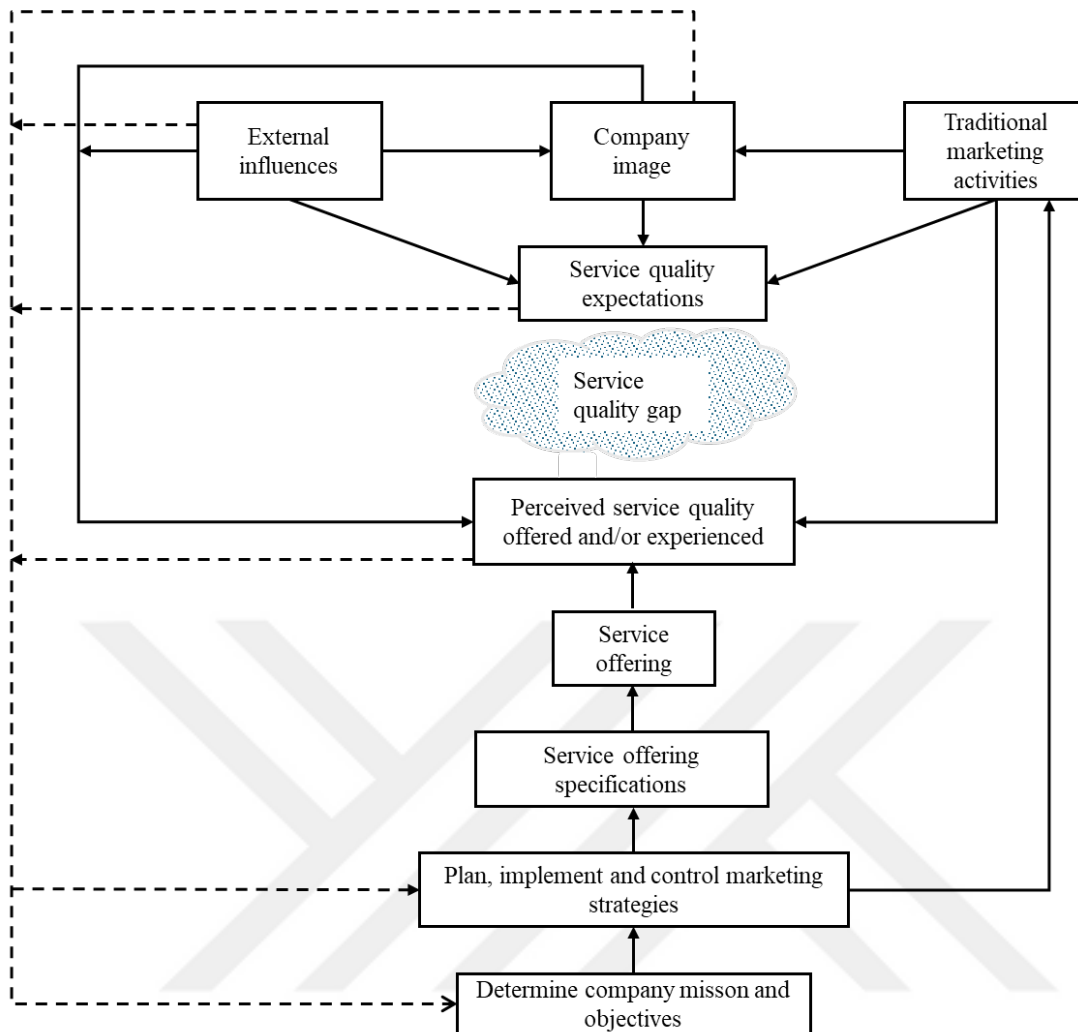


Figure 2.4. Synthesized Model of Service Quality

Source: Brogowitz et al., 1990

The approach is centered around the service quality gap, which emerges when a customer's expectations of service quality are higher than their perception of the actual service quality they receive. The service quality gap is influenced by an individual's service quality expectations, which are determined by various circumstances, as displayed in Figure 2.4. These factors encompass additional effects such as cultural norms, social hierarchy, interpersonal communication, media exposure, and market competitiveness, in addition to an individual's personal needs and previous encounters. Traditional marketing operations also impact customer expectations regarding service quality. These marketing operations aim to shape the customer's expectations of service quality, but sometimes they make promises of higher service quality than what

the service provider can actually offer, resulting in a gap between expected and actual service quality.

Cronin and Taylor (1992) proposed a model that evaluates service quality based on performance, arguing that the SERVQUAL model was inadequate. They contended that there is neither theoretical or empirical evidence establishing a connection between service quality and gap theory. Therefore, the authors offered and tested the performance-based SERVPERF scale that they have introduced as an alternative to the SERVQUAL scale. Furthermore, Cronin and Taylor (1992) investigated the connection between service quality and purchase intentions. As a third objective, they have examined the causality direction of service quality and customer satisfaction and they have asserted that satisfaction is the processor of service quality. They compared the final model, which they label as SEVPERF, with the SERVQUAL model, suggesting that SEVPERF had wider applicability and that SERVQUAL's applicability is limited to the fast food and banking sectors.

Teas (1993) investigated the issues in the conceptualization of the gap theory. According to the author, there is a substantial missing link between service quality and customer satisfaction. He proposed an Evaluated Performance Framework and Normed quality models alternative to SERVQUAL.

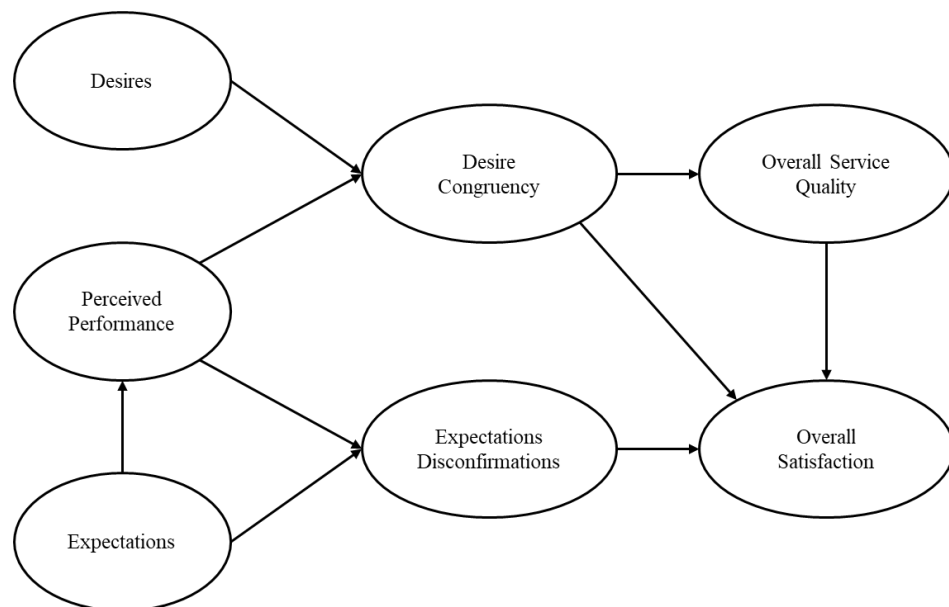


Figure 2.5. Satisfaction Service Quality Model

Source: Spreng & Mackoy, 1996

Spreng & Mackoy (1996) conducted a study that thoroughly examined many crucial components of the perceived relationship between service quality and customer satisfaction. This study aims to evaluate the differentiation between perceived service quality and customer satisfaction and investigate the influence of varying benchmarks for comparison. The model presented in Figure 2.5 demonstrates that Service Quality and Service Satisfaction are distinct yet related concepts. Additionally, when examining the antecedents influencing these two concepts, it is evident that they differ. A notable finding is that the disconfirmation of expectations does not have a direct effect on the perceived quality of service. Nevertheless, the lack of a direct impact does not mean that it is unimportant in shaping the perceived quality of service. Research has demonstrated that there is a positive indirect relationship between expectations and perceived service quality, mediated by perceived performance.

To assess the service quality from the perspective of its consumers, Philip and Hazzlet (1997) offered a hierarchical model consisting of three primary categories of attributes: Pivotal, Core, and Peripheral (P-C-P). The proposed model defines the hierarchical tiers of a service. The pivotal attributes, located at the top of the pyramid, symbolize the ultimate result that consumers anticipate and obtain after the service process is over. Thus, it is widely acknowledged that these fundamental characteristics have the greatest influence on why consumers select a specific company. The core attributes revolve around and can be most accurately defined as the components that individuals, procedures, and customers must engage with in order to attain the essential standard. The third tier of the model concentrates on "peripheral attributes", which refer to supplementary features intended to enhance the service experience and provide complete satisfaction for the user (Figure 2.6).

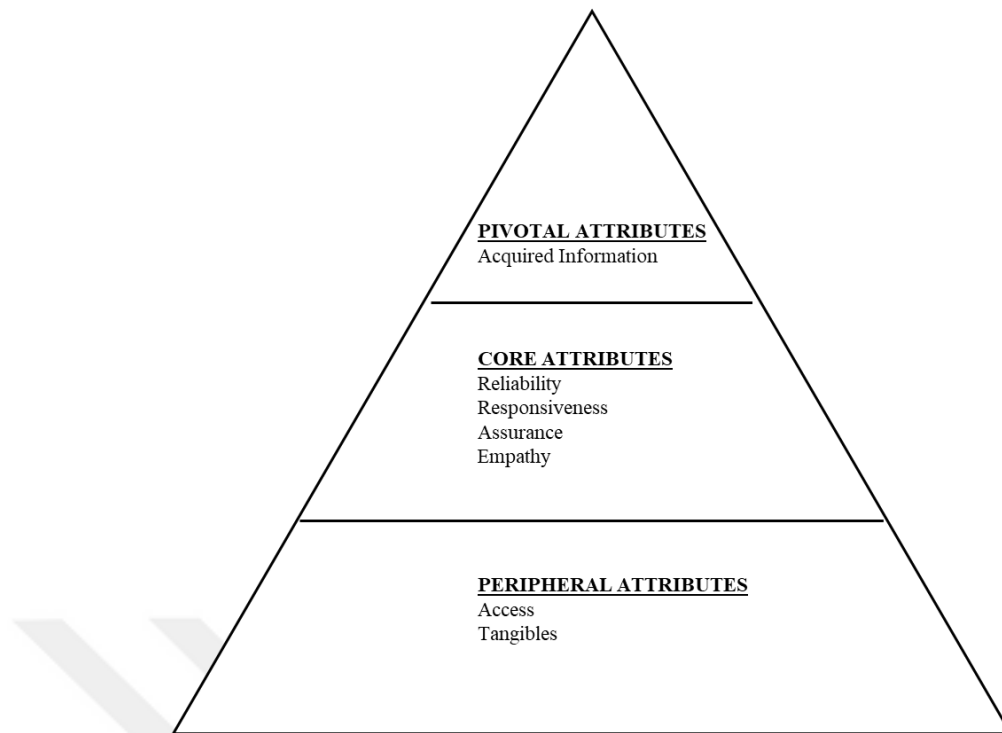


Figure 2.6. P-C-P Model for Service Quality

Source: Philip & Hazlett, 1997

The concept of P-C-P(Philip & Hazlett, 1997) suggests that consumers, when assessing their service experience, assign greater significance to the satisfaction of fundamental criteria. Therefore, it is asserted that if all components of the pivotal attributes are satisfied, indicating that the primary expected outcomes are either met or exceeded the customer will be reasonably content even if the core and peripheral qualities are realized to a lower degree. Nevertheless, it is noted that this principle may not always be valid, particularly in scenarios where the service is utilized regularly, since the fundamental and secondary characteristics may begin to assume more significance. Likewise, when the service consistently provides its main feature at a high level, users may examine other characteristics more carefully and thoroughly.

Frost and Kumar (2000) designed the internal service quality model based on the five service dimensions of SERVQUAL and their effects on internal service quality were investigated. The authors emphasized the lack of consideration about the service quality framework that consists of the internal customers and emphasized an apparent need for a conceptual internal service quality model. For this purpose, they developed an adaptation of the SERVQUAL instrument for measuring the internal service quality

called INTSERVQUAL and empirically tested it on the Singapore Airlines company's front-line staff, support services, and internal suppliers. Additionally, it was concluded that the responsiveness dimension, one of the service quality dimensions proposed by Parasuraman et al. (1988), has the most significant impact on internal service quality. In this regard, they diverge from Parasuraman et al. (1988), who emphasize that the reliability dimension has the greatest impact on perceived service quality.

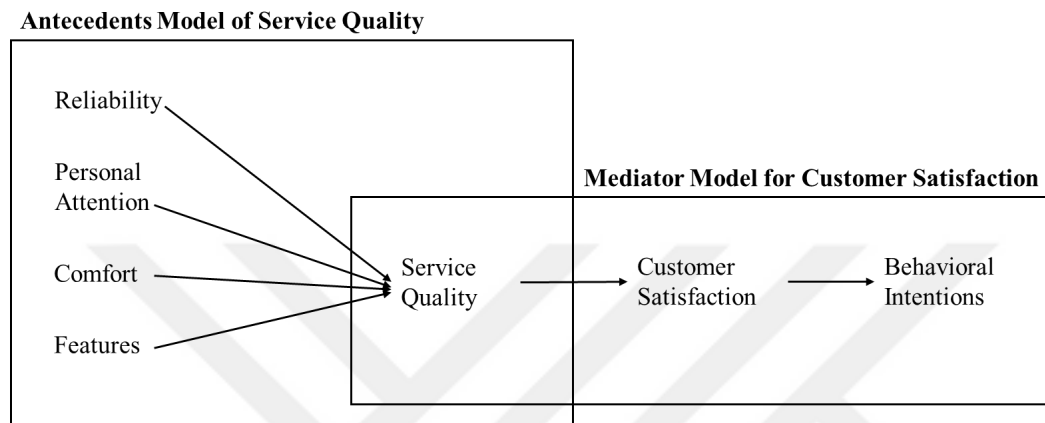


Figure 2.7. Antecedents and Mediator Model of Service Quality

Source: Dabholkar et al. ,2000

The question of whether factors influencing service quality should be seen as components that constitute the perception of service quality or as antecedents to the perception of service quality was examined by Dabholkar et al. (2000). The results of the study support the conceptual model designed with these factors as antecedents. The proposed model in the study argues that antecedent factors such as Reliability, Personal Attention, Comfort and Features influence behavioral intentions in the service quality assessment. Additionally, it is emphasized that this model gives benefits to service providers by suggesting that measuring the antecedents to predict outcomes is easier than attempting to measure the components directly. Moreover, the model incorporates the mediating role of customer satisfaction, concluding that it has a significant impact on customer behavior. This finding also supports the view in Spreng & Mackoy (1996) that customer satisfaction and service quality should be considered as distinct dimensions. Based on the paradigm depicted in Figure 2.7. The characteristics initially provided serve as precursors to service excellence. Service

quality impacts behavioral intentions, and this impact is mediated by customer satisfaction.

2.1.3. After-Sale Services and Service Quality

After-sales services are a set of supporting activities ranging from delivery, installation, maintenance/upgrade, repair, and spare parts supply of the product to warranties, technical support, solution providing, feedback, help desks, and more. The term "after-sales services" is generally used to describe all services provided after the delivery of a product. A subset of these services, often referred to as "field service," typically involves the provision of services by technical personnel across a wide geographical area. This subset is characterized by its heterogeneous nature, encompassing both customer and service-providing technical personnel perspectives (Simmons, 2001).

After-sales services offer companies opportunities to generate new revenue. The potential of these services to create new income streams necessitates their management as a part of the core operations, especially for manufacturing firms. After-sales services are provided over an extended period during the product's life cycle. Considering the duration and variety of services provided, the revenue generated from after-sales services can exceed the revenue from the product's initial sale (Knecht et al., 1993).

As a result of intense competition in the market, consumers are presented with a wide range of options when it comes to goods and services. Therefore, corporations are actively searching for methods to differentiate their products and services in order to avoid engaging in price-based competition and ensure the maintenance of their profit margins. In this regard, after-sales services are promoted as a differentiation strategy (Agnihothri et al., 2002).

Information obtained from after-sales services can provide significant opportunity for product-based innovation and improvement, as well as guide sales and promotional efforts that generate value (Saccani et al., 2006).

In this section, studies that apply quality theory to after-sales services are reviewed.

The literature on after-sales service quality can be traced back to the late 80's, when Kasper and Lemmink (1989) first coined the applicability of the methodology developed on behalf of service quality almost directly to after-sales services. In

congruence with the studies reviewed in the service quality literature review section, authors claim that after-sales services should meet with the perceptions of the customers (Kasper and Lemmink, 1989). If the company successfully translates these needs into its offerings, then the marketing concept is applied to its full meaning. Therefore, all it needs to be done is to explore customers' service quality expectations and check their congruence with the managements' opinions.

Supporting Kasper and Lemmink's (1989) proposal that service quality measures are almost directly applicable to after-sales services, Hussian, Bhatti and Jilani (2011) applied SERVQUAL methodology to after-sales service quality measurement. In particular, the authors aimed to measure service satisfaction using the dimension in SERVQUAL but primarily focused on after-sales services. Empirical findings stated in the report of Hussian, Bhatti and Jilani (2011) motivate SERVQUAL methodology to be a successful predictor for customer satisfaction for after-sales services. Related to the question proposed in this thesis, the authors state the relevance of geographical and cultural (somehow related concepts) distinctions to the effect on satisfaction levels of customers due to differences in needs and level of perception.

Alzona and Robiana (2005) relates SERVQUAL to after-sales services by underlining the effect of guarantees on the subsequent dimensions of SERVQUAL for online transactions. Even though this study lacks empirical justification for the suggestions and proposals, it presents the authors' assertion on the interrelatedness of online purchases and after-sales services, which is particularly important for the overall stance of this thesis. First, it directly emphasizes the role of after-sales services in durable goods since durable goods are being purchased online at an increasing pace. Second, customer satisfaction is again directly related to after-sales services since with online-purchase so called core services are ultimately by-passed. Third, SERVQUAL is still valid as a framework to measure overall service quality. The last line puts forward the fact that service quality methodology, despite the loss of the dominance of core activities over after-sales, stays valid.

Rigopoulou et al. (2008) assert that after-sales services are related to the technical quality of the overall service and influence the way the service is delivered. In this regard, the focus is on the functional aspect of the service quality and relates to both service products. Rigopoulou et al. (2008) further focused on the effect of overall

satisfaction on recommendation behavior. Building on the literature, Rigopoulou et al. (2008) argue that the direction of the effect needs to be from Overall Satisfaction to Recommendation Behavior, which they call Word of Mouth (WOM). The authors are particularly focused on two dimensions; the Delivery and Installation dimensions of after-sales service quality. They have concluded that after-sales services contribute to both satisfaction and behavioral intention.

In today's manufacturers, the physical product cannot be distinguished from the intangible aspects of customer service. It is not enough for the product to be technically competitive in order to provide value for customers; the dimensions of the service offered with the product must also be considered (Murali et al., 2016).

2.2. Service Quality Assessment

In Section 2.1.2, the attributes of the service, the conceptual service quality models proposed in the literature and other concepts related to services are explained. Literature conveys two main currents when referring to service quality. First, the meaning of the term: aspects of service quality. This literature mainly focuses on subdimensions. Second, measuring service quality. In addition to determining the characteristic features of service quality, creating a conceptual model, and investigating its relationship with other concepts, such as customer satisfaction, developing a tool to measure service quality has been one of the important efforts in the literature.

In this section of the thesis, the working principles of the most commonly used SERVQUAL scale in literature and its derivative SERVPERF scale, which is preferred in this study, are discussed, along with how these scales measure service quality. Following this, the weighting of service quality dimensions and the factor models and AHP method that can be used for this purpose are explained.

2.2.1. SERVQUAL

SERVQUAL is a factor model composed of latent variables and item scales developed by Parasuraman et al. (1988) to measure consumer perceptions. The scale is based on the concept of 'perceived' quality, which refers to consumers' judgments of services resulting from the comparison between their expectations and actual perceptions. The conceptual foundation of SERVQUAL is a gap model derived from exploratory

research that included focus group interviews with customers and in-depth interviews with executives (Parasuraman et al., 1985). The gap model delineates perceived service quality as the magnitude and orientation of the disparity between consumers' perceptions and expectations.

The gap model proposes that the perception of service quality can be assessed by quantifying the disparity between consumers' evaluations of their perceived service quality (P) and their anticipated service quality (E), expressed as $\text{Service Quality} = P - E$. The gap score is evaluated using the disconfirmation paradigm, where perceived quality is considered satisfactory if P equals E, exceeds satisfaction if P is larger than E, and falls below satisfaction if P is less than E.

In the initial phase of the SERVQUAL development process, the 10 key determinants of service quality identified in their previous study (Parasuraman et al., 1985) were utilized. Adopting an exploratory perspective without imposing a predetermined dimensional structure on the data, the authors subjected these 10 dimensions to an empirical “purification” procedure.

This process preserved the dimensions of Tangibles, Reliability, and Responsiveness, while the other seven determinants of service quality converged into two dimensions: Assurance and Empathy. These five dimensions are represented by a set of 22 variables/items. Since service quality was operationalized as the gap between customers' expectations (22 items) and perceptions (22 items) of performance on these variables, the service quality measurement scale comprises a total of 44 items. Dimension extraction was conducted using Exploratory Factor Analysis (EFA).

The resulting five dimensions are:

Tangibles: Refers to the physical evidence and visible aspects of the service. This dimension encompasses all material elements that reflect the quality of the service. It includes up-to-date and functional technology, equipment, and tools used in service delivery. The appearance of the service provider's personnel is also included in this dimension. Professional, clean, and well-dressed staff instill confidence in customers and reflect the seriousness of the service.

Reliability: The ability to deliver the promised service dependably, accurately, and consistently. It involves the firm's ability to charge fair fees for its services and maintain accurate records regarding its processes.

Responsiveness: The willingness and eagerness to help customers and provide prompt service. It pertains to how quickly a firm responds to customer needs and requests. This dimension also encompasses the firm's ability to provide swift and effective solutions in unexpected situations.

Assurance: Encompasses the expertise and politeness of staff members and their capacity to build confidence and comfort in customers. It signifies the possession of the necessary knowledge, skills, and experience to perform the service effectively. This dimension highlights that the quality and nature of the service depend on the competence level of the personnel providing it.

Empathy: The ability of the firm to put itself in the customer's shoes and provide individualized attention. It involves making efforts to understand customers' needs, expectations, and preferences. This dimension emphasizes the importance of providing personalized service and meeting specific customer requirements. It also considers cultural and demographic differences, ensuring that the service is appropriately tailored to different customer groups.

They state that the resulting five-dimensional structure may vary depending on the type of service considered in the study.

“[I]tems under each of the five dimensions can be suitably reworded and/or augmented to make them more germane to the context in which the instrument is to be used.”(Parasuraman et al., 1988: p28)

The distribution of items to dimensions of SERVQUAL model is depicted in Figure 2.8. Each X_j corresponds to the gap between customers' expectations and perceptions on item j . Thus, SERVQUAL's dimensions are extracted from gap data.

SERVQUAL can be operationalized to evaluate a firm's service quality by averaging the gap scores of items within each of the five dimensions. This operationalization of SERVQUAL can be expressed as follows:

$$ServQual_i = \sum_{j=1}^k (P_{ij} - E_{ij}) \quad (1)$$

where i is one of the five SERVQUAL dimensions, k is the number of items j loading to the dimension i , P_{ij} is the average perception (among respondents) of item j in dimension i , and finally E_{ij} is the average expectation (among respondents) of item j in dimension i .

Additionally, it can provide a comprehensive measure of overall service quality by calculating an average score across all five dimensions. This operationalization of SERVQUAL can be expressed as:

$$SERVQUAL = \sum_{i=1}^5 \sum_{j=1}^k (P_{ij} - E_{ij}) \quad (2)$$

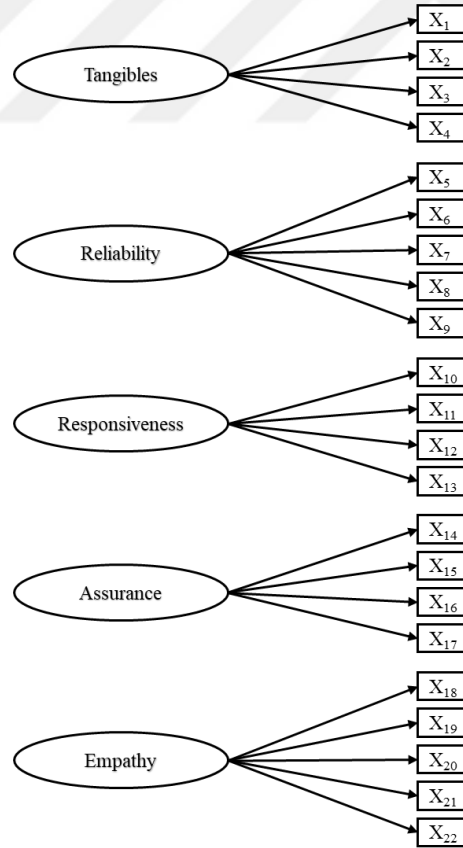


Figure 2.8. The distribution of 22 items to SERVQUAL dimensions.

Both expectations (22 items) and perceptions (22 items) of performance items can be found in the Appendix of Cronin and Taylor (1992) in Figure 2.8. It has been reproduced based on the numeration therein.

Parasuraman et al. (1988) also utilized SERVQUAL to predict an overall quality variable, which they included in their questionnaire as a general evaluation of service quality. When the dimension scores calculated in this manner were regressed against the overall quality question (rated on a four-point scale), the resulting R-squared values, while deemed respectable by the authors, were still below 50%.

Despite its widespread use, the SERVQUAL scale has faced substantial criticism on both conceptual and operational fronts. Key objections include the use of (P-E) gap scores and the validity of the five-dimension structure (Cronin & Taylor, 1992; Jain & Gupta, 2004).

The use of (P-E) gap scores in constructing the service quality measure and the validity of the five distinct service quality components remain contentious (Cronin and Taylor, 1992). Some empirical evidence indicates that the proposed delineation of these five components is inconsistent in cross-sectional analyses (Carman, 1990). It was found that certain items did not consistently load onto the same components when compared across different types of service providers.

Empirical research has found that SERVQUAL's dimensions lack discriminant validity, and there is uncertainty regarding the content validity due to overlapping conceptual definitions of certain factors (Buttle, 1996). More precisely, the aspects of 'empathy' and 'reliability' were determined to be perplexing. Although the dimensions of 'tangibles' and 'reliability' were found to be separate, the other dimensions seemed to constitute a single dimension (Buttle, 1996). Even Parasuraman et al. (1991) later research found that there might be shared items representing responsiveness and assurance dimensions. As a result, some researchers (Durvasula et al., 1999) have proposed that SERVQUAL should be redesigned to include two or three dimensions in order to more accurately measure functional quality and hence offer a more appropriate assessment of service quality.

2.2.2. SERVPERF

Cronin and Taylor (1992) introduced the SERVPERF model as an alternative to the SERVQUAL model. They removed the expectation component and focused solely on performance assessment. Additionally, they presented actual data that showed the superiority of the SERVPERF instrument over the SERVQUAL scale in certain industries.

The primary motivation for developing SERVPERF stems from forming the items in SERVQUAL dimensions based on the gaps between customers' expectations and perceptions. In this case, the items that converge on a common denominator are neither the expectations nor the perceptions but their gaps, which complicates the interpretation of the resulting structure. In other words, when a researcher applies factor analysis to expectation data in an empirical study using SERVQUAL, the factor groups obtained from expectation data may differ from those obtained from perception data. An entirely different factor formation may emerge as if this contradictory situation were not enough when each item is calculated based on the gaps between perceptions and expectations. This is the point emphasized by Cronin and Taylor (1992). Instead of explaining the most difficult of the conflicting results, the researcher should focus on the elements of perceived quality from the beginning, namely the perception component of SERVQUAL. SERVPERF assumes the five-dimensional structure of SERVQUAL with the same names but completely ignores the expectations component.

SERVPERF can be operationalized to evaluate a firm's service quality by averaging the performance scores of items within each of the five dimensions. This operationalization can be expressed as follows:

$$ServPerf_i = \sum_{j=1}^k P_{ij} \quad (3)$$

where i is one of the five SERVQUAL dimensions, k is the number of items j loading to the dimension i , and P_{ij} is the average perception (among respondents) of item j in dimension i .

The development of SERVPERF allowed for the use of modern techniques in factor analysis. Confirmatory Factor Analysis (CFA) is conducted to determine the fit of the assumed factor structure with the data, and composite scales are used for validity research (Qin & Prybutok, 2009). Since factor loadings are directly calculated based on the responses to the relevant questions, they are easier to interpret. The model results can be correlated with demographic variables (Fragoso & Espinoza, 2017). In the context of service satisfaction perception, demographic variables can be used as explanatory factors (Dimitriades & Maroudas, 2007). Additionally, SERVPERF, due to its compact structure, can be easily augmented into Structural Equation Models (Jain & Gupta, 2004).

Moreover, the integration of SERVPERF with modern statistical techniques enhances its robustness and applicability across different service industries (Devaraj et al., 2002; Leong et al., 2015). By focusing solely on the perception component, SERVPERF eliminates the potential inconsistencies arising from the expectation-perception gap, resulting in more reliable and valid measures of service quality. This focus also allows for more straightforward data collection and analysis, as it eliminates the need to separately measure and compare expectations and perceptions.

Confirmatory Factor Analysis (CFA) can be conducted to determine the fit of the assumed factor structure with the data, and composite scales can be used for validity research (Qin & Prybutok, 2009). Since factor loadings are directly calculated based on the responses to the relevant questions, they are easier to interpret. The model results can be correlated with demographic variables, providing a more nuanced understanding of the data (Fragoso & Espinoza, 2017). In the context of service satisfaction perception, demographic variables can be used as explanatory factors, which can help in identifying specific trends and patterns across different customer segments (Dimitriades & Maroudas, 2007). Additionally, SERVPERF, due to its compact structure, can be easily augmented into Structural Equation Models, allowing for more complex and comprehensive analyses (Jain & Gupta, 2004).

2.2.3. Importance Weighting – Regression, Factor Analysis & AHP

Parasuraman et al. (1988) highlight in their later work that while the SERVQUAL (and SERVPERF) dimensions are fundamental for evaluating service quality from a customer's perspective, their importance varies among customers (Zeithaml et al.,

1990). This variability necessitates differentiating the importance of each dimension in service quality assessments. Different customers prioritize different aspects of service quality, and assigning appropriate weights allows companies to capture and reflect these priorities accurately. For example, reliability may be more important to some customers than responsiveness. Recognizing and incorporating this heterogeneity enables the development of more nuanced models. Weighting dimensions based on their importance to customers also improves the precision of service quality evaluations. In this context, the weighting of dimensions is not dependent on whether SERVQUAL or SERVPERF is used. Both methods allow for the weighting of dimensions; however, the interpretation of results will undoubtedly depend on the method employed. In the following discussion, but limited to this section, we have used SERVPERF to refer to both SERVQUAL and/or SERVPERF.

Properly weighting dimensions within the SERVPERF framework allows organizations to focus on the most critical aspects of service delivery and enhances customer satisfaction by addressing or surpassing expectations in the most valued areas identified through weighted SERVPERF analysis. In this way, organizations ensure that they assess customer expectations and perceptions in the areas that matter most to them.

The weights of dimensions can be determined using various methods. Weights are often derived from customer surveys where customers rate the importance of various service quality dimensions through direct questions (Cronin and Taylor, 1992).

In other studies, techniques such as factor analysis or regression analysis are used to determine the relative importance of dimensions. To employ a regression model, it is essential to have a questionnaire item that measures overall service quality (Zeithaml et al., 1990). By regressing the $ServPerf_i$ values for each dimension on the overall service quality variable, the beta coefficients for each dimension can be found. These coefficients can be standardized to calculate item weights (Cronin & Taylor, 1992, 1994).

Some studies suggest using factor loading coefficients to calculate weighted factor scores, or in other words, weighted dimension scores (Pakdil & Aydın, 2007).

$$weightedScore_i = \sum_{j=1}^k \lambda_{ij} \times P_{ij}, \quad (4)$$

where i is a SERVPERF dimension, λ_{ij} is the factor loading j among k items that load to dimension i , and finally P_{ij} is the perception (if averaged, then averaged among respondents) of the item j in dimension i .

As an example, if items P_1, P_2, P_3 and P_4 load on a particular factor, say $i = \text{Tangibles}$, and if the responses (on a Likert scale) of a respondent for these items are $P_1 = 3$, $P_2 = 4$, $P_3 = 2$, and $P_4 = 5$ with factor loadings $\lambda_1 = 0.644, \lambda_2 = 0.782, \lambda_3 = 0.903$, and $\lambda_4 = 0.581$, the weighted score of the Tangibles dimension for that respondent is calculated as follows:

$$\begin{aligned} weightedScore_{Tangibles} &= 0.644 \times 3 + 0.782 \times 4 + 0.903 \times 2 + 0.581 \times 5 \\ &= 9.771 \end{aligned}$$

Some researchers who applied factor analytic methods to elicit dimension weights have preferred a more systematic approach by using the Structural Equation Modeling (SEM) methodology (Qin and Prybutok, 2009). Specifically, they define a second-order service quality latent variable, as shown in the figure below. This second-order latent variable, usually depicted as Service Quality, is hypothesized to explain the shared variance of the five dimensions of SERVPERF. In such a model, the computed second-order factor loadings, namely (w_i), act as dimension weights. In some recent studies, more sophisticated model structures are proposed that blends SEM methodology with machine learning techniques (Leong et al., 2015)

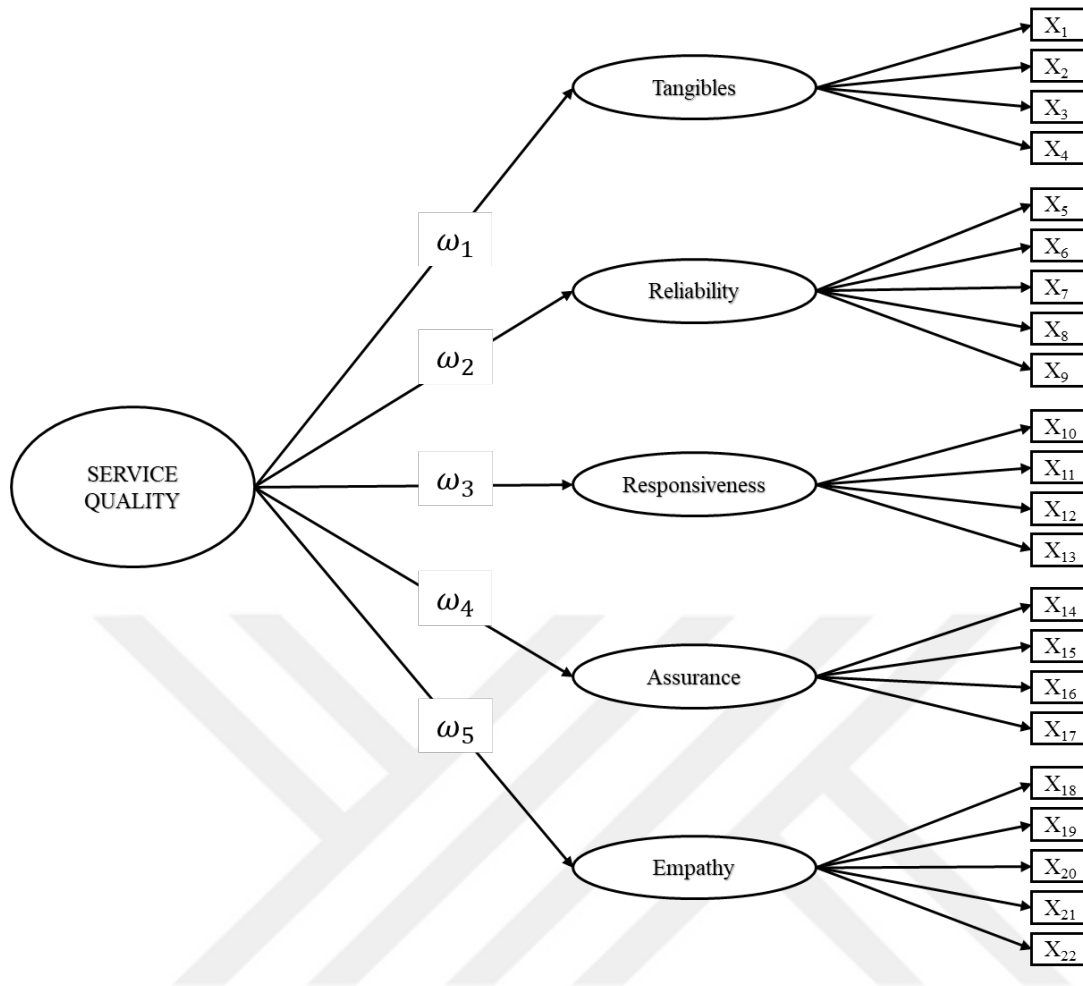


Figure 2.9. Conceptual model of a hypothetical SERVPERF model with CFA.

In the literature, most methods used to calculate weights for the five fundamental dimensions of SERVPERF are based on statistical techniques. These statistical methods require the data to be exogenous to compute these weights accurately. This requirement is typically operationalized by including item importance rating questions in SERVPERF surveys (Cronin and Taylor, 1992).

On the other hand, certain studies contend that inquiring respondents (customers) about the importance they give to the five characteristics of SERVPERF can introduce biases, hence reducing the model's acceptability (Alam & Mondal, 2019; Parasuraman et al., 1994). Based on these studies, the respondent's insufficient comprehension of the five dimensions and the surveyor's misinterpretation or biased interpretation can potentially distort the model's outcomes.

The Analytical Hierarchical Process (AHP) model can be employed to facilitate the elicitation of weights of the five fundamental dimensions of service quality based on

subject-matter experts' opinions (Alam & Mondal, 2019; Bhattacharya et al., 2023; Liu et al., 2005; Mardani et al., 2015). AHP is a robust and widely recognized methodology designed to address complex decision-making scenarios involving multiple criteria. It allows for decomposing a problem into more easily understandable sub-problems. This decomposition is particularly beneficial for weighing SERVPERF dimensions, as it enables the breakdown of each dimension into its constituent questionnaire items and even sub-dimensions, enhancing the overall decision-making process.

AHP involves a process of ranking alternatives based on how well they meet the specified criteria, generating scores for the ranking process. AHP is a method that logically and validly combines knowledge, experience, intuition, and foresight (Emrouznejad & Marra, 2017). Another significant feature of AHP is its reliance on comparative judgments or pairwise comparisons.

The AHP process consists of three stages: the formation of a hierarchy, the analysis of priorities, and the verification of consistency (Ho, 2008). In the context of using AHP with SERVQUAL, forming a hierarchy involves identifying the criteria and their sub-items, such as the five dimensions of SERVQUAL. Subsequently, the analysis of priorities entails pairwise comparisons of these criteria and/or sub-items. Decision-makers evaluate the relative importance of each pair, and these evaluations are translated into numerical values. These values are used to compute priority vectors that reflect the relative weights of each criterion. This stage enables the systematic quantification of subjective judgments, allowing for the calculation of weights that determine which aspects of service are most critical from the experts' perspectives. Finally, verifying consistency checks the consistency ratio of the pairwise comparisons. The results are considered reliable if the consistency ratio falls within an acceptable range. If not, the comparisons must be revisited (Alam & Mondal, 2019; Liu et al., 2005).

2.3. Spatial Variability in Service Quality

Companies offering after-sale services through a dispersed network of outlets serving various geographical areas face many challenges. Within these diverse regions, customers exhibit a range of attitudes and requirements influenced by factors like

cultural values (Kahle, 1986), socioeconomic status, and demographics (Mittal et al., 2004). Bronnenberg and Mahajan (2001) integrate diverse geo-demographic information and demonstrate that their geographical location significantly influences consumer reactions to pricing strategies and promotional activities. Consequently, when serving a dispersed market, it becomes essential to account for the diverse actual needs, particularly in service and product quality. One approach to addressing this complexity is segmentation, which takes into consideration the varied needs and attitudes of customers alongside their geographical locations, thereby laying the groundwork for strategies geared towards enhancing customer satisfaction.

Gentry et al. (1988) investigated the explicative potential of spatial data concerning consumer behavior constructs such as perceived risk and motivation to try new products. The data collection spanned various regions, namely Wisconsin, Washington, Oklahoma, and Massachusetts, with sample sizes of 67, 98, 189, and 135, respectively, to collect insights into geographical heterogeneity. Due to the existence of multiple dependent variables, canonical correlation analysis was employed. Their results indicate that spatial elements display higher efficacy in segmenting customers compared to diverse demographic factors. Similarly, Chen (2022) documented that consumers in different Chinese urban areas exhibit distinctive preferences. Consumers in Southern regions prioritize quality and substance, while their counterparts in the North emphasize brand and packaging. Hence, there exists evidence suggesting that consumer values, attitudes, and perceptions are significantly influenced by cultural and social variances across regions. Although this perspective incorporates geographical information into the model, geographical heterogeneity is defined as exogenous outside the model. In other words, geographical information is treated like a dummy variable within the model, stripping away the intrinsic geographical data. Proximity and distance information are removed.

In endogenous models, spatial regions are determined by the patterns present in the data rather than being predefined. To identify consumer segments for yogurt using data from 11 countries, Ter Hofstede et al. (1999) discovered four international segments associated with sociodemographics, consumption habits, media usage, and personality traits. Notably, these segments did not align with political borders. Their model captures both the differences in consumer behavior between countries and the heterogeneity within countries. The segments are represented by a finite mixture

model, making segment sizes endogenous to the model and allowing for variation across countries.

While specifying spatial data endogenous to the model enhances its explanatory power and highlights inconsistencies between segments and spatial regions, it cannot include spatial variables, as the resulting segments and spatial regions may not align. This is the main advantage of exogenous spatial information models. Hybrid approaches combine the strengths of both methods. They allow flexibility in defining regions, provided that these regions overlap with some spatially defined areas. In ecology, for instance, all "segments" align with spatial regions since spatial data can be divided almost infinitely. However, this is not the case in social sciences, where social spatial regions are not infinitely divisible.

The rates of change in physical and psychological environments are likely to vary between different regions. The psychological environment in large cities can undergo rapid changes across short distances due to the significant variation in elements such as demographic variables. Conversely, these changes are likely to occur more slowly with distance in rural areas due to the lower population density and heterogeneity. (Govind et al., 2017).

Mittal et al. (2004) adopted a spatial approach to model customer satisfaction. Earlier studies such as Mittal and Kamakura (2001) indicated that models based on customer characteristics fail to explain more than 90% of the overall customer satisfaction data variation. Mittal et al. (2004) examined customer satisfaction data by analyzing regional differences to identify systematic patterns in consumer characteristics. Their method for identifying spatial segments follows a two-step procedure:

1. Specify the region with the lowest spatial hierarchy from which spatial demographic information can be obtained. This region is considered a unit.
2. Combine these regions to form larger regions using the concept of spatial relatedness. Neighboring regions with positive correlation indicate similarity in the measured context, while those with negative correlation indicate dissimilarity.

3. CHAPTER: RESEARCH DESIGN AND METHODOLOGY

In this section, the data obtained from the HOAM (Home Appliance Manufacturer) company, basic information about the methodologies used in the thesis study, the variables used in these methods, notations, and the principles regarding the evaluation of outputs are explained.

First, the data obtained from the HOAM (Home Appliance Manufacturer) company is summarized. Second, the fundamentals, assumptions, and interpretation of model outputs of Exploratory Factor Analysis (EFA) are explained. Subsequently, the Confirmatory Factor Analysis (CFA) model, which is used to test the fit of a hypothesized factor structure to the data, along with the fit indices measuring the model's alignment with the data and suggested model modifications to improve the fit, are summarized. In the fourth section, the Analytical Hierarchy Process (AHP) method, which allows for weighting based on the opinions and experiences of experts on the obtained factor structure, is described. Finally, the principles of Geographically Weighted Regression (GWR) models, which reveal the geographical heterogeneity of a constructed model and can use this information to enhance the model's performance, are demonstrated.

3.1. Data Collection

The data primarily consists of two groups.

- Installation/Setup Service
- Repair /Maintenance Service

3.1.1. Survey Questions

The data was collected using a survey prepared by the company. The survey includes binary Yes/No questions and questions where participants are asked to select one of the Likert-type discrete rating options ranging from 1 to 5. Company officials coded the responses to the Likert-type questions on a scale of 0-100. A common set of questions is asked to participants who received either Installation or Repair services

(Appendix 1). This set includes one Yes/No question and eleven questions, ten of which are Likert-type questions ranging from 1 to 5. Following this, there are different questions for each group. The survey directed at the Installation group includes two Yes/No questions and six Likert-type questions ranging from 1 to 5. The Repair group survey includes two Yes/No questions and seven Likert-type questions ranging from 1 to 5. The complete list of questions is provided in Appendix 1. In addition to the survey questions, information on the type of service received by the participants, region, province, district, neighborhood, and product group was included. Due to an agreement made before the study, no descriptive or identifying information such as contact details, education level, or gender of the participants is shared with us.

3.1.2. Determination of the Sample and Sample Size

The company selected the respondents using a geographical stratification and quota method among customers who received service within a specific date range. Cities were designated as the stratum unit, and participants were selected through random sampling proportional to the number of customers who received services from each stratum.

In fact, the company managers' belief that customer satisfaction parameters are influenced by geographical factors can be inferred from the method they chose to select survey participants. Therefore, it is understood that they preferred the stratified sampling method. If a researcher suspects that certain population characteristics might influence the results, they may use stratified sampling in such a case (Lohr, 2021). The parameters calculated in the model through stratified sampling have a lower sampling error than those calculated using the simple random sampling method.

Company officials calculated the sample size by checking the width of the confidence interval for the means calculated for each survey question. According to the Central Limit Theorem, since the sample mean of responses to any question is usually distributed with parameters $\bar{x} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$, the sample size n is calculated using the formula:

$$n = \left(\frac{Z_{\alpha/2} S}{E} \right)^2 \quad (5)$$

In the formula, $\alpha = 5\%$, s is the standard deviation calculated from the sample, and E is the margin of error acceptable to the researchers. Of course, it is impossible to know the sample's standard deviation before collecting it, so a preliminary standard deviation value can generally be determined by conducting a pilot study. Alternatively, researchers can equate the [0-100] range in which the mean will fall under the normal distribution version to six standard deviations. However, since the company has been conducting these surveys for a while, a sample standard deviation from past data can be used in the formula to determine the sample size.

The company conducts survey studies at regular intervals, with the sample size determined by this method. Therefore, the data they provide us with comes from different times. During the thesis study, these data were periodically added, and the models were developed at each step. The last data set belongs to the 2022 period.

3.1.3. Descriptive Statistics and Data Cleaning

This section conducted analyses separately for the Repair and Installation Data groups. However, for both groups, the responses to the Likert-type questions were predominantly concentrated at the highest value (5 or 100 after conversion). It can be said that the use of survey results in determining the bonuses given to services contributed to this concentration. Nevertheless, it is not possible to analyze the data with this distribution. To partially overcome this issue, surveys where the highest value was given to all Likert-type questions were excluded from the analysis. Similarly, observations where the lowest value was given to all questions were also excluded from the analysis. The proportion of data eliminated in this way is reported in the sections providing descriptive statistics for each group.

3.1.3.1 Repair/Maintenance Service

There are 2,366 survey data points in the Repair/Maintenance group. In the first step, 1,003 observations where the lowest or highest responses were given to all Likert-scaled questions were excluded from the analysis, as they induced zero variance per record. No variability in 16 Likert-scaled items in a questionnaire is considered unrealistic. This left 1,363 survey data points to be used in the analyses.

The main characteristics of Yes/No items in the questionnaire are as follows:

- It is observed that 54.4% of the customers who received Repair/Maintenance services were not charged any fees. 1% of the participants responded that they do not remember this question.
- 60% of the survey participants reported that they were called before the repair service, while 33% did not respond to this question.
- To determine the importance of cleanliness by the service personnel, 70% of the participants responded positively to whether the personnel wore shoe covers. This question was left unanswered by 12% of the participants.
- Finally, the repair service was completed in a single visit at a rate of 67.5%.

Table 3.1. Descriptive Statistics for the Yes/No Questions

	Yes	No	NA
Were You Charged for the Service You Received?	44.7	54.4	1.0
Were You Contacted by the Authorized Service for an Appointment Regarding Your Repair Process?	60.1	6.5	33.4
Did the Authorized Service Personnel Who Visited You Wear Shoe Covers?	70.0	18.0	12.0
Was Your Repair Service Completed Correctly in One Go?	67.5	32.3	0.2

It is observed that the responses to the Likert-scaled questions are concentrated at the highest value Table 3.2. Participants chose the highest value 57% of the time, and the remaining four choices were used 43% of the time. This is naturally reflected in the average value calculated after converting the preferences made within the 1-5 range to the 0-100 range. The average value for the questions is 79.42.

The ratio of missing values is given under the NA heading in the Table 3.2, and the values are generally below 2% per item. Observations with missing values cause significant efficiency loss as they cannot be used in factor analysis and geographic modeling. The *Cost-Benefit Ratio* question, which has 57% missing values, causes significant data loss. In the previous table (Table 3.1), it was reported that 54.2% of repair services were carried out without charge. For these observations, the participant cannot be expected to calculate any cost-benefit ratio. The call center agent

administering the survey might not have asked this question after informing them that the repair was free. Therefore, the missing value problem here is systematic. For analyses like EFA and CFA where effective use of the variable is required, the *Cost-Benefit Ratio* question should be carried out with cases with no missing values.

Table 3.2. Descriptive Statistics for Likert-Type Questions

	Low	2	3	4	High	NA	Mean
Overall Service Satisfaction	11.5	3.3	8.8	29.7	46.6	0.1	74.10
Recommendation	13.1	3.5	8.3	30.2	44.6	0.2	72.30
Reputable Image	6.9	2.9	7.0	28.5	54.0	0.7	79.60
Understanding Service	5.6	1.9	5.4	22.3	64.3	0.4	84.23
Attentive Service	5.5	1.9	5.1	23.4	63.6	0.4	84.21
Knowledgeable Service	5.9	3.6	9.4	26.7	52.2	2.2	77.86
Innovative Service	6.5	2.8	9.6	27.1	46.8	7.2	72.67
Cost-Benefit Ratio	8.3	3.0	8.1	10.8	68.2	1.6	81.07
Quality Service	9.1	2.9	8.9	28.0	50.6	0.5	76.80
Fast Service	7.2	4.5	9.7	24.1	53.8	0.7	77.81
Technician's Keeping the Appointment Time	6.3	1.8	5.0	18.6	60.1	8.3	76.94
Technician Appearance	2.2	1.4	6.2	24.1	62.9	3.2	84.41
Technician Behavior and Communication (Attitude)	3.9	1.5	5.1	19.5	69.0	1.0	86.54
Knowledgeable About Faults	8.6	3.4	7.1	20.4	59.2	1.3	78.91
Technician Post-Work Cleanliness	4.0	1.7	5.3	21.1	64.7	3.2	83.62

	Low	2	3	4	High	NA	Mean
Technician's Provision of Repair Process Information	9.0	3.1	7.0	19.8	57.4	3.7	76.52
Technician Accessibility	6.3	2.1	8.3	23.3	58.3	1.7	80.48
Average	7.0	2.6	7.3	24.2	56.8	2.2	79.19

3.1.3.2 Installation/Setup Service

There are a total of 2,378 observations in the Installation/Setup Service surveys, which include 16 Likert-type questions. In 1,146 of these observations, all 16 Likert-type questions were marked with the highest option. This situation reduces the variability in the data, causing the models to struggle to explain observations with options other than the highest one due to the predominance of the highest option in most of the data. Marking the highest option for all survey questions is unrealistic and can also be considered a weakness in the survey design. As a result, 1,232 observation points remained for analysis.

Values for nominal binary (Yes/No type) questions can be summarized as follows:

- Almost 95% of the installation/setup services are free.
- 93% of the participants replied that they were called before the installation/setup service. There are no missing observations regarding this question.
- 68% of the survey participants affirmed that the service personnel wear shoe covers during sale installation/setup service.

Table 3.3. Descriptive Statistics for the Yes/No Questions

	Yes	No	NA
Were You Charged for the Service You Received?	5.5	94.5	0.0
Were You Contacted by the Authorized Service for an Appointment Regarding Your Repair Process?	92.9	7.1	0.0

	Yes	No	NA
Did the Authorized Service Personnel Who Visited You Wear Shoe Covers?	68.6	17.8	13.6

The responses to the Likert-scaled questions tend to cluster at the highest value, as shown in Table 3.4. Participants selected the highest value 60% of the time, while the other four options were chosen 40% of the time. This pattern affects the average value, which, when adjusted to a 0-100 scale, is 83.34.

The proportion of missing values is listed under the NA column in Table 3.4, typically below 2% per item. These missing values lead to significant efficiency losses since they can't be used in factor analysis and geographic modeling. Notably, with 95% missing values, the Cost-Benefit Ratio question results in unacceptable data loss and is therefore removed from the analysis.

Table 3.4. Descriptive Statistics for Likert-Type Questions

	Low	2	3	4	High	NA	Mean
Overall Service Satisfaction	2.8	2.4	5.4	32.7	56.6	0.2	87.50
Recommendation	3.8	1.7	4.8	33.6	55.7	0.4	86.88
Opt for Service	3.5	1.1	5.0	27.9	61.3	1.1	87.79
Reputable Image	1.5	1.5	5.5	29.1	60.8	1.5	88.34
Understanding Service	1.8	0.9	4.1	22.5	70.1	0.6	91.27
Attentive Service	1.9	1.1	3.7	22.8	70.0	0.5	91.30
Knowledgeable Service	2.4	1.2	6.6	27.6	60.1	2.1	87.06
Innovative Service	1.7	0.7	7.7	28.4	54.5	6.9	82.53
Cost-Benefit Ratio	1.0	0.5	0.3	1.6	2.0	94.6	3.91
Quality Service	2.4	1.1	4.6	26.0	65.7	0.2	90.19

	Low	2	3	4	High	NA	Mean
Fast Service	2.5	1.8	6.5	23.1	66.0	0.2	89.55
Technicians Keeping the Appointment Time	2.7	1.1	3.7	15.1	70.5	6.9	85.76
Technician Appearance	0.8	1.3	5.0	28.8	62.3	1.9	88.96
Technician Behavior and Communication	1.9	0.7	4.5	20.2	72.5	0.2	92.01
Knowledgeable About Product	3.2	2.1	5.9	22.2	65.7	0.8	88.57
Technician Post-Work Cleanliness	3.9	2.8	6.7	23.9	61.4	1.2	86.53
Technician Accessibility	1.9	1.3	5.8	22.6	66.2	2.2	88.64
Average	2.3%	1.4%	5.0%	24.0%	60.1%	7.1%	83.34

3.2. Exploratory Factor Analysis

EFA is a multivariate technique used to understand the relationships between observed variables, to explore how to group the variables in larger datasets in a smaller number and meaningful way, and to uncover the implicit structures underlying these variables. (Hair et al., 2010).

EFA discriminates variables into two broad categories:

Hypothetical constructs: Hypothetical constructs are sometimes labeled as factors, dimensions, latent variables, or synthetic variables (Watkins, 2018). Hypothetical constructs are conceived as unobservable attributes or qualities of individuals (Trochim et al., 2016).

Measured variables: Measured variables are sometimes called observed variables, manifest variables, indicators, or surface attributes (Watkins, 2018).

Thus, in the EFA perspective, a factor is an unseen variable that affects multiple observed measures and explains the correlations between them. Essentially, these

observed measures are connected because they share a common underlying cause (Brown, 2015).

On a technical basis, EFA involves steps of eigenvalue decomposition applied to the covariance matrix of manifest variables. This perspective coincides with a famous dimension reduction approach called Principal Component Analysis (PCA).

PCA aims to reduce dimensionality by transforming measured variables into a set of orthogonal components, which are linear combinations of these variables. The new variables, known as factors, are concise demonstrations of the original variables but are not underlying constructs (Cudeck, 2000). Conversely, the components are immediately affected by the measured variables. Thus, it is advisable for users of PCA to designate these linear combinations as components rather than factors.

Conversely, Exploratory Factor Analysis (EFA) takes a conceptually opposing perspective, positing that the interrelationships among measured variables can be accounted for by a few underlying latent factors.

Generally, EFA is employed when the structure between the manifest variables is unclear. It is a data-driven approach without any prior specification of the number of factors or the relationships between the items (Brown, 2015). Thus, EFA can be considered a method for developing a theory of construct measurement (Govindasamy et al., 2024).

3.2.1. Variables to Include

When selecting measured variables as indicators of targeted factors, the focus is on their content, convergent, and discriminant validity (Izquierdo et al., 2014). This selection process ensures that each variable accurately represents a specific hypothetical construct, aligns well with other variables measuring the same construct, and does not significantly overlap with variables measuring different constructs.

A comprehensive literature review is an essential element in establishing content validity, but it may not be sufficient on its own. The literature review facilitates the delineation of the construct and aids in identifying all pertinent dimensions and facets that should be incorporated into the measurement instrument. Nonetheless, this should be augmented with contributions from experts. It is imperative to engage a panel of

experts to assess the relevance and comprehensiveness of the items, drawing upon their extensive professional experience and expertise (Trochim et al., 2016).

The advanced stages of factor analysis address convergence and discriminant validity considerations. Convergent validity is evidenced by high loadings of indicators expected to cluster under the same factor, while discriminant validity is demonstrated by low correlation values between different factors.

It is generally recommended that a minimum of three measured variables (indicators) be associated with each factor to facilitate its statistical identification (Izquierdo et al., 2014; Watkins, 2018). The model's parameters, including factor loadings, variances, and covariances, are expected to possess unique values that are determinable from the data. This requirement helps avoid scenarios in which multiple parameter value sets could equivalently account for the observed data, thereby rendering the factor indeterminate.

3.2.2. Participants to Include

Regarding the requisite minimum sample size for factor analysis, Hair et al. (2010) have provided a rule of thumb, suggesting a minimum of 200 observations or 5 to 10 observations per factor. However, simulation studies reported by Mundfrom et al. (2005) indicate that the necessary sample size is related to the common variance or commonality among the indicators.

They recommend a minimum necessary sample size of 1200 in scenarios where there are three to six factors, each associated with three indicators, and the indicator communality values are low. When the communality values are medium, the sample size requirement falls to 500. Furthermore, the sample size requirement increases with the number of factors to be extracted. In this regard, 500 exceeds minimum requirements when the factor count would be 3.

3.2.3. Distributional Properties of Data

Deviation from normality and linearity have a major impact on the Pearson correlation coefficients between measured variables, which are crucial for calculating EFA results. Pearson correlations assume linear relationships and deviations from these

assumptions can distort the correlation matrix, leading to inaccurate factor extraction and misleading EFA findings (Watkins, 2018) .

Pearson correlation calculations for ordinal variables (such as Likert-type items) do not satisfy the linearity and normality assumptions (Fabrigar & Wegener, 2011). Instead, polychoric correlations have been extensively considered because they assume that each observed categorical variable is underlain by an unobservable, normally distributed continuous variable and estimate the Pearson correlation between those underlying hypothetical variables (Olsson, 1979). Studies using both simulated and real data have consistently shown that polychoric correlations are more likely to accurately recover the true factor model than Pearson correlations (Baglin, 2014).

Not all data sets are suitable for factor analysis. The measured variables expected to load onto the same factor should exhibit shared correlations, a requirement known as the factorability assumption. This can be assessed using the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and Bartlett's test of sphericity. The KMO test evaluates the proportion of variance in the observed variables that can be attributed to underlying factors (Kaiser, 1974). KMO values above 0.7 are considered acceptable. Bartlett's test assesses the correlation between the observed variables and the identity matrix to detect the presence of interrelationships among the variables. In order to justify the use of EFA (Bartlett, 1954), it is essential to have a chi-square value that is statistically significant. Nevertheless, due to the sensitivity of Bartlett's test to large sample sizes, it is recommended to complement its findings with metrics of sampling adequacy (Watkins, 2018).

3.2.4. Factor Extraction Methods

We have previously discussed the fundamental conceptual difference between EFA and PCA. The research question addressed in this thesis necessitates the use of the EFA approach, and the factor extraction methods employed in EFA are referred to as common method factor models (Govindasamy et al., 2024).

Each factor extraction technique has distinct objectives, procedures, and methods for computing commonalities. Maximum likelihood aims to find the factor solution that maximizes the probability of the observed data given the model, assuming multivariate normality. Principal axis factoring focuses on identifying the factors that account for

the common variance in the data is more robust to deviations from linearity and normality (Fabrigar et al., 1999). Unweighted least squares minimize the sum of squared differences between observed and reproduced correlations, providing a non-iterative solution that can handle smaller sample sizes. Minimum rank aims to find the smallest number of factors that can account for the observed correlations, often used for exploratory purposes (Govindasamy et al., 2024).

3.2.5. Number of Factors

Determining the number of factors to retain following factor extraction is influenced by both conceptual and methodological requirements. As discussed in Section 3.2.1, ensuring content validity involves conducting a comprehensive literature review and consulting experts to frame the expected factor structure. Thus, EFA is not entirely a data-driven technique; it is guided by an underlying hypothesis regarding the factor structure. Previously independent, these two frameworks converge within the EFA procedure when deciding the number of factors to retain. The decision is impacted by the congruence between the results of methods used to determine the number of factors and the hypothesized factor structure. In this context, the criteria for retention remain uncertain (Humphreys & Montanelli Jr, 1975).

Eigenvalues calculated using Principal Component Analysis (PCA) can be used as a criterion for factor retention (Carroll, 1978). In PCA, the correlation matrix is decomposed. Since the matrix has a square structure with dimensions equal to the number of measured variables, the decomposition results in as many factors as there are measured variables. The eigenvalue associated with each factor, calculated in this manner, is used to create a ratio that indicates how much of the total variance that factor explains independently of the other factors. For example, in a scenario with 30 measured variables, if an eigenvalue of 6 is associated with a factor, it means that the factor alone explains $6/30 = 20\%$ of the total variance independently of the other factors. Under this rationale, factors with eigenvalues less than one are not retained because the variance explained by such a factor is less than the variance explained by a single measured variable. Although this method is the default in many software applications, it is emphasized that it can produce erroneous results for EFA and should not be used (Fabrigar & Wegener, 2011; Izquierdo et al., 2014; Watkins, 2018).

The scree plot method, developed by (Cattell, 1966), is a graphical technique where eigenvalues are ordered from largest to smallest and plotted on a graph. Each eigenvalue is placed on the vertical axis according to its rank order on the horizontal axis. The goal is to identify a distinct "elbow" in the slope of the graph, which indicates a significant change in the rate of decline of the eigenvalues. The position of this elbow corresponds to the number of factors to be retained. (Watkins, 2018) describes this method as highly speculative.

With the increase in computing power, the parallel analysis method (Horn, 1965), an empirical approach, has become one of the most preferred algorithms for factor retention estimation in recent times (Govindasamy et al., 2024). The method begins by generating random data with the same dimensions as the original dataset. PCA is then applied to the correlation matrix of this random data to calculate eigenvalues, which are ordered from largest to smallest. This process is repeated hundreds of times, and the average eigenvalues are calculated. By comparing the ordered eigenvalues from the original data with those from the simulations, the number of eigenvalues from the original data that exceed the simulated eigenvalues determines the retention value.

3.2.6. Rotation of Factors

There are infinitely many possible solutions to the EFA, a phenomenon known as rotational indeterminacy. By rotating the factors, the loadings can be adjusted to align with a desired structure closely (Zhang, 2023). In factor rotation, the objective is to optimize the factor loads of variables to prevent measured variables from being loaded on more than one factor, thus making the conceptual framework of the factors clearer (Osborne, 2015). Factor rotation can be classified into two types: orthogonal and oblique. In orthogonal rotation, the factors remain closely associated with each other. In oblique rotation, however, the correlation between the factors can be distorted at the expense of increasing the explanation of the results. In this case, the inter-factor correlation differs from zero.

3.2.7. Interpretation of the Results

Before starting the interpretation of the results, the following checklist is controlled (Watkins, 2018):

- a. Is there a sufficient degree of freedom in the model? Or is the model overdetermined? For this, the presence of at least three measured variables that are significantly loaded to each factor is considered.
- b. Is there a single factor where each variable is loaded at a high rate? So, is there a complex or cross-loading?
- c. Is the internal reliability of each factor greater than 0.70?
- d. Are the factors consistent with conceptual expectations?

The values listed in the factor loading table (structure table) represent the correlation between the measured variable and the factor (Hair et al., 2010). Consequently, squaring the factor loadings provides the covariance between the factor and the measured variable. For example, in a scenario where the factor loading is 0.30, the factor explains approximately 10% of the variance of the measured variable. Therefore, only when the loading exceeds 0.70 does the factor explain half of the variance of the measured variable. If the factor loading value is negative, it indicates that the measured variable is inversely coded concerning the factor, according to Hair et al. (2010), the factor loadings greater than 0.7 are considered as acceptable. Statistical significance of loadings is guaranteed when the sample size exceeds 350 (Watkins, 2018).

3.3. Confirmatory Factor Analysis

CFA is a statistical method capable of testing measurement theory, which assumes how measured variables explain hypothetical constructs (Hair et al., 2010).

CFA is employed to test a hypothesized model. The hypothesized model might have been developed previously, used in practice, or recently proposed as an outcome of EFA. Researchers ascertain the number of factors and frequently mark several loadings as zero. As a result, the number of loadings calculated in CFA is lower than in EFA since not all factors are expected to have an influence on all observable variables. Variables that have loadings that are not equal to zero on a factor serve as indicators for that factor. During the process of fitting the model to the data, researchers make determinations about which parameters are considered free and can be estimated, and which parameters are fixed and limited to a given value.

Due to the greater number of parameter restrictions compared to EFA, CFA models are not subject to rotational indeterminacy (p.47). As a result, CFA does not make use of eigenvalues and eigenvectors, and factor correlations are not incorporated through rotation. Researchers determine if correlations among factors are free or fixed. In exploratory factor analysis (EFA), researchers often estimate unique variances and identify correlations among factors (Woods & Edwards, 2011).

3.3.1. Basic Definitions and Concepts

CFA measurement models typically have visual representations known as path diagrams. This section provides definitions related to CFAs and discusses elements covered in the thesis, utilizing the explanation method via diagrams as presented by Ross & Bauldry (2021) for a clear conveyance of the topic. Figure 3.1 illustrates a measurement model comprising a single construct and four measured variables. Under the widely accepted labeling convention in CFA literature, measured variables are denoted by x , constructs by ξ , and measurement errors by δ and factor loadings with λ .

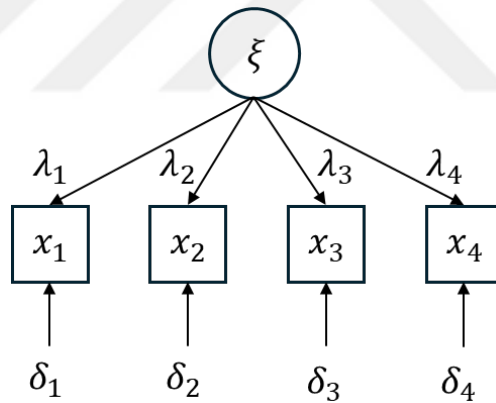


Figure 3.1. Measurement model with one construct and four measured variables

Source: Roos&Bauldry, 2021

The measurement model illustrated above can also be represented as a system of equations as:

$$\begin{aligned}
x_{1i} &= \lambda_1 \xi_i + \delta_{1i} \\
x_{2i} &= \lambda_2 \xi_i + \delta_{2i} \\
x_{3i} &= \lambda_3 \xi_i + \delta_{3i} \\
x_{4i} &= \lambda_4 \xi_i + \delta_{4i}
\end{aligned} \tag{6}$$

This measurement model specifically includes the following assumptions:

- All four variables are measurement variables of a single construct (unidimensionality).
- Each of the four variables is an imperfect measurement variable of the construct, resulting in measurement error terms.
- The construct explains the shared variance of the four variables.
- The measurement model is reflective.

As illustrated, there are a total of 8 parameters¹. The four-factor loadings and four error variances are to be estimated in the system. The system calculates these parameters using the covariance matrix of the measured variables. The covariance matrix of the four measured variables has 10 unique terms. Since the number of data points in the covariance matrix exceeds the number of parameters, the system is considered over-identified (Hair et al., 2010).

$$\begin{pmatrix}
\sigma_{x_1}^2 & \sigma_{x_1 x_2} & \sigma_{x_1 x_3} & \sigma_{x_1 x_4} \\
\sigma_{x_1 x_2} & \sigma_{x_2}^2 & \sigma_{x_2 x_3} & \sigma_{x_2 x_4} \\
\sigma_{x_1 x_3} & \sigma_{x_2 x_3} & \sigma_{x_3}^2 & \sigma_{x_3 x_4} \\
\sigma_{x_1 x_4} & \sigma_{x_2 x_4} & \sigma_{x_3 x_4} & \sigma_{x_4}^2
\end{pmatrix} \tag{7}$$

The system is considered just identified when the number of data points in the covariance matrix equals the number of parameters. In this case, there is a single solution for the parameters. For example, a measurement model with three variables associated with a single construct is just identified. If the model is over-identified, the parameters follow a chi-square (χ^2) distribution with degrees of freedom determined

¹ $(\lambda_1, \lambda_2, \lambda_3, \lambda_4, Var(\delta_{1i}), Var(\delta_{2i}), Var(\delta_{3i}), Var(\delta_{4i}))$. As in a typically regression model, means of errors are assumed to be zero and therefore only variances are parameters of interest.

by excess data points. In the example of the measurement model shown in Figure 3.1, the degrees of freedom is 2 (Hair et al., 2010).

If an additional construct is needed to explain the shared variance of the measured variables, the measurement model may take the form exemplified in Figure 3.2. This measurement model consists of two sub-models, each comprising two measured variables associated with a single construct.

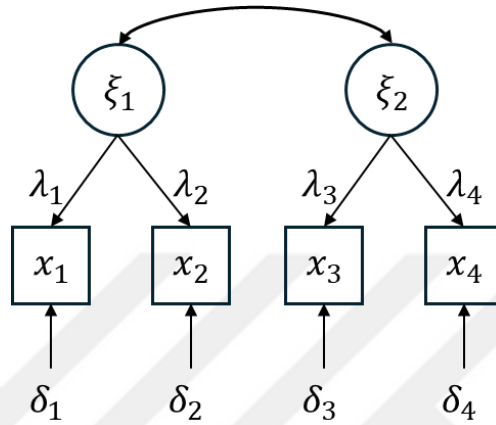


Figure 3.2. Measurement model with two constructs and four measured variables

Source: Roos&Bauldry, 2021

The system of equations describing the model is quite similar to the previous one, and it also includes four-factor loadings and four error variance parameters.

$$\begin{aligned}
 x_{11} &= \lambda_1 \xi_1 + \delta_{11} \\
 x_{21} &= \lambda_2 \xi_1 + \delta_{21} \\
 x_{32} &= \lambda_3 \xi_2 + \delta_{32} \\
 x_{42} &= \lambda_4 \xi_2 + \delta_{42}
 \end{aligned}
 \tag{8}$$

However, an additional parameter, the covariance between constructs, is to be estimated in this system. This measurement model is under-identified, yet it aids in the straightforward communication of some essential elements within the scope of CFA.

The measurement model illustrated in Figure 3.2 is established with the following hypotheses:

- Each sub-model possesses the characteristics of the model shown in Figure 3.1.

- Additionally, the shared variance between x_1 and x_2 is greater than the shared variance between x_1 and x_3 or x_1 and x_4 . This property is similarly applicable to the other measured variables.

In practice, situations where a measured variable is associated with multiple constructs can occur (Figure 3.3). This issue is also addressed in this thesis (p.78). Measurement models where measured variable loads onto n constructs are called "factor complexity n (FC n)."

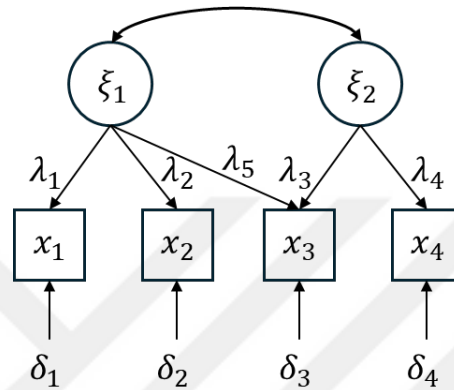


Figure 3.3. FC2 Measurement model with two constructs and four variables

Source: Roos&Bauldry, 2021

In measurement models like the FC2 example shown above, identification problems or estimation errors such as negative error variances, known as Heywood cases (Hair et al., 2010), can occur.

Another common issue encountered in CFA practice is correlated measurement errors. In the example shown in Figure 3.4, there is shared variance between x_3 and x_4 that the construct cannot explain. The mathematical breakdown of the model is exactly the same as the models shown in Figure 3.1, but an additional parameter to be estimated is the covariance between x_3 and x_4 , resulting in the loss of one degree of freedom.

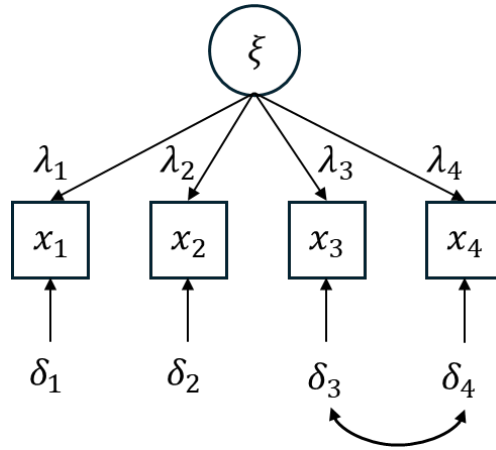


Figure 3.4. One construct and four measured variables and correlated errors

Source: Roos&Bauldry, 2021

Up to this point, most of the fundamental concepts within the scope of CFA have been discussed. Two more definitions need to be addressed before concluding this section. The first is the concept of the nested measurement model. In CFA studies, alternative measurement models are often compared. If the parameters to be estimated (factor loadings, error variances, and covariances) in one measurement model are a strict subset of the parameters in another model, the simpler model is said to be *nested* within the larger one. For example, the measurement model shown in Figure 3.1 is nested within the measurement model in Figure 3.4. Additionally, the measurement model with the most parameters to be estimated is called the *unrestricted model*. Models nested within the unrestricted model are referred to as *restricted models*.

Finally, if a measurement model has no cross-loadings, i.e., it is not FCn (Figure 3.3), and there are no within or between covariances among error variances (Figure 3.4), it is called a *congeneric measurement model*. The measurement models in the examples shown in Figure 3.1 and Figure 3.2 are congeneric. As discussed below (Hair et al., 2010), congeneric models are preferred within the construct validity framework.

3.3.2. Developing & Estimating Measurement Model

CFA requires a prespecified measurement model and is initialized with the selection of an estimation method (Nye, 2023). Most software defaults to the maximum likelihood (ML) method at this stage. Beyond the statistical advantages provided by ML, the normality and linearity assumptions for all measured variables apply to CFA, and they apply to EFA. Additionally, ML is only suitable for continuously scaled

variables. However, in social sciences, most studies (including this thesis) use ordinal scale measured variables such as responses to survey items, psychological questionnaires, and educational tests, and the scale of data is essentially discrete. A typical factor analysis model applied to ordinal data results in model misspecification due to the inability to establish a linear link between discrete ordinal data and a continuous factor (Shi et al., 2020).

When the data fails to meet the assumptions required by ML, biased parameter estimation, incorrectly specified errors, and inaccurate fit index calculations may occur (DiStefano, 2002; Nye, 2023). Therefore, the ML method should only be used for continuous data that has marginal deviations from normality for each measured variable (Rogers, 2024). In this regard, CFA researchers are expected to document, justify, and report the estimation method they have used in their studies (Crede & Harms, 2019).

Different estimation methods have been proposed to overcome this problem. Among these, unweighted least squares (ULS) and diagonally weighted least squares (DWLS) are preferred because they can work with small sample sizes and offer statistical advantages (Nye, 2023; Nye & Drasgow, 2011). At the time of writing this thesis, there are two freeware statistical software packages where ULS and DWLS can be used as estimation methods in CFA; JASP (JASP Team, 2024), Jamovi (The jamovi project, 2024) that offer user-friendly, SPSS-like interfaces, eliminating the need for programming skills.

3.3.3. Model Evaluation

The model evaluation consists of two components. The first component involves testing and assessing how well the measurement model reproduces the means, variances, and covariances among the measured variables. For this purpose, various overall model fit metrics can be used, each with its own strengths and weaknesses. The second component involves evaluating various parameter estimates to identify more precise or localized issues with model fit that cannot be detected by overall model fit metrics alone (Hair et al., 2010; Nye, 2023; Nye & Drasgow, 2011; Roos & Bauldry, 2021).

This thesis adheres to the warnings from numerous recent studies that advocate against the use of Pearson correlation in factor analytic methods for Likert-type ordinal scaled discrete data. Instead, it employs polychoric correlation in the development of the EFA model (see p.44). Accordingly, ULS and DWLS methods were preferred in the CFA models. However, it has been noted that the use of these methods might lead to incorrect calculations of some fit indices (Shi et al., 2020). The warnings regarding fit indices are explained in each respective subsection.

3.3.3.1 Chi-square (χ^2) Fit Index

The null hypothesis of the chi-square (χ^2) fit index posits that the disparity between the covariance matrix anticipated by the model and the observed covariance matrix for the variables in the model is equal to zero. Consequently, elevated χ^2 values signify inadequate conformity (Nye, 2023). Low χ^2 values compared to degrees of freedom (df), indicate a good fit with the data.

The fundamental weakness of the chi-square fit index is that it is immensely sensitive to sample size, assumption of normality, and estimation technique (Nye & Drasgow, 2011; Xia & Yang, 2019). Nevertheless, the most widely used fit indices in the literature, such as CFI, RMSEA, and SRMR, are functions of the chi-square fit index.

3.3.3.2 Root Mean Square Error of Approximation (RMSEA)

RMSEA is currently the most widely used (mis)fit index that applies factor analytic models such as CFA or Structural Equation Models (SEM) (Shi et al., 2020).

$$RMSEA = \sqrt{\frac{F_0}{df}} = \sqrt{\frac{\chi^2 - df}{N \times df}} \quad (9)$$

where F_0 is the aforementioned discrepancy, df is degrees of freedom, χ^2 is the chi-square test statistic of the model fit and N is the sample size (Hayashi et al., 2007).

The RMSEA has no upper limit but has a minimum of zero. Since it is a misfit index, lower RMSEA values indicate a better fit with the data. RMSEA values below 0.06 are considered indicative of a good fit. However, RMSEA is not a standardized fit index. Consequently, any RMSEA value must be interpreted differently for different CFA structures, as it has different implications depending on these factors. (Groskurth et al., 2023; Maydeu-Olivares, 2017; McNeish, 2023).

Shi et al. (2020) claim that the RMSEA calculation in current software is biased when polychoric correlations are used for Likert-type data with either ULS or DWLS as estimation methods in CFA. For the ULS fitting function;

$$RMSEA_{ULS} = \sqrt{\frac{1}{df} \sum_{i \leq j} (\rho_{x_i x_j} - \rho_{x_i x_j}^0)^2}, \quad (10)$$

where $\rho_{x_i x_j}$ is the polychoric correlation between variables i and j and $\rho_{x_i x_j}^0$ is the polychoric correlation under the fitted model.

3.3.3.3 Standardized Root Mean Square Residual (SRMR)

This is essentially a scaled version of the chi-square (χ^2) fit index and is based on the difference between the observed and estimated correlation matrices. It can be roughly interpreted as the average standardized residual correlation (Hayashi et al., 2007; Maydeu-Olivares, 2017).

$$SRMR = \sqrt{\frac{1}{t} \sum_{i \leq j} \frac{(\sigma_{x_i x_j} - \sigma_{x_i x_j}^0)^2}{\sqrt{\sigma_{x_i x_i} \sigma_{x_j x_j}}}}, \quad (11)$$

where $t = p(p + 1)/2$, p is the number of measured variables, $\sigma_{x_i x_j}$ denotes population covariance between variables i and j (or variance if $i = j$) and finally $\sigma_{x_i x_j}^0$ denotes the covariance (or variance) under the fitted model.

The SRMR takes values between 0 and 1, with smaller SRMR values indicating a better fit of the model to the data. SRMR values $\leq .08$ are considered to indicate a moderate fit, while SRMR values $\leq .06$ are required for a good fit (Hu & Bentler, 1999). Contemporary research suggests that SRMR is more robust against differing estimation methods when compared to RMSEA (Shi & Maydeu-Olivares, 2020).

Shi et al. (2020) claim that the SRMR calculation in current software is biased when polychoric correlations are used for Likert-type data with either ULS or DWLS as estimation methods in CFA.

$$SRMR_{ULS} = \sqrt{\frac{1}{t} \sum_{i \leq j} (\rho_{x_i x_j} - \rho_{x_i x_j}^0)^2} \quad (12)$$

This thesis follows Shi et al.'s (2020) advice and calculates unbiased RMSEA and SRMR by using their R code, provided as supplementary material.

3.3.3.4 Incremental Fit Indices (TLI & CFI)

The fundamental rationale underlying incremental fit indices involves comparing the model being tested with a baseline model that assumes the measured variable covariances are zero. Therefore, in some sources baseline models are also referred to as independence models. If the covariances between measured variables are zero, then the concept of shared variance is not applicable; in other words, a factor analytic structure is not present. This is clearly a void model and the model to be tested is expected to perform superior to the baseline model. There are two most widely used incremental fit indices (Roos & Bauldry, 2021):

1. **Tucker-Lewis Index (TLI):** The Chi-square test value of the baseline model is χ_b^2 with df_b degrees of freedom whereas for test model it is χ_m^2 with df_m degrees of freedom.

$$TLI = \frac{\left(\frac{\chi_b^2}{df_b}\right) - \left(\frac{\chi_m^2}{df_m}\right)}{\left(\frac{\chi_b^2}{df_b}\right) - 1} \quad (13)$$

From a dataset appropriate for factor analytic models, we expect that $\chi_b^2 \gg df_b > df_m > \chi_m^2$. Therefore, for a model that has a good fit with the data, $\left(\frac{\chi_m^2}{df_m}\right)$ is expected to be greater than one. Consequently, TLI would be very close to 1. Simulation studies suggest that a good model data fit is indicated by $TLI \geq 0.95$ and a moderate fit is indicated by $TLI \geq 0.90$ (Hu & Bentler, 1999).

2. **Comparative Fit Index (CFI):**

$$CFI = 1 - \frac{\max[(\chi_m^2 - df_m), 0]}{\max[(\chi_m^2 - df_m), (\chi_b^2 - df_b), 0]} \quad (14)$$

For a good model, we would expect the value of $(\chi_m^2 - df_m)$ to be small and possibly even negative. In the worst case, the tested model would be as poor as the baseline model. Consequently, CFI takes values between zero and one. The same cut-off points apply also for CFI; a good model data fit is indicated by $CFI \geq 0.95$ and a moderate fit is indicated by $CFI \geq 0.90$ (Hu & Bentler, 1999).

3.3.3.5 Information Criterion-Based Fit Indices

There are a couple of information criteria-based indices that can be used to compare the fit of multiple models, as discussed in Section 3.3.4. One of these indices is the Schwartz Modified Bayesian Information Criterion (SBIC), which can be calculated using the chi-square statistics as follows:

$$SBIC = \chi^2 - df * \ln(N) \quad (15)$$

The working principle of SBIC is based on comparing the current model to the saturated model. The saturated model is one in which all measured variables are correlated with each other. This is the exact opposite of the baseline model described above. In the saturated model, the degrees of freedom are zero, and there are no free parameters. Despite the presence of a factor analytic structure in the saturated model, there is a single factor to which all variables are connected, and the common variance is shared through this factor. The SBIC value of the saturated model is zero, and the lower the SBIC value, the better the model fits the data. Therefore, if the SBIC value for the tested model is negative, it means that the model fits the data more successfully compared to the saturated model. Conversely, if the SBIC value is positive, it indicates a poorer fit compared to the saturated model.

3.3.4. Modifying the Measurement Model

A key principle in CFA research is to not only evaluate the proposed model, but also to estimate several other feasible alternative models. (Nye, 2023). As discussed in Section 3.3.3 above, the fit indices used to evaluate CFA only determine the overall fit of the model, but they do not indicate that the hypothesized model is the best one, except for extremely rare cases where CFI is identical to one. While a specific model

may appear to have a strong fit based on conventional criteria for assessing model fit, it is possible that a slightly modified model could offer a more precise depiction of the data. An essential aspect to consider is the exploration and examination of several alternative models in order to conduct a thorough and comprehensive assessment of the proposed model (Nye, 2023). There are two different approaches to testing with alternative models: chi-square difference tests and information criterion tests.

To explain this method, the concepts of restricted and unrestricted models defined in Section 3.3.1 are needed. Briefly, a model in which the set of parameters (factor loadings, error variances, and covariances) is a subset of the parameter set of another model is defined as a restricted model. According to this definition, the restricted model is nested within the unrestricted model. When the restricted model is tested against the unrestricted model using the chi-square difference test if the restricted model achieves a similar fit with fewer parameters, the principle of parsimony dictates that the restricted model is preferred (Roos & Bauldry, 2021). The test statistic is defined as follows:

$$\begin{aligned}\chi_d^2 &= \chi_R^2 - \chi_U^2 \\ df_d &= df_R - df_U\end{aligned}\tag{16}$$

χ_R^2 (χ_U^2): The chi-square test statistic for the restricted (unrestricted) model.

df_R (df_U): Degrees of freedom of test statistic for restricted (unrestricted) model.

In chi-square difference tests, a modification suggestion is developed for a single parameter at each step, which is why this test is also referred to as modification indices in the literature (Hair et al., 2010; Roos & Bauldry, 2021). Since a change in a single parameter will result in a difference in degrees of freedom, modern software filters changes that exceed 3.84 in the chi-square distribution and presents them as options to the researcher.

One form of model modification involves the use of correlated error terms (Figure 3.4). This problem is frequently encountered in CFA analysis, and it may be acceptable to incorporate the correlation of error terms for specific indicators. In a typical CFA model, responses to each indicator are determined by one or more underlying causes and measurement inaccuracies. The measurement error in this model accounts for the distinct variability in the indicator that is not accounted for by the underlying causes.

Hence, when considering the underlying factors, if two or more indicators have correlated measurement errors, it indicates that they also have additional shared variation. It is important to discuss the practice of correlating error terms because it is generally not recommended in CFA (Nye, 2023). However, it is recommended to add correlated errors with theoretically justified reasons and to keep their number to a minimum. For example, if two survey questions have roughly the same meaning from the respondent's perspective, there is theoretical support for allowing a correlation between the two questions (Nye, 2023; Rogers, 2024).

If the compared measurement models are not nested, then it is not possible to use the chi-square difference test. In this case, information criteria such as Akaike Information Criterion (AIC) or Bayesian Information Criterion (BIC) can be used for comparison. According to this criterion, the model with a lower AIC or BIC value should be preferred (Roos & Bauldry, 2021).

3.3.5. Construct Validity

Construct validity involves evaluating whether the theoretically defined and developed construct is effectively captured and reflected by the specified measurement variables in the model. Construct validity ensures that the instrument accurately reflects the construct, encompasses all its dimensions through sharing a high proportion of variance in common, and distinguishes it from other related constructs (Hair et al., 2010; Westen & Rosenthal, 2003).

Validating or rejecting a theory through data demands the validation or rejection of the measurement instrument derived from that theory. Within this framework, factor analytic methods create statistical structures of the construct- set of measured variable relationships derived from the theory. Consequently, construct validity determines whether there is a higher correlation among items within different subsets compared to other subsets by examining the dimensionality among constructs (Tabachnick et al., 2013). Hence construct validity has two main components: convergent validity and discriminant validity.

3.3.5.1 Convergent Validity

Convergent validity concerns whether the construct explains a high proportion of the shared variance of the variables it measures. For convergent validity, all factor loadings should be significant, at minimum.

On the other hand, it is not sufficient for factor loadings to be significant; the standardized values must also be high. A high coefficient indicates that a specific factor explains the item well. According to studies using simulation methods, in terms of convergent validity, factor loadings above 0.70 are considered excellent, above 0.63 are very good, above 0.55 are good, above 0.4 are fair, and above 0.32 are poor (DiStefano & Hess, 2005).

Average Variance Extracted (AVE), developed as a measure of the extent to which a construct explains shared variance of the items to which it is structurally connected, is also used in assessing convergent validity. AVE is the average of the squared factor loadings and is calculated for each construct.

$$AVE = \frac{\sum_i \lambda_i^2}{n} \quad (17)$$

where n is the number of items of a construct and λ_i is a standardized factor loading. AVE scores of each construct should be greater than 0.5.

Reliability is also regarded as a measure of construct validity (Hair et al., 2010). There are two most widely used reliability measures in the literature: Cronbach's alpha, which is generally used in EFA, and construct reliability, which is used in CFA.

The formula for Cronbach's alpha does not take factor loadings into account, or in other words, it assumes that factor loadings are equal for all items (called tau-equivalence).

$$\alpha_i = \frac{\sum \sigma_{ij}}{\sigma_i^2 + \sum \sigma_{ij} \times \frac{n-1}{n}} = \frac{n \times \overline{\sigma_{ij}}}{\sigma_i^2 + (n-1) \times \overline{\sigma_{ij}}} \quad (18)$$

where σ_{ij} is the covariance between item i and item j , σ_i^2 is the variance of item i , n is the number of items related to construction and $\overline{\sigma_{ij}}$ is the average covariance of the items of the construct. Cronbach's alpha less than 0.6 is deemed unacceptable, 0.60 to

0.65 as undesirable, 0.65 to 0.70 as acceptable, 0.70 to 0.80 as good, and finally greater than 0.80 as very good (DeVellis & Thorpe, 2021).

In literature, construct reliability is sometimes referred to as composite reliability or coefficient omega (Cheung et al., 2024; Hair et al., 2010).

$$CR = \frac{(\sum_{i=1}^n \lambda_i)^2}{(\sum_{i=1}^n \lambda_i)^2 + \sum_{i=1}^n (1 - \lambda_i^2)} \quad (19)$$

where n is the number of items of a construct and λ_i is a standardized factor loading. CR values are expected to be greater than 0.7.

3.3.5.2 Discriminant Validity

Discriminant validity examines the extent to which a construct is distinct from other constructs in the model (Hair et al., 2010). Therefore, discriminant validity is a characteristic that all constructs in a measurement model should demonstrate. It can be measured in two ways.

First, a model is established in which the correlations between the constructs are set to one, and its performance is compared with the original model. The original model is expected to perform significantly better than this restricted model. However, there are objections to the effectiveness of this method in adequately measuring discriminant validity (Hair et al., 2010). This is because a correlation of 0.9 between constructs would also be high enough to threaten discriminant validity, and testing such situations is practically impossible, especially when the number of constructs exceeds two.

A more rigorous method for measuring discriminant validity involves comparing the Average Variance Extracted (AVE) values of two constructs with the square of the correlation between them. The logic behind this test is that a construct's ability to explain the shared variance of its associated measured variables should be higher than its ability to explain the shared variance of unrelated measured variables (hence, the correlation is squared).

Finally, discriminant validity necessitates that each indicator loads exclusively onto a single construct. If the same measured variable is utilized to measure multiple constructs, it becomes challenging to assert that the constructs are distinct. Unidimensionality requires the absence of cross-loaded indicators, which is essential for attributing meaning to a latent construct (Gerbing & Anderson, 1988).

3.4. Analytic Hierarchy Process (AHP)

The Analytic Hierarchy Process (AHP) is a methodology in decision-making developed by Saaty (1977). This method is capable of handling both rational and intuitive aspects in order to choose the most optimal option from a range of possibilities, which are evaluated based on multiple criteria. During this procedure, the individual responsible for making decisions makes straightforward comparisons between pairs of options. These comparisons are then utilized to establish the general priorities for rating the different alternatives (Saaty & Vargas, 2001).

AHP is defined in a hierarchical structure consisting of at least three levels. The top level represents the objective. Below that, there are sub-criteria and at the lowest level, decision alternatives. When defining criteria, they can be classified according to their common characteristics. This definition stage is also very important in ensuring the consistency of pairwise comparisons between these criteria set. In addition, since the data collected in pairwise comparisons contains subjective information, the participation of experienced and knowledgeable experts should be ensured.

Wind and Saaty (1980) explain the methodology in three main parts:

Step 1: Problem definition with a hierarchical manner

AHP begins by breaking down a complicated problem into a hierarchy. Each level of the hierarchy contains a small number of manageable criteria, and each criterion is further broken down into another set of elements. The process extends to the most precise components of the problem. Organizing a decision problem hierarchically is an effective method for managing complexity and finding the key elements of the problem. The AHP does not have a fixed hierarchical structure. One can create a hierarchy that suits their own requirements.

Step 2: Determine a measurement scale and collect expert opinions

This step aims to determine the relative importance of the elements within each level of the hierarchy. This is achieved by requesting the experts involved to assess each group of items pairwise. The hierarchy is divided into a sequence of pair comparison matrices, where experts assess the relationship in one half of each matrix. Each pair is assessed individually based on the extent to which one item surpasses the other in terms of the elements from the higher level in the hierarchy. To make accurate

numerical assessments during pairwise comparisons, it is essential to have a dependable and practical scale. The scale commonly employed in analytic hierarchy investigations is as follows:

Table 3.5. The typical scale for AHP

Importance Level	Verbal Judgement
1	Equally important
3	Weakly important
5	Strongly important
7	Very strongly important
9	Absolutely important
2,4,6,8	Intermediate values
Reciprocals	When activity i is compared to activity j, and activity j is assigned one of the nonzero numbers listed above, then activity j is assigned the reciprocal value while compared with activity i.

Source: Saaty, 2004

Step 3: Determine the priorities of the hierarchy and check the consistency of the data

The main objective of this phase of the methodology is to compute the priorities that arise from pairwise comparisons of each element within each level of the hierarchy.

Let for n objects with known weight vector of $w = (w_1, w_2, \dots, w_n)$. Then pairwise comparison of wieghts can be written as follows (Wind & Saaty, 1980):

$$A = \begin{bmatrix} w_1/w_1 & \cdots & w_1/w_n \\ \vdots & \ddots & \vdots \\ w_n/w_1 & \cdots & w_n/w_n \end{bmatrix} \quad (20)$$

The vector of weights w can be found by multiplying A on the right side by w :

$$Aw = nw \quad (21)$$

n is the largest eigenvalue of A . Since n is the eigenvalue of $Aw = nw$ has a nontrivial solution. The ratios w_i/w_j are not known; however, they can be estimated from expert opinions. The problem of estimating the weight vector of w can be solved by solving the problem $Aw = \lambda_{max}w$. A pairwise comparison reciprocal matrix is utilized to compare the relative contributions of the components in each level of the hierarchy to an element in the upper level that is adjacent to them. A subsequent step involves deriving the principal eigenvector of this matrix, which is then weighted according to the priority of the property being compared.

In a matrix A , instead of assigning two unknown numbers w_i and w_j , a single number from the fundamental scale of absolute numbers shown in Table 3.5 can be assigned to represent the ratio w_i/w_j which is approximated by the nearest integer. The extraction of the principal eigenvector and the hierarchical weighting procedure results in a one-dimensional scale that represents the priority of elements at each level of the hierarchy. The resulting priorities reflect the respondent's subjective view of the relative importance of the elements in the hierarchy.

In order for priorities to be meaningful, they must be produced from matrices that are consistent. Therefore, it is necessary to perform a consistency check that can be calculated with the consistency index (CI) introduced by Saaty (1977).

$$CI = \frac{\lambda_{max} - n}{n - 1} \quad (22)$$

where n is the dimension of the matrix A . The consistency ratio (CR) is given equation:

$$CR = \frac{CI}{RI} \quad (23)$$

A CR value below 10% indicates an acceptable level of consistency (Ishizaka & Labib, 2011). The random index (RI) is calculated as the average of the composite index (CI) of 50,000 randomly filled matrices and the Table 3.6 shows the random indices from Saaty (2004).

Table 3.6. The Random Index (RI)

n	1	2	3	4	5	6	7	8	9	10
RI	0	0	.52	.89	1.11	1.25	1.35	1.40	1.45	1.49

In this study, AHP is employed to assess the relative significance of the questionnaire items and latent variables after EFA and CFA.

3.5. Geographically Weighted Regression (GWR)

In the empirical models used in the marketing science literature, the key assumption is that individuals are conditionally independent of each other, and therefore, information about one individual cannot be used to make inferences about another individual. This conditional independence allows for the pooling of data collected at the individual level, and parameters can be estimated using maximum likelihood principles. However, in time-stamped data, conditional independence cannot be assumed. Knowledge about observations that are close to each other in time affects the predictions made about each other, indicating a correlation along the time axis. Thus, when modeling time series, treating data independently of the temporal axis by shuffling the temporal ranking can lead to biased parameter estimates. Similarly, in spatial models, the location of an observation relative to other observations negates its conditional independence and introduces correlation along the spatial axis. Spatial dependence implies that observations located near each other influence each other. Furthermore, the influence is itself a function of the distance between observations. Therefore, the premise of spatial models is to identify the spatial sensitivity of parameters rather than pooling the data. Otherwise, the error terms of the model will carry spatial information, leading to issues of heteroskedasticity (Bradlow et al., 2005; Mittal et al., 2004).

Classical regression models that allow the pooling of data to take the well-known form,

$$Y = X\beta + \varepsilon \quad (24)$$

where Y is the dependent variable, X is the set of covariates, β is the set of coefficients and finally ε is the unexplained part of the dependent variable not explained by the covariates. Here the model coefficients β are constant throughout the dataset.

If it is believed that the model fails to capture the heterogeneity across the spatial distribution of observations, a geographically weighted regression (GWR) should be employed in place of the classical regression model. To effectively capture spatial heterogeneity, the coefficients in GWR vary depending on the spatial location.

$$Y = X\beta(g) + \varepsilon \quad (25)$$

where g represents some form of a geographic information vector, possibly taking the form of latitude-longitude pairs. This indicates that the parameters of the model vary across different regions specified by this vector. By incorporating g , the model can account for the spatial variability in the data, allowing analysis of regional differences.

The operationalization of geographic information is achieved through the weighting of observations. In traditional regression methodology, each observation plays an equal role in determining the model parameters. However, the Generalized Least Squares (GLS) methodology allows for the processing of observations with varying weights, thus accommodating some form of heterogeneity. In this context, parameter estimation in GWR works by applying the potential of GLS to handle different weights to geographic data.

In traditional regression, parameters are determined based on the covariance matrix of the covariates. The parameter estimation formula is expressed as:

$$\beta = (X'X)^{-1}X'y \quad (26)$$

In contrast, in GWR, the parameter estimation incorporates a weight matrix W , which adjusts the influence of each observation based on its distance from the focal location. The adjusted formula for parameter estimation in GWR is:

$$\beta = (X'WX)^{-1}X'Wy \quad (27)$$

Here, W is a matrix where each element w_{ij} represents the weight assigned to the ij^{th} observation, reflecting its proximity to the focal point (Fotheringham et al., 2009).

For a specific geographic region g , the weights assigned to each observation are determined based on their distances from the focal location (of the region g). As this distance increases, the assigned weight decreases. This inverse relationship ensures that observations further from the focal point have a diminished influence on the parameter estimates for that region. The functional form of this relationship is typically specified using either an exponential or a Gaussian function, which dictates how rapidly the weights decay with distance. Exponential form is:

$$w_{ij} = e^{-\frac{d_{ij}^2}{\theta}} \quad (28)$$

where d_{ij} is the distance between observation i and j , and θ is the bandwidth parameter exogenous to the system that defines the size of the geographical region g . When latitudes and longitudes are used as the coordinate system locating observations to the geographical space the unit of bandwidth is kilometers (A. S. Fotheringham et al., 2009).

Bandwidth parameter controls every aspect of GWR (Comber et al., 2023). If a very high bandwidth is set, observations are treated as if they come from a single geographic region. Although observations influence parameter estimation through weighting, the parameters remain constant across the entire dataset. In this scenario, it is not appropriate to refer to the model as GWR; essentially, the system reverts to traditional regression. As the bandwidth decreases, the number of geographic regions increases, and unique parameters are calculated for each region. Consequently, GWR begins to diverge from traditional regression. However, as the bandwidth continues to decrease, the number of geographic regions increases proportionally, meaning fewer observations are available for parameter estimation within each region. This situation jeopardizes the robustness of GWR. Another issue is the potential loss of the model's ability to generalize the data, leading to the risk of overfitting (Fotheringham et al., 2009).

Charlton et al. (1998) introduced a cross-validation method designed to optimally determine the bandwidth value for models. As the method's name implies, the dataset is divided into training and test subsamples. Models are then constructed using only the training dataset, and the bandwidth values are iterated to identify the value that

performs best on the overall sample. This iterative process ensures that the chosen bandwidth maximizes the model's predictive accuracy and generalizability.

In a subsequent development, Fotheringham et al. (2009) devised a hypothesis test to compare the performance of a Geographically Weighted Regression (GWR) model with its Ordinary Least Squares (OLS) counterpart. According to this method, a GWR model that successfully passes the hypothesis test is considered superior to the OLS model. The selection of the optimal bandwidth among the models that pass this test is then contingent upon the specific context of the research, considering factors such as the geographical distribution of the data and the scale of spatial heterogeneity.



4. CHAPTER: FINDINGS & RESULTS

In this chapter, the outputs of the Exploratory Factor Analysis (EFA) investigating the compatibility of after-sales service data with SERVPERF dimensions are presented. The fit of the Confirmatory Factor Analysis (CFA) model, based on these outputs, to the data and the possibilities for modification are reported. Following these two steps, the results of the Analytical Hierarchy Process (AHP) analysis, which is constructed based on the obtained factor model, are presented. Outputs demonstrating the performance of the geographically weighted regression model, which is developed using the weights obtained from the AHP analysis, are provided. GWR models are compared to the OLS models. Additionally, the geographical distribution and statistical properties of the GWR Local models are reported.

4.1. Exploratory Factor Analysis Results

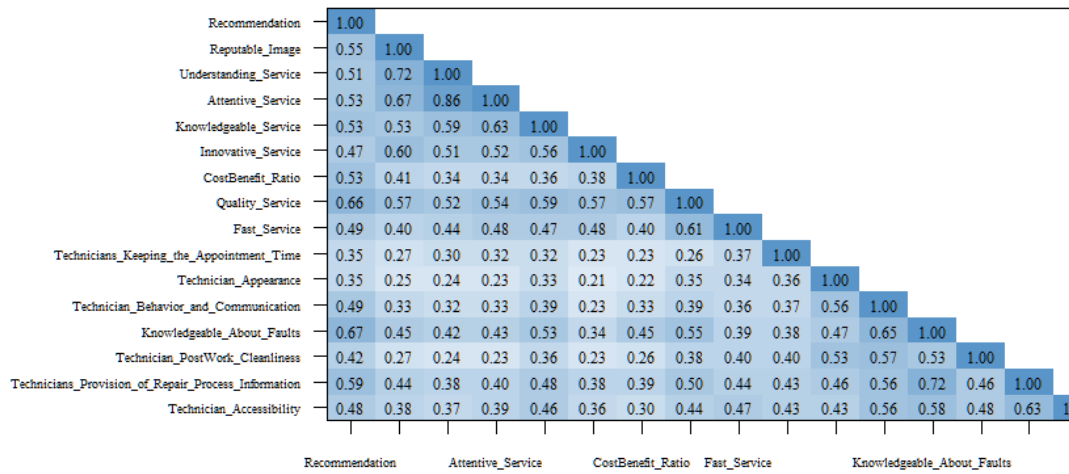
In this section, the EFA analyses conducted on the Repair/Maintenance Service dataset are reported in detail. On the other hand, the main outputs of the EFA analyses performed on the Installation/Setup Service dataset are shown here, with detailed tables provided in Appendix 2.

4.1.1. Data Characteristics

In this section, the findings obtained from the data are presented by following the perspective outlined in Sections 3.2.1, 3.2.2, and 3.2.3.

The building block of EFA is estimating the correlation matrix from data. Since all the questions included in the dataset are Likert-type questions, the normality and linearity assumptions required by the Pearson correlation matrix are not met. Therefore, polychoric correlations were used in calculating the correlation matrix for the EFA instead of the traditional approach. All pairwise correlations of the questions are positive, and it is noteworthy that polychoric correlation values are higher than Pearson correlation values when comparing the correlation matrices (Figure 4.1).

a.



b.

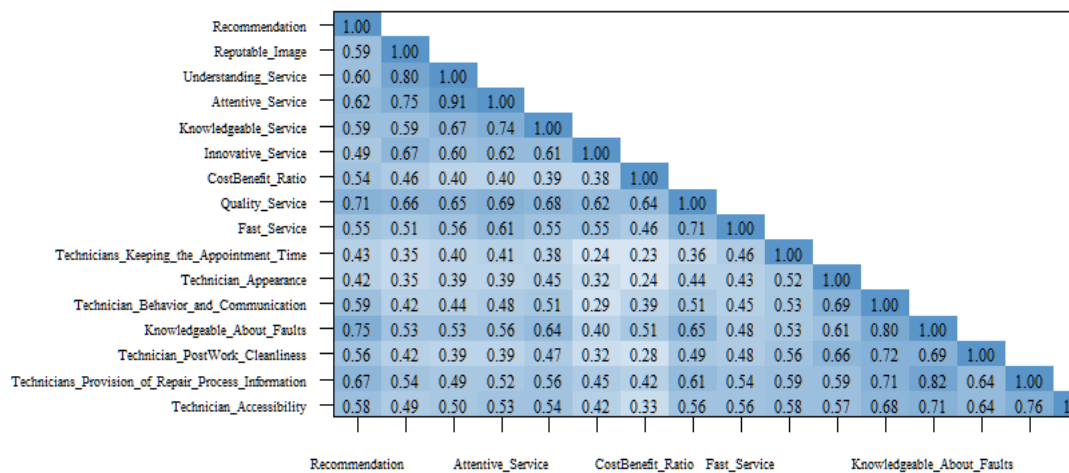


Figure 4.1. The Pearson correlation matrix (a) and polychoric correlation (b)

To assess the suitability of the data for factor analysis, in other words, to assess factorability of the data, the KMO test was conducted, yielding an overall measure for sampling adequacy of 92.5% (Table 4.1). Comparing this with the reference values, it is evident that the data are appropriate for factor analysis and possess adequate communality values. Table 4.1 also provides the sampling adequacy values for each measured variable. It is observed that the measure of sampling adequacy for each item is higher than the benchmark. Additionally, none of the items negatively impact the integrity of the data, as indicated by the absence of divergent sampling adequacy values.

Table 4.1. KMO values obtained from the Repair Dataset

Overall	92.5%
Recommendation	96.5%
Reputable Image	91.8%
Understanding Service	87.3%
Attentive Service	87.8%
Knowledgeable Service	94.7%
Innovative Service	92.2%
Cost Benefit Ratio	90.1%
Quality Service	93.3%
Fast Service	92.8%
Technicians Keeping the Appointment Time	94.4%
Technician Appearance	93.2%
Technician Behavior and Communication	92.6%
Knowledgeable About Faults	91.6%
Technician Post-Work Cleanliness	93.5%
Technicians' Provision of Repair Process Information	93.7%
Technician Accessibility	96.3%

In addition to the KMO test, the results of Bartlett's test of sphericity ($\chi^2(120) = 8666$, $p < .001$) indicate that the hypothesis that the data are only different from the identity matrix at the level of sampling error is rejected, based on the p-value.

4.1.2. Factor Extraction

Principal axis factoring has been chosen as the extraction method since it focuses on identifying the factors that account for the common variance in the data and is more robust to deviations from linearity and normality. As with all EFA algorithms, principal axis factoring demands a predetermined number of factors. To determine this, the eigenvalues resulting from the decomposition of the correlation matrix were compared with the results of parallel simulation, yielding a four-factor solution. In the scree plot, the eigenvalues obtained from the data coincide with the average eigenvalues from the simulation at the fourth step.

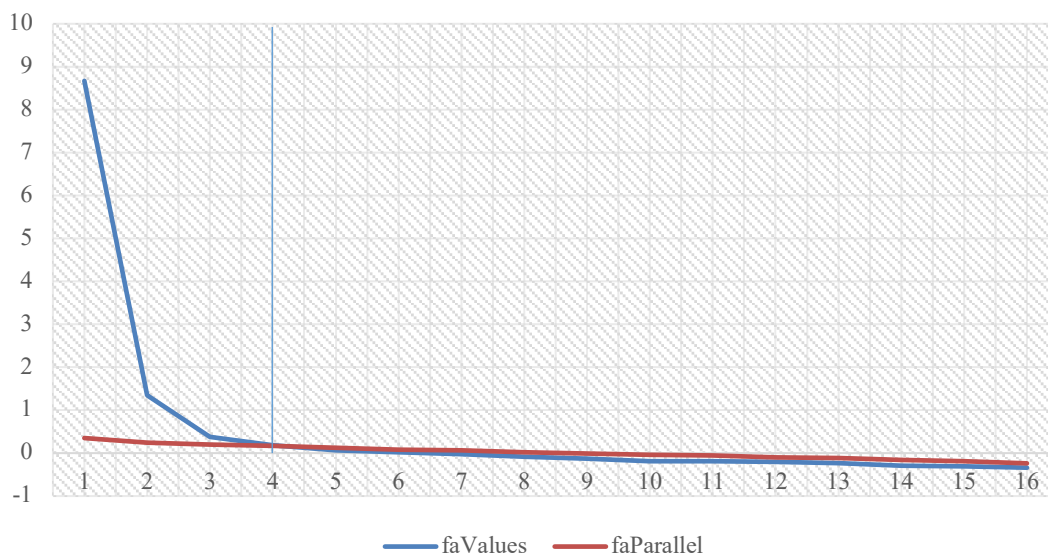


Figure 4.2. Screeplot of eigenvalues against parallel simulation

In the pre-rotation outputs table of the four-factor model there are six columns alongside the factors (Table 4.2). In the first column, the eigenvalue corresponding to each factor is listed. If PCA had been used as the factor extraction method, this value would represent the explanatory power of the first factor. Since the sum of all eigenvalues would equal the number of measured variables, which is 16, the explanatory power of the first factor in this case would be $9.103/16 = 56.9\%$. However, since principal axis factoring was used as the extraction method, the factor explanatory power is 4.844 as indicated in the column labeled Sum of Squares. Accordingly, the calculated explanatory power is 30.27%. This value is shown in the Proportion of Variance column because the factor alone can explain 30.27% of the variance in the data. In the adjacent Cumulative Proportion of Variance column, the proportion of

variance values are accumulated on a factor-by-factor basis. The difference between the total eigenvalue sums and the total Sum of Squares values for all factors is due to the principal axis factoring method's focus on explaining the common variance rather than the total variance in the data. If PCA had been used, the model's explanatory power would have been:

$$\frac{9.103 + 1.816 + 0.908 + 0.644}{16} = 77.94\%$$

According to the last row in the Cumulative Proportion of Variance column, the model as a whole can explain 70.85% of the variance in the data. The first factor explains 30.3% of the total variance and 42.7% of the common variance (shown in the Proportion Explained column). The second factor explains 21.9% of the total variance and 30.9% of the common variance. The third factor accounts for 16.1% of the total variance and 22.7% of the common variance. The explanatory power of the first three factors for both the general variance and common variance is quite high. Collectively, these three factors explain 68.28% of the total variance and 96.37% of the common variance (shown in the Cumulative Proportion Explained column), indicating that the fourth factor does not play a significant role in the model. This is clearly shown in the table. The fourth factor explains only 2.57% of the total variance and 3.63% of the common variance. It is important to note that the decision for a four-factor solution was guided by the parallel simulation output in the scree plot. The insignificance of the fourth factor is sufficient to redirect the analysis towards a three-factor solution. It is expected that a potential redistribution of loadings among factors may occur after rotation.

Table 4.2. Unrotated factor solution

	eig	SS	Var	Cum Var	Expl(%)	Cum. Exp (%)
<i>Factor 1</i>	9.103	4.844	30.27%	30.27%	42.73%	42.73%
<i>Factor 2</i>	1.816	3.504	21.90%	52.17%	30.91%	73.64%
<i>Factor 3</i>	0.908	2.577	16.11%	68.28%	22.73%	96.37%
<i>Factor 4</i>	0.644	0.412	2.57%	70.85%	3.63%	100.00%

When the varimax rotation is applied to the factor solution (Table 4.3), the situation regarding the fourth factor does not improve. After rotation, the fourth factor's explanatory power within the general framework is only 3.37%, and within the common variance, it is merely 4.76%. Consequently, transitioning from a four-factor solution to a three-factor solution becomes a necessity.

Table 4.3. 4-Factor solution after varimax rotation

	eig	Sum of Square	Var	Cum Var	Expl(%)	Cum. Exp (%)
<i>Factor 1</i>	9.103	4.656	29.10%	29.10%	41.08%	41.08%
<i>Factor 2</i>	1.816	3.950	24.69%	53.79%	34.85%	75.92%
<i>Factor 3</i>	0.908	2.190	13.69%	67.48%	19.32%	95.24%
<i>Factor 4</i>	0.644	0.539	3.37%	70.85%	4.76%	100.00%

The results for the three-factor solution after varimax rotation are shown below (Table 4.4). The three-factor model can explain 68.1% of the variance in the data. The distribution of loadings across the factors appears reasonable.

Table 4.4. 3-Factor solution after varimax rotation.

	eig	Sum of Square	Var	Cum Var	Expl(%)	Cum. Exp (%)
<i>Factor 1</i>	9.103	4.692	29.33%	29.33%	43.06%	43.06%
<i>Factor 2</i>	1.816	4.095	25.59%	54.92%	37.58%	80.64%
<i>Factor 3</i>	0.908	2.110	13.19%	68.10%	19.36%	100.00%

In the three-factor solution, the factor loadings of the measured variables after varimax rotation are shown below. In this section, the relationships between the measured variables and the factors are discussed purely from a technical perspective. A conceptual discussion will be addressed later. There are four columns to the right of the variables in Table 4.5. The first three columns show the degree of association of the measured variables with the factors. The Communality column indicates how much each variable is explained by the three factors in total. Accordingly, the sum of the squared loadings for Factor 1 is 4.692, which is equal to the sum of squares value for Factor 1 in Table 4.4.

The degree loading for each measured variable corresponds to the correlation between the variable and the factor (Hair et al., 2010). For example, the loading of the *Technical Appearance* question on Factor 1 is 0.72. Thus, the correlation between the question and the factor can be interpreted as 72%. It can be observed that the question also loads on other factors, albeit to a lesser extent. Loadings below 40% are printed in gray in the table. Looking at the loading values, there are more than three measured variables loading above 40% on each factor. This indicates that over-determination, which is critical for trusting the parameter estimations of the EFA, is achieved.

The Communality column represents the shared variance a measured variable has with the solution of EFA. For example, the sum of the squared loading values for the *Technical Appearance* question shows how much this question is explained by the three factors in total. According to the table, the variance of the *Technical Appearance* question is explained by the three-factor solution at 57.4%. Since we are discussing shared variance rather than regression, the 57.4% value can also be interpreted as the shared variance between the question and the three-factor solution (Field, 2000).

Finally, the Complexity column is located at the far right of the table. This column shows the proportion of each measured variable's loadings on the factors. Item complexity or more specifically Hofman's index of complexity (Hofmann, 1978) is calculated using,

$$\frac{(\sum_i \lambda_i^2)^2}{\sum_i \lambda_i^4} \quad (29)$$

where λ_i is the loading of an item on i^{th} factor. If an item loads solely on one factor, it equals 1 and it equals 2 if it loads exactly equally on two factors. If a variable loads highly on one factor and lower on the others, the Complexity value ranges between 1 and 2. However, if the largest loading values of a variable occur close to multiple factors, the Complexity value begins to exceed 2. This situation complicates the conceptual interpretation of the EFA results. In the table, the Complexity value of a total of 5 measured variables exceeds 2. However, except for the *Recommendation* question, which has the highest Complexity value, it is not difficult to make conceptual choices for the other questions, i.e., to decide under which factor they fall. However, further analysis will show that it is more beneficial to keep the *Recommendation* question within the model.

Table 4.5. Factor loadings table for 3-Factor solution

	F 1	F 2	F 3	Comm.	Cmplx.
<i>Recommendation</i>	0.472	0.422	0.521	0.672	2.916
<i>Reputable Image</i>	0.256	0.731	0.312	0.697	1.623
<i>Understanding Service</i>	0.268	0.884	0.149	0.876	1.243
<i>Attentive Service</i>	0.289	0.885	0.177	0.897	1.299
<i>Knowledgeable Service</i>	0.388	0.611	0.305	0.617	2.229
<i>Innovative Service</i>	0.164	0.630	0.336	0.537	1.684
<i>Cost-Benefit Ratio</i>	0.177	0.262	0.654	0.528	1.477

	F 1	F 2	F 3	Comm.	Cmplx.
<i>Quality Service</i>	0.321	0.542	0.667	0.842	2.403
<i>Fast Service</i>	0.379	0.486	0.378	0.522	2.821
<i>Technicians Keeping the Appointment Time</i>	0.632	0.244	0.059	0.462	1.311
<i>Technician Appearance</i>	0.724	0.206	0.092	0.574	1.195
<i>Technician Behavior and Communication</i>	0.812	0.188	0.242	0.753	1.290
<i>Knowledgeable About Faults</i>	0.743	0.278	0.437	0.820	1.939
<i>Technician Post-Work Cleanliness</i>	0.781	0.188	0.179	0.677	1.226
<i>Technicians' Provision of Repair Process Information</i>	0.732	0.295	0.357	0.750	1.809
<i>Technician Accessibility</i>	0.718	0.333	0.215	0.673	1.614

4.1.3. Installation/Setup Dataset EFA Results

Resultant factor structure elicited from Installation/Setup Dataset following the rigorous steps outlined in Appendix 2 is presented in the table below (Table 4.6).

Table 4.6. Factor loadings table for 3-Factor solution

	F1	F2	F3	Uniqueness
<i>Attentive Service</i>	0.849	0.212	0.206	0.191
<i>Understanding Service</i>	0.826	0.242	0.215	0.213
<i>Reputable Image</i>	0.734	0.139	0.281	0.363

	F1	F2	F3	Uniqueness
<i>Innovative Service</i>	0.714	0.165	0.127	0.447
<i>Knowledgeable Service</i>	0.652	0.306	0.200	0.442
<i>Fast Service</i>	0.620	0.251	0.171	0.524
<i>Opt for Service</i>	0.400	0.222	0.715	0.280
<i>Technician Accessibility</i>	0.370	0.639	0.287	0.373
<i>Technicians Keeping the Appointment Time</i>	0.328	0.478	0.178	0.632
<i>Technician Appearance</i>	0.269	0.658	0.063	0.491
<i>Recommendation</i>	0.250	0.292	0.816	0.186
<i>Quality Service</i>	0.241	0.390	0.724	0.266
<i>Technician Behavior and Communication</i>	0.205	0.775	0.362	0.226
<i>Knowledgeable About Product</i>	0.175	0.716	0.314	0.358
<i>Technician Post Work Cleanliness</i>	0.099	0.682	0.142	0.505

Note. Applied rotation method is varimax.

4.2. Confirmatory Factor Analysis Results

This section provides a comprehensive analysis of the Repair/Maintenance Service dataset undertaken by the CFA. Additionally, the primary results of the CFA analyses conducted on the Installation/Setup Service dataset are presented here, while more extensive tables may be found in Appendix 3.

4.2.1. Estimating CFA

Since the data is on an ordinal scale, a polychoric correlation matrix was calculated instead of the Pearson covariance matrix. Correspondingly, the unweighted least

squares (ULS) method was used instead of the Maximum Likelihood estimation method, which requires multivariate normality and pairwise linearity assumptions. The *lavaan* package in R can be utilized to use the ULS method with polychoric correlation for CFA. However, users without coding knowledge may prefer JASP, a completely free software.

The CFA model was established for a model with three constructs. This structure is based on the factor model resulting from the EFA discussed in the previous section.

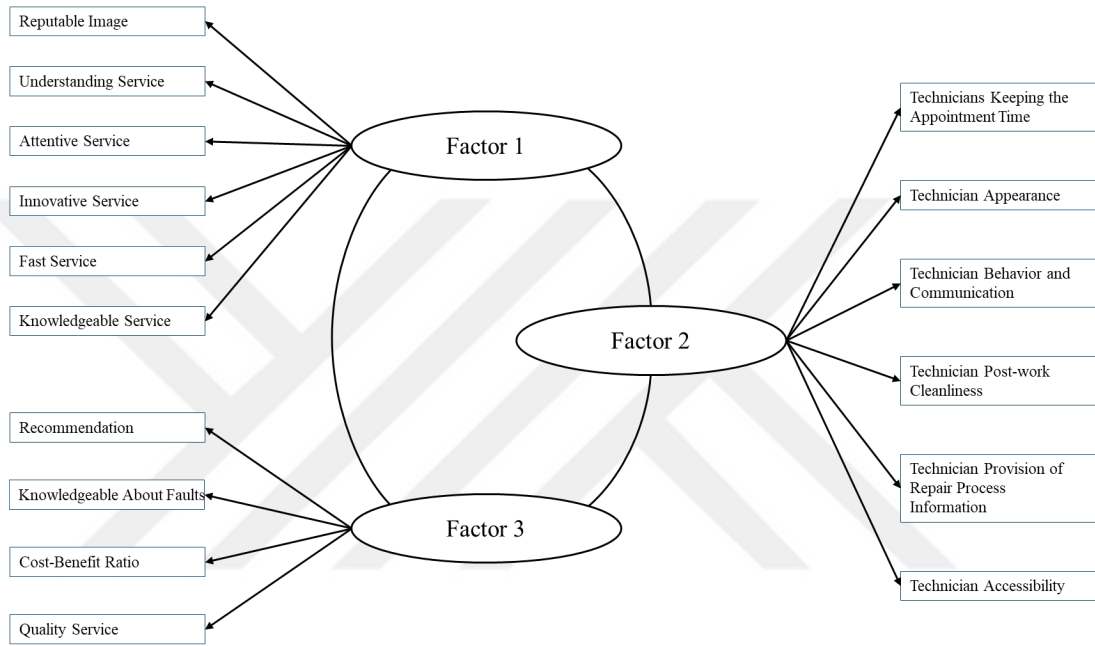


Figure 4.3. Initial CFA model

This structure imposes a baseline model with 120 degrees of freedom² to estimate 16 loadings and 3 covariances. After the estimation, 19 degrees of freedom are used, and we end up with the chi-square model with 101 degrees of freedom, whose fit values are displayed in Table 4.7.

² $120 = \frac{16 \times (16-1)}{2}$

Table 4.7. Chi-square test for the initial (16 measured variable) model.

Model	X²	df	p
Baseline model	16167	120	0
Factor model	232	101	0

4.2.2. Model Evaluation – Initial Model

The results based on the fit indices calculated from the comparisons with the baseline model are presented in Table 4.8. The CFI and TLI values are above 0.95, indicating an excellent fit between the model and the data according to these indices. The RMSEA value is below the cut-off of 0.06, which indicates a successful model. However, this positive indication is overshadowed by the 90% confidence intervals calculated for RMSEA, where the upper bound exceeds 0.06. Additionally, there is only a 25.7% probability that RMSEA is below 0.06 in sampling simulations. Finally, since the SRMR value does not fall below 0.06, the model's fit with the data according to this index is minimal.

Table 4.8. Fit indices for the initial (16 measured variable) model.

Index	Value
Comparative Fit Index (CFI)	0.992
Tucker-Lewis Index (TLI)	0.990
Root mean square error of approximation (RMSEA)	0.053
RMSEA 90% CI lower bound	0.044
RMSEA 90% CI upper bound	0.063
RMSEA p-value	0.257
Standardized root mean square residual (SRMR)	0.061

To accurately determine the model's fit with the data, all fit indices must remain on the correct side of the established cut-off values (Nye, 2023). Consequently, the model that uses all 16 measured variables does not adequately fit the data. Before proceeding with other aspects of the model evaluation, we shortcut to the modification process and then return to model evaluation.

4.2.3. Modifying the Measurement Model

1-Modified model:

In this section, the search for more successful modified models derived from the model containing 16 measured variables, which did not achieve the desired success in fit indices, was conducted. Accordingly, χ^2 values were calculated for various possible scenarios, and statistically significant options that have χ^2 test values above 3.84 are selected. Software filtered 20 such recommendations, and to save from space those with a chi-square test value above 10 (the first 9 out of 20 recommendations) are presented in Table 4.9.

Table 4.9. Modification indices for measured variables in the initial model

			Mod. Ind.
Factor 2	→	Knowledgeable About Faults	70.349
Factor 1	→	Knowledgeable About Faults	42.691
Factor 1	→	Quality Service	28.095
Factor 3	→	Fast Service	25.894
Factor 2	→	Quality Service	23.075
Factor 2	→	Fast Service	22.002
Factor 2	→	Cost-Benefit Ratio	17.318
Factor 3	→	Technicians Provision of Repair Process Information	12.118
Factor 2	→	Knowledgeable Service	11.008

It is crucial to take one step at a time in the procedures conducted according to the modification indices. For example, according to Table 4.9, it is suggested that the variable *Knowledgeable About Faults*, which is currently associated with Factor 3, be linked to Factor 2 as well, and the corresponding chi-square value is very high, indicating a high increase in the fit to data. The second option of the table also promises a considerable increase in the fit values, but we make only the first modification and connect *Knowledgeable About Faults* to Factor 2 (Figure 4.4). The reason for taking a single step at a time is shown below.

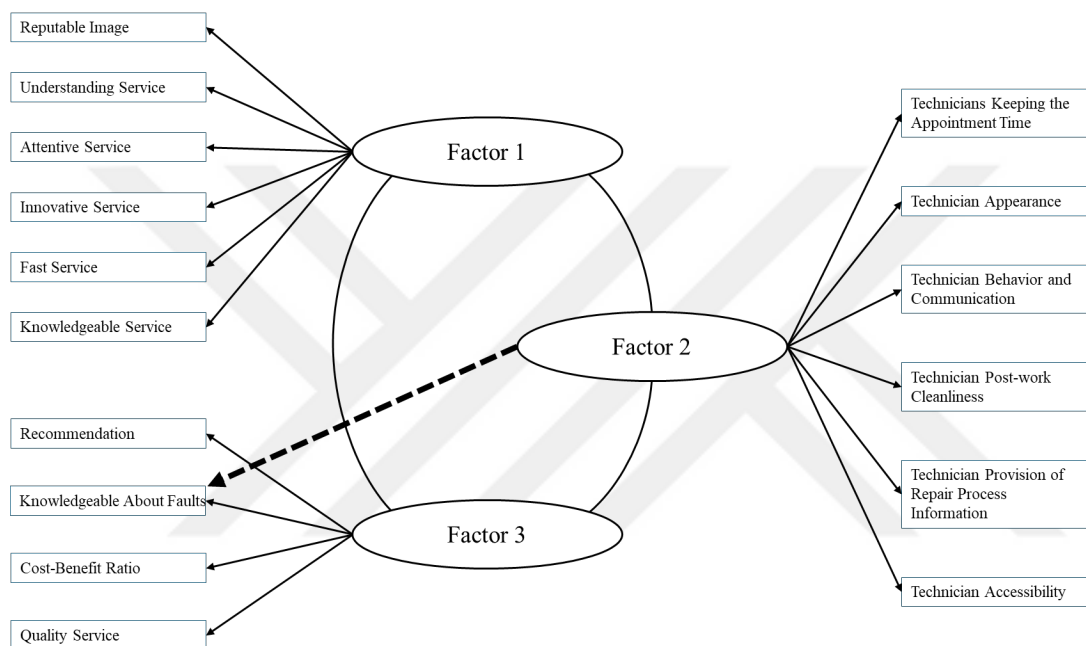


Figure 4.4. 1-Modified CFA model

The fit indices for the 1-Modified model are given in Table 4.10. When these values are compared to the fit indices of the initial model (Table 4.8), the improvement in the fit to the data obtained with the modification is clearly evident. The fit values of the 1-Modified model imply a perfect fit with the data, and in fact, the search for a better model should stop, and this model should be reported as the outcome of the CFA. However, the modification introduced a cross-loading to the initial model, which previously had no cross-loadings, and cross-loadings are undesirable for validity considerations. Therefore, instead of augmenting the suggested link (Factor 2 → *Knowledgeable About Faults*), we need to exclude the *Knowledgeable About Faults* variable entirely from the model.

Table 4.10. Fit indices for 1-Modified model.

Index	Value
CFI	0.996
TLI	0.995
RMSEA	0.038
RMSEA 90% CI lower bound	0.027
RMSEA 90% CI upper bound	0.048
RMSEA p-value	0.977
SRMR	0.052

Before removing the variable *Knowledgeable About Faults* from the model and reconducting the analysis from the start, we would like to present the modification indices table filtered to those with a chi-square test statistic value greater than 10 (Table 4.11). We observe that the top-ranked modifications from Table 4.9, particularly the second-ranked modification Factor 2 → *Knowledgeable About Faults*, are not among the first four ranks in Table 4.11. This demonstrates why modification steps should be implemented one at a time.

Table 4.11. Modification indices for measured variables in 1-Modified model

	Mod. Ind.
Factor 3 → Fast Service	22.444
Factor 2 → Fast Service	17.077
Factor 2 → Knowledgeable Service	14.363
Factor 2 → Recommendation	13.696

2-Modified model:

In this section, based on the information in Table 4.9, the performance of the 2-Modified model, resulting from the removal of the *Knowledgeable About Faults* variable due to the exclusion of the cross-loading option, was evaluated (Table 4.12). When the values in the table are considered as a whole, it is evident that a model with a very high fit to the data has been achieved. However, before concluding the search procedure, whether there is still a way to improve the model was examined.

Table 4.12. Fit indices for 2-Modified model

Index	Value
CFI	0.996
TLI	0.995
RMSEA	0.039
RMSEA 90% CI lower bound	0.028
RMSEA 90% CI upper bound	0.050
RMSEA p-value	0.954
SRMR	0.052

When examining the modification indices (Table 4.13) the *Fast Service* variable is suggested for cross-loading with high chi-square test values. Although the model's fit indices are currently at ideal values, the chi-square value greater than 20 in the modification table prompts us to remove the variable from the model and take another step forward.

Table 4.13. Modification indices for measured variables in 2-Modified model.

	Mod. Ind.
Factor 3 → Fast Service	23.814
Factor 2 → Fast Service	17.870

			Mod. Ind.
Factor 2	→	Recommendation	12.605
Factor 2	→	Knowledgeable Service	12.131
Factor 3	→	Technicians' Provision of Repair Process Information	11.287
Factor 1	→	Technicians' Provision of Repair Process Information	7.371
Factor 2	→	Innovative Service	7.098
Factor 1	→	Technician Accessibility	6.764
Factor 3	→	Understanding Service	4.823
Factor 3	→	Technician Accessibility	4.768
Factor 2	→	Understanding Service	4.768
Factor 2	→	Cost-Benefit Ratio	4.691
Factor 1	→	Technician Behavior and Communication	4.352
Factor 3	→	Innovative Service	4.261
Factor 2	→	Attentive Service	4.004

3-Modified model:

So far, starting from the initially postulated model (Figure 4.3), the Knowledgeable About Faults variable was first removed, followed by the Fast Service variable. At this point, the fit indices indicate that the model's fit to the data is highly satisfactory. However, the effort to achieve a more successful model with a more straightforward structure—pursuing the most parsimonious model—continues. After removing the Fast Service variable, the diagnostics of the model are presented in Table 4.14. Notably, there was significant improvement in RMSEA, which decreased from 0.039 to 0.031.

Table 4.14. Fit indices for 3-Modified model

Index	Value
CFI	0.997
TLI	0.996
RMSEA	0.031
RMSEA 90% CI lower bound	0.017
RMSEA 90% CI upper bound	0.044
RMSEA p-value	0.995
SRMR	0.047

When examining the modification indices table (Table 4.15), the only suggestion with a chi-square test statistic above 20 is the cross-loading proposal for the *Knowledgeable Service* variable. In light of the previous discussions, this variable has also been removed from the model.

Table 4.15. Modification indices for measured variables in 3-Modified model

	Mod. Ind.
Factor 2 → Knowledgeable Service	21.683
Factor 3 → Technicians' Provision of Repair Process Information	11.779
Factor 3 → Knowledgeable Service	11.395
Factor 2 → Recommendation	10.696
Factor 1 → Technicians' Provision of Repair Process Information	7.189
Factor 1 → Technician Accessibility	5.880
Factor 2 → Cost-Benefit Ratio	5.073
Factor 1 → Technician Post-Work Cleanliness	4.562

	Mod. Ind.
Factor 3 → Technician Accessibility	4.160

4-Modified model:

The final step in the modification process involves removing the Knowledgeable Service variable from the model. After this step, no further modification suggestions lead to a more successful outcome for the model. We refer to the model obtained after these four modifications as the final model and thoroughly examine its diagnostics.

4.2.4. Model Evaluation – Final Model

The final model structure imposes a baseline model with 13 measured variables. Consequently, there are 78 degrees of freedom to estimate 13 loadings and 3 covariances. After the estimation, 17 degrees of freedom are lost, and we end up with the χ^2 model with 62 degrees of freedom, whose fit values are displayed in Table 4.16. The right tail probability (p-value) of the factor model test statistic is 6.4%, greater than 5%. Hence, the Chi-square test does not reject the model. This is a rare phenomenon in literature since the Chi-square test almost always rejects the model. Researchers often refer to other fit indices by ignoring the Chi-square test results, claiming that this test is only applied as it provides a foundational analytical basis for forthcoming tests most widely used in the literature. However, the Final CFA model of this thesis manages to pass the intransigent chi-square test.

Table 4.16. Chi-square test for the initial (13 measured variable) model

Model	X ²	df	p
Baseline model	9973.211	78	0.000
Factor model	79.805	62	0.064

The fit values displayed in Table 4.17 are far beyond the necessary cut-off values. The CFI and TLI values are 99.8%. The RMSEA, which is considered a conservative index and sometimes relaxed up to 0.08 to indicate a working model, is 0.025 with a confidence interval of (0.000-0.040). Even the right-hand side of the confidence interval of RMSEA is better than the required 0.050 cut-off value. Thus, the software

calculates the probability that RMSEA would be below 0.050, considering the sampling error, as 99.9%. The SRMR value easily passes the required cut-off value as well.

Table 4.17. Fit indices for the Final model

Index	Value
CFI	0.998
TLI	0.998
RMSEA	0.025
RMSEA 90% CI lower bound	0.000
RMSEA 90% CI upper bound	0.040
RMSEA p-value	0.999
SRMR	0.044

The modification index in Table 4.18 does not recommend feasible modification alternatives anymore since following the top-ranked modification by removing the Recommendation variable is not admissible.

Table 4.18. Modification indices for measured variables in the final model

	Mod. Ind.
Factor 2 → Recommendation	11.855
Factor 3 → Technicians' Provision of Repair Process Information	11.600
Factor 2 → Cost-Benefit Ratio	7.263
Factor 1 → Technicians' Provision of Repair Process Information	6.712
Factor 1 → Technician Accessibility	5.499

	Mod. Ind.
Factor 1 → Technician Post-Work Cleanliness	4.779
Factor 3 → Technician Accessibility	3.908

For the sake of completeness, we also report measurement error residual covariances. The chi-square values are all low, and there is very little room left for model improvement since all diagnostics are nearly perfect.

Table 4.19. Residual covariance modification indices for measured variables

	Mod. Ind.
Cost-Benefit Ratio ↔ Quality Service	7.795
Quality Service ↔ Recommendation	7.323
Technicians' Provision of Repair Process Information ↔ Recommendation	5.634
Technician Appearance ↔ Technician Behavior and Communication	5.572
Understanding Service ↔ Attentive Service	4.980
Technicians Keeping the Appointment Time ↔ Quality Service	4.832
Technician Appearance ↔ Technician Post-Work Cleanliness	4.693
Technician Behavior and Communication ↔ Technician Post-Work Cleanliness	4.681
Reputable Image ↔ Innovative Service	4.022
Technician Appearance ↔ Cost-Benefit Ratio	3.946

4.2.5. Construct Validity Diagnostics

There are two pillars of construct validity: convergent validity and discriminant validity. The factor loadings table (Table 4.20) demonstrates that all loadings are statistically significant and that measured variables clustered under the same construct share a high proportion of their variances. Ten out of the thirteen-factor loadings are above 0.7, and according to the cut-offs discussed in Section 3.3.5.1, these values are excellent from a convergent validity perspective. Two-factor loadings are above 0.63, making them very good indicators of convergent validity. Finally, only one indicator has a loading of 0.545, which falls into the good indicator category of convergent validity.

Table 4.20. Factor loadings of the Final Model

Factor	Indicator	Estimate	Std. Error	Lower	Upper
Factor 1	Reputable Image	0.873	0.023	0.828	0.918
	Understanding Service	0.913	0.023	0.867	0.958
	Attentive Service	0.918	0.023	0.872	0.963
	Innovative Service	0.726	0.022	0.683	0.768
Factor 2	Technicians Keeping the Appointment Time	0.675	0.021	0.635	0.716
	Technician Appearance	0.698	0.021	0.657	0.738
	Technician Behavior and Communication	0.795	0.021	0.753	0.836
	Technician Post-Work Cleanliness	0.793	0.021	0.751	0.835
	Technicians Provision of Repair Process Inf.	0.887	0.022	0.844	0.929
	Technician Accessibility	0.839	0.022	0.796	0.881
Factor 3	Recommendation	0.842	0.026	0.791	0.893
	Cost-Benefit Ratio	0.545	0.022	0.503	0.588
	Quality Service	0.915	0.028	0.860	0.969

Each factor's average variance extracted value is expected to exceed 50%. According to Table 4.21, all factors have an AVE value well above 50%, supporting the convergent validity of the final model.

Table 4.21. Average Variance Extracted table of the Final Model

Factor	AVE
Factor 1	0.741
Factor 2	0.615
Factor 3	0.614

The AVE table is decomposed to demonstrate the R^2 values of each measured variable in Table 4.22. Except for the Cost-Benefit Ratio variable, all values are above 40%, indicating that the model explains a high proportion of the variance of each measured variable.

Table 4.22. R^2 values of each measured variable of the Final Model

	R^2
Reputable Image	0.762
Understanding Service	0.833
Attentive Service	0.842
Innovative Service	0.527
Technicians Keeping the Appointment Time	0.456
Technician Appearance	0.487
Technician Behavior and Communication	0.631
Technician Post-Work Cleanliness	0.629
Technicians Provision of Repair Process Info.	0.786
Technician Accessibility	0.703

	R²
Recommendation	0.709
Cost-Benefit Ratio	0.297
Quality Service	0.836

Below (Table 4.23) we report two final indicators of convergent validity. McDonald's omega, sometimes referred to as composite reliability or coefficient omega scores, are expected to be greater than 70%. For all three factors, this cut-off is exceeded, supporting convergent validity. Cronbach's alpha is rarely used since it requires tau-equivalence, it is reported for the sake of completeness. Even though the final CFA model developed in this thesis does not fulfill the tau-equivalence requirement, the scores indicate that the model's convergent validity is excellent.

Table 4.23. McDonald's omega and Cronbach alpha values of the Final Model

Factor	McDonald's Omega	Cronbach's Alpha
Factor 1	0.864	0.887
Factor 2	0.855	0.831
Factor 3	0.737	0.793

Next, we report our findings on the discriminant validity of the final model. First, we compare the final model with a model in which correlations between the constructs are set to one. The final model is expected to perform significantly better. For ease of comparison, the fit index values of the final model are placed beside the fit indices of the constrained model, where correlations between the constructs are set to one (Table 4.24). The result clearly indicates the final model's supremacy over the constrained model and supports discriminant validity.

Table 4.24. Fit index values of the Final Model and the Constrained Model

	Index	Final Model	Constrained Model
	CFI	0.998	0.888
	TLI	0.998	0.866
	RMSEA	0.025	0.192
	RMSEA 90% CI lower bound	0.000	0.182
	RMSEA 90% CI upper bound	0.040	0.202
	RMSEA p-value	0.999	0.000
	SRMR	0.044	0.167

The second approach to assessing discriminant validity involves comparing the Average Variance Extracted (AVE) values of two constructs with the square of the correlation between them. The logic behind this test is that a construct's ability to explain the shared variance of its associated measured variables should be higher than that of unrelated measured variables (hence, the correlation is squared). First, we report the correlations between the factors below.

Table 4.25. Factor correlation of the Final model.

					95% Conf. Int.	
					Lower	Upper
Factor 1	↔	Factor 2	0.627	0.017	0.593	0.661
Factor 1	↔	Factor 3	0.805	0.029	0.748	0.862
Factor 2	↔	Factor 3	0.725	0.025	0.675	0.774

Then, we borrow AVE values from Table 4.21 and compare them with the squared correlations in Table 4.25. In the first block of the table, the correlation between Factor 1 and Factor 2 is squared; 0.393. This value is smaller than the AVE values of both Factor 1 and Factor 2. In the second block of the table, the correlation between Factor

1 and Factor 3 is squared; 0.648. This value is smaller than the AVE value of Factor 1 but slightly larger than the AVE value of Factor 3. Finally, in the third block of the table, the correlation between Factor 2 and Factor 3 is squared; 0.526. This value is smaller than the AVE values of both Factor 2 and Factor 3.

Table 4.26. Comparison of AVE values and squared correlations of Factors

	AVE	Square of Factor Correlations
Factor 1	0.741	0.393
Factor 2	0.615	
Factor 1	0.741	0.648
Factor 3	0.614	
Factor 2	0.615	0.526
Factor 3	0.614	

Finally, discriminant validity necessitates that each indicator loads exclusively onto a single construct. Unidimensionality requires the absence of cross-loaded indicators, which is essential for attributing meaning to a latent construct (Gerbing & Anderson, 1988). In the Final Model, each measured variable loads onto a single factor. Hence, the Final model supports both unidimensionality and discriminant validity.

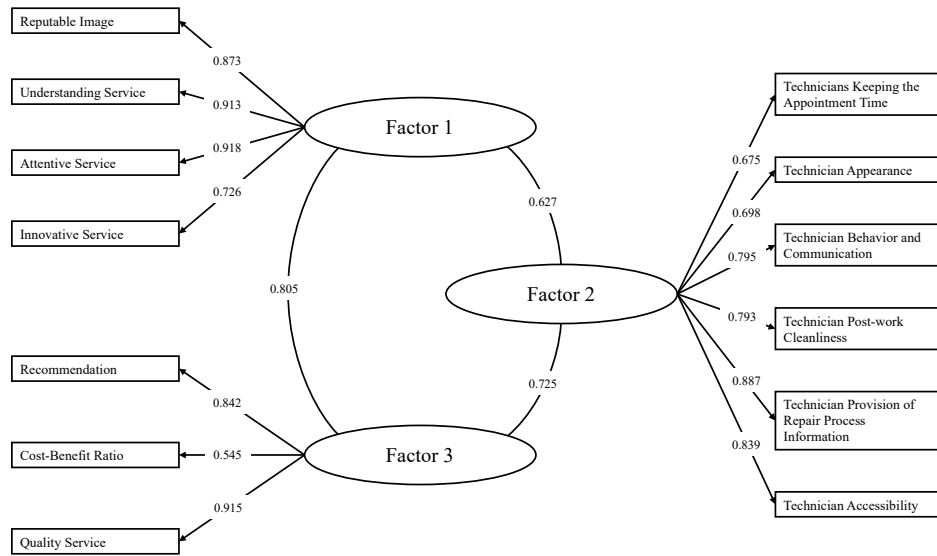


Figure 4.5. Final CFA model for Repair/Maintenance

4.2.6. Final CFA Model for Installation/Setup Dataset

While a final CFA model is created for the Installation/Setup Service, the same steps are followed in developing the Repair/Maintenance Service model. Detailed tables showing the steps of all analyses are provided in Appendix 3.

Each measured variable is loaded into a single factor in the Final Model of Installation/Setup Service. Therefore, the Final model demonstrates the quality of having a single dimension and the ability to distinguish between different variables. Demonstration of the model and the loadings of each item are presented in Figure 4.6. The discussion on the model's construct validity is in Appendix 3.

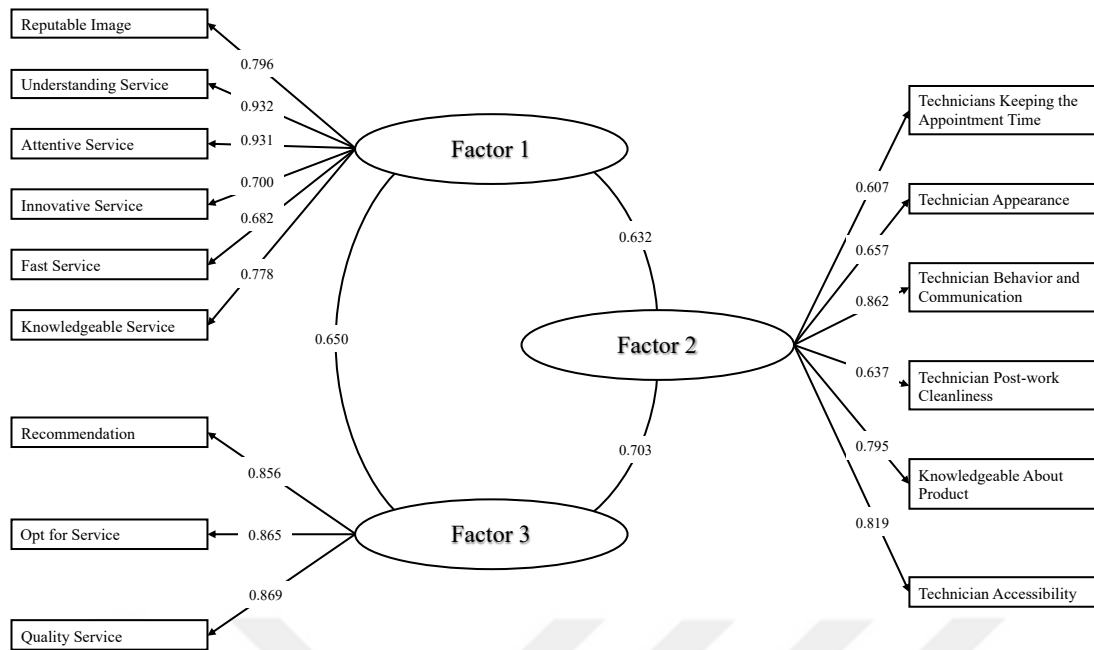


Figure 4.6. Final CFA model for Installation/Setup

4.3. Interpretation of the Factors

This section examines the correspondence of factor models created after EFA and CFA analyses and passes the construct validity tests with SERVPERF (SERVQUAL) dimensions.

First, the correspondences of the factors derived from the Repair/Maintenance Service Quality dataset with the SERVPERF dimensions were identified (Table 4.27).

According to this, the Responsiveness, Tangible, and Empathy dimensions merged into a single dimension, while the Reliability and Assurance dimensions remained separate. Three items are dropped from the factor models. Hence, their SERVPERF dimension equivalences have not been reported.

Table 4.27. Repair/Maintenance Service Quality Dimension Correspondence

Repair/Maintenance Service Quality Item	Service Quality Dimension	CFA Factors
<i>Recommendation</i>	Reliability	Factor 3

Repair/Maintenance Service Quality Item	Service Quality Dimension	CFA Factors
<i>Reputable Image</i>	Assurance	Factor 1
<i>Understanding Service</i>	Assurance	Factor 1
<i>Attentive Service</i>	Assurance	Factor 1
<i>Knowledgeable Service</i>	Dropped	
<i>Innovative Service</i>	Assurance	Factor 1
<i>Cost-Benefit Ratio</i>	Reliability	Factor 3
<i>Quality Service</i>	Reliability	Factor 3
<i>Fast Service</i>	Dropped	
<i>Technicians Keeping the Appointment Time</i>	Responsiveness	Factor 2
<i>Technician Appearance</i>	Tangibles	Factor 2
<i>Technician Behavior and Communication</i>	Empathy	Factor 2
<i>Knowledgeable About Faults</i>	Dropped	
<i>Technician Post-Work Cleanliness</i>	Tangibles	Factor 2
<i>Technicians' Provision of Repair Process Information</i>	Responsiveness	Factor 2
<i>Technician Accessibility</i>	Empathy	Factor 2

Second, the correspondences of the factors derived from the Installation/Setup Service Quality dataset with the SERVPERF dimensions were identified (Table 4.28). Similar to the Repair/Maintenance Service dimensions, the Responsiveness, Tangible, and

Empathy dimensions merged into a single dimension, while the Reliability and Assurance dimensions remained separate.

Table 4.28. Installation/Setup Service Quality Dimension Correspondence

Installation/Setup Service Quality Item	Service Quality Dimension	CFA Factors
<i>Recommendation</i>	Reliability	Factor 3
<i>Reputable Image</i>	Assurance	Factor 1
<i>Understanding Service</i>	Assurance	Factor 1
<i>Attentive Service</i>	Assurance	Factor 1
<i>Knowledgeable Service</i>	Assurance	Factor 1
<i>Innovative Service</i>	Assurance	Factor 1
<i>Opt for Service</i>	Reliability	Factor 3
<i>Quality Service</i>	Reliability	Factor 3
<i>Fast Service</i>	Assurance	Factor 1
<i>Technicians Keeping the Appointment Time</i>	Responsiveness	Factor 2
<i>Technician Appearance</i>	Tangibles	Factor 2
<i>Technician Behavior and Communication</i>	Empathy	Factor 2
<i>Knowledge About Product</i>	Responsiveness	Factor 2
<i>Technician Post-Work Cleanliness</i>	Tangibles	Factor 2
<i>Technician Accessibility</i>	Empathy	Factor 2

4.4. Analytic Hierarchy Process (AHP) Results

The AHP method was used in that phase of the study, and the results obtained for both the Installation/Setup Service and Repair /Maintenance Service datasets are given here. The purpose of using the AHP method here is to determine the priority of the dimensions relative to each other and their relative weights, as well as to calculate the dimension scores for the GWR performed in the last phase of the study.

4.4.1. Repair /Maintenance Service AHP Results

First, the final model obtained after CFA was defined in a hierarchical structure. The defined hierarchical structure, with overall service quality satisfaction at the top, service quality dimensions obtained from the survey at the lower level, and survey questions forming the dimensions at the bottom level, is given below in Figure 4.7.

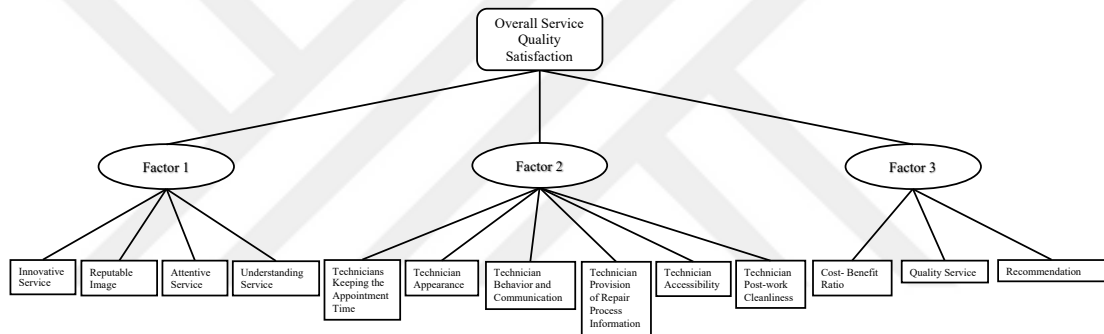


Figure 4.7. The hierarchical structure designed for Repair /Maintenance Service

The hierarchical structure designed for the AHP application is shared with five company managers who are experts in evaluating the quality of the service offered, managing customer relations, and performing evaluation of service personnel. A comparison matrix is created for the three dimensions determined at the first level and experts are asked to fill in this matrix. Additionally, definitions explaining the service characteristics included in each defined dimension are provided with the comparison matrix. The comparisons are made based on how important each criterion is compared to the other, considering the objective, that is, service quality satisfaction.

The consistency of the comparison matrices created in this way was evaluated, and it was found that the CR value for each matrix was less than 10%, which indicates consistent evaluations of the items.

Since a single judgment was desired based on pairwise comparisons, the geometric averages of the answers given to each question were taken (Ishizaka & Labib, 2011). The final comparison matrix obtained using geometric means is given in Table 4.29.

Table 4.29. Pairwise comparisons for the first level of the hierarchical concept

	Factor 1	Factor 2	Factor 3
Factor 1	1.00	3.93	1.44
Factor 2	0.25	1.00	0.40
Factor 3	0.69	2.50	1.00

After the comparison matrix given in Table 5.26 is normalized, the principal eigenvalue λ_{max} and w vector, which gives the weights of the dimensions, are calculated. The weight vector of w is $w = (0.517, 0.135, 0.348)$ with the principal eigenvalue $\lambda_{max} = 3.00083$

From Table 3.6 (Randomness Indicators), the RI value for $n = 3$ is 0.52, and the CI and CR values are obtained as follows.

$$CI = \frac{3.00083 - 3}{2} = 0.000415$$

$$CR = \frac{0.000415}{0.52} = 0.00079$$

The CR obtained as 0.00079 is less than 0.10, and this result shows consistency in pairwise comparisons.

According to the results obtained, the priority order of the dimensions and the corresponding weights for each dimension are given in the table below (

Table 4.30).

Table 4.30. Priority Weights

Dimension	Priority Weight
Factor 1	%51.7

Dimension	Priority Weight
Factor 3	%34.8
Factor 2	%13.5

At the lowest level, there are different numbers of items under each factor dimension. Considering that the number of items here is high, especially in Factor 2, and that pairwise comparisons in a two-layer hierarchical structure will be challenging for experts and will potentially lead to inconsistency problems, experts were asked to evaluate the items at the lower level with the point allocation method. Accordingly, experts distributed 100 points to the items under each factor according to their degree of importance to the factor to which it connects. The percentages distributed by the experts are converted into a comparison matrix, assuming that it is a continuous scale rather than a 1-9 discrete scale. In this way, the problem of inconsistent comparisons is eliminated. Finally, to calculate the pairwise comparison matrix, a single comparison value for each comparison was obtained by taking the geometric mean of the values given by the experts.

The final comparison matrix calculated for Factor 1 is given in Table 4.31.

Table 4.31. Pairwise Comparisons for the Items under Factor 1

Pair-wise Matrix	Reputable Image	Understanding Service	Attentive Service	Innovative Service
Reputable Image	1.00	0.88	0.87	0.80
Understanding Service	1.13	1.00	0.99	0.90
Attentive Service	1.15	1.01	1.00	0.91
Innovative Service	1.26	1.11	1.09	1.00

After the comparison matrix is normalized, w vector, which gives the weights of the dimensions, are calculated. The weight vector of w is $w = (0.220, 0.250, 0.253, 0.277)$.

The final comparison matrix calculated for Factor 2 is given in Table 4.32.

Table 4.32. Pairwise Comparisons for the Items under Factor 2

Pair-wise Matrix	Technician Appearance	Technician Behavior and Communication	Technician Provision of Repair Process Information	Technician Post-work Cleanliness	Technician Accessibility	Appointment Time
Technician Appearance	1.00	0.57	0.33	0.70	0.76	0.35
Technician Behavior and Communication	1.74	1.00	0.57	1.22	1.32	0.62
Technician Provision of Repair Process Information	3.05	1.75	1.00	2.13	2.31	1.08
Technician Post-work Cleanliness	1.43	0.82	0.47	1.00	1.08	0.51
Technician Accessibility	1.32	0.76	0.43	0.92	1.00	0.47
Technicians Keeping the Appointment Time	2.83	1.62	0.93	1.97	2.14	1.00

After the comparison matrix is normalized, w vector, which gives the weights of the dimensions, are calculated. The weight vector of w is $w = (0.088, 0.153, 0.268, 0.126, 0.116, 0.249)$.

The final comparison matrix calculated for Factor 3 is given in Table 4.33.

Table 4.33. Pairwise Comparisons for the Items under Factor 3

Pair-wise Matrix	Recommendation	Cost-Benefit Ratio	Quality Service
Recommendation	1.00	0.41	0.89
Cost- Benefit Ratio	2.42	1.00	2.17
Quality Service	1.12	0.46	1.00

After the comparison matrix is normalized, w vector, which gives the weights of the dimensions, are calculated. The weight vector of w is $w = (0.220, 0.553, 0.246)$.

According to the results obtained, the priority order of the sub-factors and the corresponding weights for each factor are given in the tables (Table 4.34, Table 4.35, Table 4.36)

Table 4.34. Priority weights for Items under Factor 1

Sub-factor	Priority Weight
Innovative Service	27.7%
Attentive Service	25.3%
Understanding Service	25.0%
Reputable Image	22.0%

Table 4.35. Priority weights for Items under Factor 2

Sub-Factor	Priority Weight
Technician Provision of Repair Process Information	26.8%
Appointment Time	24.9%

Technician Behavior and Communication	15.3%
Technician Post-work Cleanliness	12.6%
Technician Accessibility	11.6%
Technician Appearance	8.8%

Table 4.36. Priority weights for Items under Factor 3

Sub-Factor	Priority Weight
Cost- Benefit Ratio	53.3%
Quality Service	24.6%
Recommendation	22.0%

4.4.2. Installation/Setup Service AHP Results

First, the final model obtained after CFA was defined in a hierarchical structure. The defined hierarchical structure, with overall service quality satisfaction at the top, service quality dimensions obtained from the survey at the lower level, and survey questions forming the dimensions at the bottom level, is given below in Figure 4.8.



Figure 4.8. The hierarchical structure designed for Installation /Setup Service
The final comparison matrix obtained using geometric means is given in Table 4.37.

Table 4.37. Pairwise comparisons for the first level of the hierarchical concept

	Factor 1	Factor 2	Factor 3
Factor 1	1.00	2.72	1.00
Factor 2	0.37	1.00	0.41
Factor 3	1.00	2.46	1.00

The weight vector of w is $w = (0.426, 0.162, 0.412)$ with the principal eigenvalue $\lambda_{max} = 3.00012$

From Table 3.6 (Randomness Indicators), the RI value for $n = 3$ is 0.52, and the CI and CR values are obtained as follows.

$$CI = \frac{3.00012 - 3}{2} = 0.00006$$

$$CR = \frac{0.00006}{0.52} = 0.00011$$

The CR obtained as 0.00011 is less than 0.10, and this result shows consistency in pairwise comparisons.

According to the results obtained, the priority order of the dimensions and the corresponding weights for each dimension are given in the table below (Table 4.38).

Table 4.38. Priority Weights

Dimension	Priority Weight
Factor 1	42.6%
Factor 3	41.2%
Factor 2	16.2%

Final comparison matrices of subfactors for Installation/Setup are given in Appendix 4. According to the results obtained, the priority order of the sub-factors and the corresponding weights for each factor are given in the tables (Table 4.39, Table 4.40, Table 4.41).

Table 4.39. Priority weights for Items under Factor 1

Sub-factor	Priority Weight
Knowledgeable Service	22.3%
Fast Service	19.2%
Attentive Service	17.7%
Reputable Image	15.4%
Understanding Service	13.7%
Innovative Service	11.7%

Table 4.40. Priority weights for Items under Factor 2

Sub-Factor	Priority Weight
Technician Knowledge about Product	21.1%
Appointment Time	20.4%
Technician Appearance	16.2%
Technician Behavior and Communication	15.3%
Technician Post-work Cleanliness	14.4%
Technician Accessibility	12.6%

Table 4.41. Priority weights for Items under Factor 3

Sub-Factor	Priority Weight
Recommendation	40.6%
Opt for Service	30.8%
Quality Service	28.6%

4.5. Geographically Weighted Regression (GWR) Results

After determining the dimensional structure based on the factor analyses conducted above, weights were assigned to these dimensions using the Analytic Hierarchy Process (AHP) based on expert opinions. Subsequently, scores for three SERVQUAL dimensions, AHP Factor1, AHP Factor2, and AHP Factor3, were calculated. In the next phase of the study, Ordinary Least Squares (OLS) and Geographically Weighted Regression (GWR) models were developed to explain the responses to the Overall Satisfaction question in the survey using these three factors.

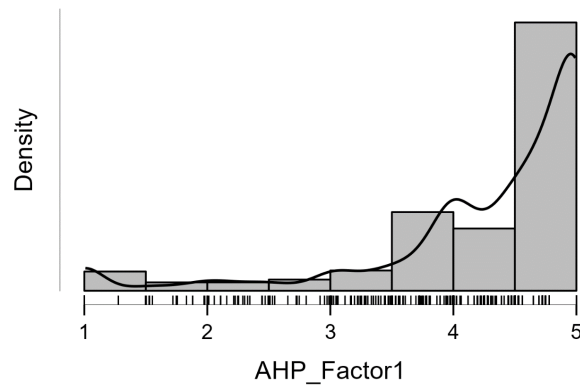
Before constructing the models, the descriptive statistics of the factor scores were examined. Since participants mostly preferred the highest ratings for the Likert-type questions in the survey, the factor scores were calculated to be relatively high. The standard deviations of the factors were like each other. The skewness index indicated that the distribution of all three factors was not symmetric but left-skewed. The ratio of the skewness values to their standard errors, given in the following row, was greater than 2, indicating that the skewness values significantly differed from zero. The kurtosis values for the first two factors were high and statistically significant, while the kurtosis index for the third factor was not significantly different from zero. Lastly, the results of the Shapiro-Wilk normality test are presented in the table. The null hypothesis of this test posits that the data come from a normally distributed population. Due to the p-values in the subsequent row, the null hypothesis is rejected.

Table 4.42. Descriptive statistics of factor scores

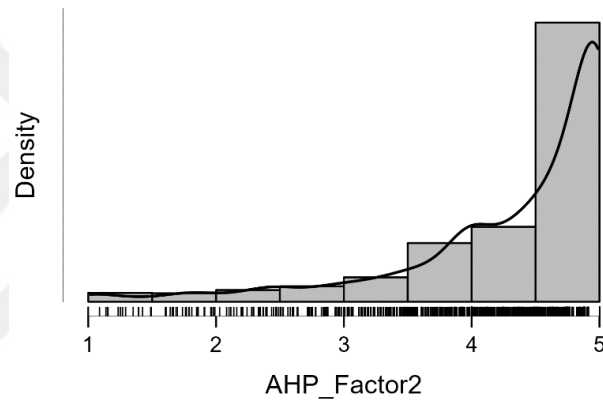
	AHP Factor1	AHP Factor2	AHP Factor3
Valid	1363	1363	1363
Missing	0	0	0
Median	4.695	4.67	4
Mean	4.269	4.328	3.698
Std. Deviation	0.975	0.872	1.11
Skewness	-1.923	-1.765	-1.052
Std. Error of Skewness	0.066	0.066	0.066
Kurtosis	3.372	2.905	0.239
Std. Error of Kurtosis	0.132	0.132	0.132
Shapiro-Wilk	0.737	0.769	0.881
P-value of Shapiro-Wilk	< .001	< .001	< .001

The distribution of the factor scores allows for a clearer understanding of the results presented in the table above. In the following figure, the histogram of each factor score is displayed in separate frames. Over each histogram, a distribution curve is superimposed, and rug marks are placed along the x-axis to indicate the locations of individual observations.

a.



b.



c.

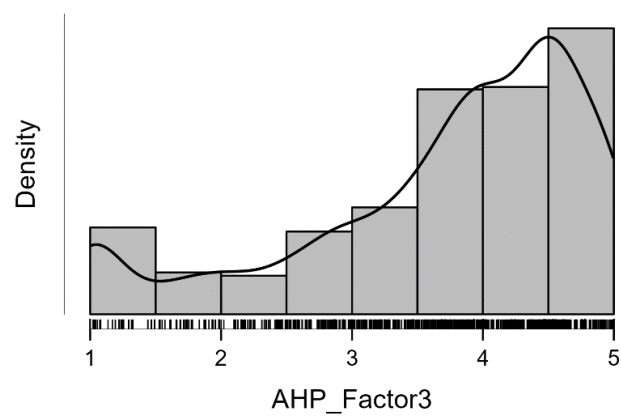


Figure 4.9. Histograms, approximate distributions and rug marks of AHP Factor1 (a), AHP Factor2 (b), and AHP_Factor3(c)

It is evident that the factor scores, given their distributions, do not meet the normality assumption required for OLS. However, various transformation attempts on the factor scores, such as those suggested by Box and Cox (1964), did not yield valuable results. When normality is not achieved in OLS, the residuals are not normally distributed, leading to biased estimates of the standard errors of the estimations. Nevertheless, as can be seen from the OLS tables, the calculated values for the model parameters are robust and significantly distant from any critical thresholds.

In the models where the factor scores served as covariates, the dependent variable used was Overall Satisfaction, which was assessed using a Likert scale item. Consistent with the survey's overall trend, the most frequently chosen value (46.6%) was the highest option.

Table 4.43. Summary Statistics and Frequency Table

Descriptive Statistics		Overall Satisfaction	Frequency	%	Cumulative %
Valid	1363	1	156	11.445	11.445
Missing	0	2	45	3.302	14.747
Mode	5	3	120	8.804	23.551
Median	4	4	407	29.861	53.412
Mean	3.968	5	635	46.588	100
Std. Deviation	1.31				

As can be seen from the bar chart, the dependent variable also has a left-skewed distribution. Additionally, since it was not measured on a continuous scale, directly applying OLS could lead to unreliable results. In situations where the dependent variable has an ordinal scale, then ordinal logistic regression is a preferable method (Harrell, 2015).

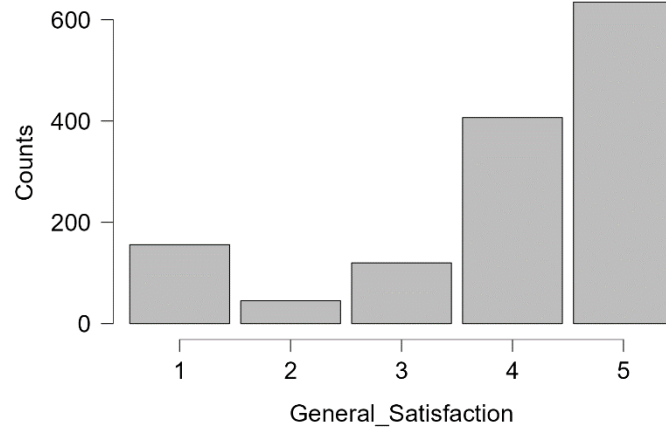


Figure 4.10. Frequency Distribution of Overall Satisfaction

On the other hand, although the Overall Satisfaction variable was measured on an ordinal scale, it should not be overlooked that the values between the options are conceptually meaningful. In typical ordinal variables, the values between options are usually meaningless (e.g., educational level, military rank). However, for the Overall Satisfaction question, intermediate values are meaningful. Therefore, the statement that the mean is approximately equal to the median in the table above makes sense. This perspective motivates OLS analysis.

4.5.1. Ordinary Least Squares for Repair/Maintenance Service

Every GWR report necessarily involves an OLS section. The analysis starts with the traditional OLS definition form:

$$GS = \beta_0 + \beta_1 \times F_1 + \beta_2 \times F_2 + \beta_3 \times F_3 + \varepsilon \quad (30)$$

where GS is the Overall Satisfaction variable that acts as the dependent variable in OLS, F_i represents AHP_Factor $_i$ covariate, β_i is the coefficient of i^{th} AHP_Factor covariate, and finally ε is the error of OLS model.

ANOVA table shows that the overall OLS model with factor scores as covariates and Overall Satisfaction as the dependent variable is meaningful as the p-value is less than 1%. From the ANOVA table we deduce that $R^2 = 0.597$ and $Adjusted R^2 = 0.596$. Thus, the OLS model explains almost 60% of the variation in the dependent variable. The AIC of the model is 3375.

Table 4.44. ANOVA table of the OLS model

	Sum of Squares	df	Mean Square	F-stat	p-value
Regression	1395.07	3	465.02	670.47	< 2.2e-16
Residual	942.58	1359	0.69		
Total	2337.64	1362			

Regarding the model coefficients and collinearity diagnostics table, it can be observed that all the parameters have t-statistics with absolute values greater than 3, indicating that they are statistically significantly different from zero at the 1% level ($p < 0.01$). The table shows that the coefficient of the intercept term is negative. This implies that the Overall Satisfaction calculated from factor scores overestimates the actual Overall Satisfaction in the data. Considering that all Likert-type variables have values ranging from 1 to 5, the factor scores representing the experts' estimates overestimate Overall Satisfaction by $0.69/5 = 13.8\%$ compared to the customers' assessments. The VIF scores suggest that multicollinearity is not a concern for the factor scores.

Table 4.45. Model coefficients and collinearity diagnostics table.

	Coef	Std.Error	t	Lower	Upper	Tolerance	VIF
(Intercept)	-0.69	0.12	-5.56	-0.94	-0.45		
AHP_Factor1	0.23	0.03	7.29	0.17	0.29	0.55	1.84
AHP_Factor2	0.37	0.03	11.47	0.30	0.43	0.65	1.53
AHP_Factor3	0.57	0.03	20.05	0.51	0.62	0.52	1.94

Examining the coefficients of the factor scores reveals that Factor 3 has the highest impact on Overall Satisfaction. Improvements made solely in this factor affect Overall Satisfaction approximately equal to the combined effect of the other two factors.

A Q-Q plot of the standardized residuals shows that most of the data points fall along the reference (red) line, suggesting that the data follows a normal distribution

reasonably well. However, there are some deviations at both the lower and upper tails. The points at the extreme left deviate from the line, indicating that the data may have heavier tails than a normal distribution.

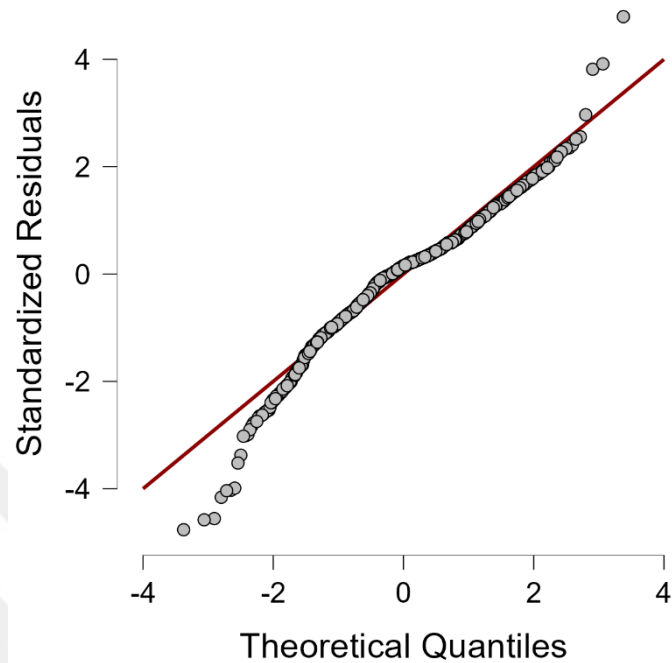


Figure 4.11. Q-Q Plot Standardized Residuals

4.5.2. Geographically Weighted Regression Repair/Maintenance Service

A GWR analysis typically begins with an optimal bandwidth study. Since geographic coordinates in the data are represented by latitude and longitude information, the bandwidth unit is kilometers. Because the survey sample size is selected relative to the population of a geographically defined area (city), the data do not have a uniform distribution in spatial space. Therefore, even through visual inspection, it is evident that a fixed bandwidth parameter would not yield successful results. The number of observations is higher in densely populated coastal areas compared to regions with lower population density.

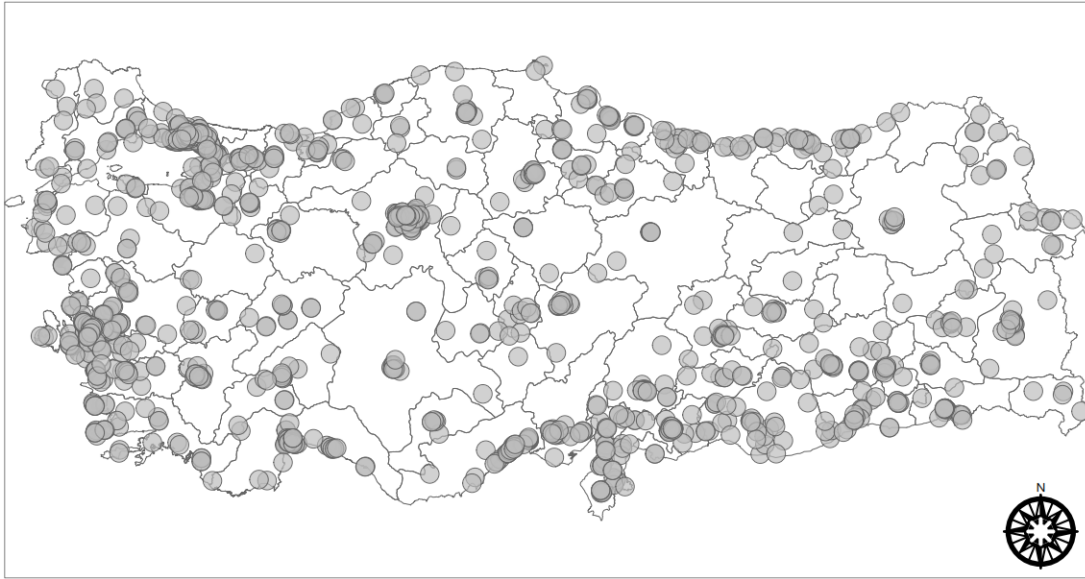


Figure 4.12. Geographic distribution of observations.

In cases where data are not uniformly distributed across spatial space, the use of adaptive bandwidth is recommended (Bidanset & Lombard, 2014). While optimal fixed bandwidth attempts to maximize the use of spatial information in the data through the weights assigned to observations, it remains a constant value across the entire map. On the other hand, optimal adaptive bandwidth achieves the same goal by determining an optimal number of neighboring observations around the focal point. Adaptive bandwidth allows for different bandwidth values to be used throughout the map. For instance, in the map's interior regions, where the number of observations is relatively low, the adaptive bandwidth value will be higher than in larger cities and coastal areas. The distribution of these values is presented below.

The decision regarding adaptive bandwidth revolves around determining the minimum number of neighboring observations to be used. In this thesis, various trials were conducted. The adaptive bandwidth value is expected to be set as a proportion of the total number of observations. Trials indicated that as the determined proportion decreased, the R^2 value of the GWR model increased. This was an anticipated outcome, and similar to determining the number of factors in EFA, a criterion was sought due to the absence of a context-independent method for shortcutting the solution.

Based on the conducted trials, a criterion was established to maximize R^2 while controlling the standard errors of the calculated parameters as much as possible.

Accordingly, two degrees of freedom were allocated for each covariate on a focal point basis. Since the number of covariates is 3, the adaptive bandwidth was determined as $p = 6/1363 = 0.0044$. Since the model is restricted to calculating parameters with at least six observations, the geographical bandwidth varied substantially. This variation is summarized in the six-point summary table below and the distribution chart beneath the table. Bandwidth values vary between 1.5 km to 165 km.

Table 4.46. Six-point summary table of the bandwidth values.

Min	1 st Quad	Median	Mean	3 rd Quad	Max
1.510	3.849	8.872	20.541	33.399	164.338

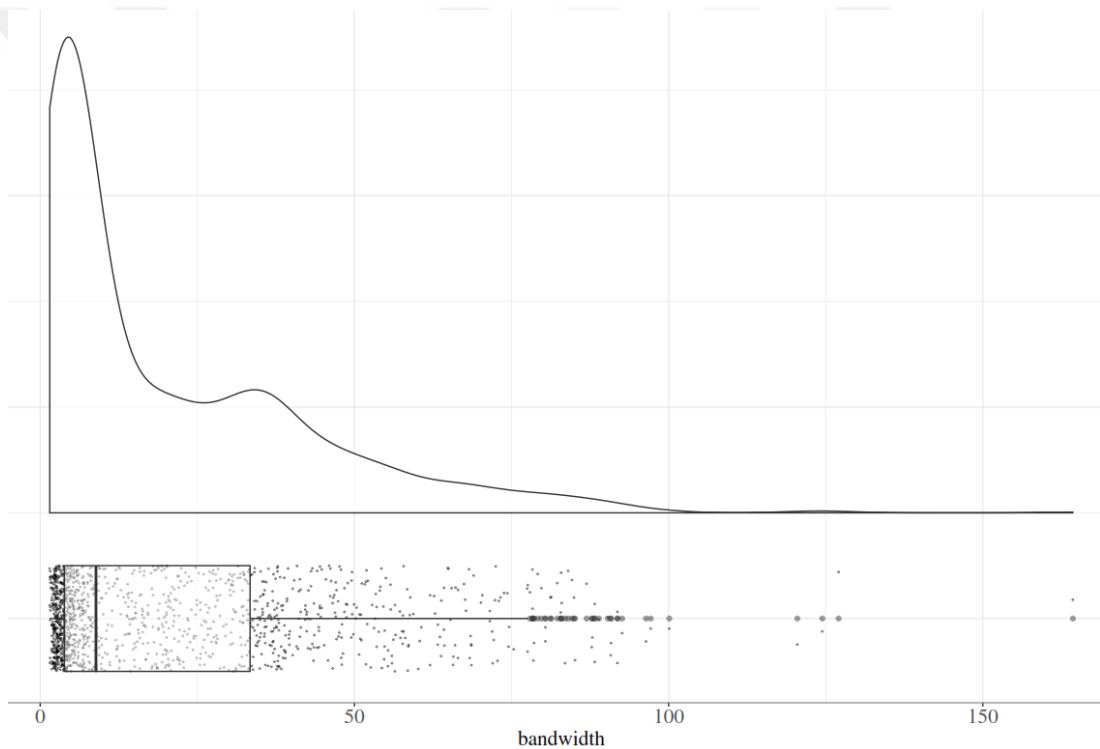


Figure 4.13. Distribution of bandwidth values in km

The model diagnostics are provided in the table below. The last three values in the table are primarily used for comparison with OLS. The GWR model explains 77.22% of the variation in Overall Satisfaction. This rate is calculated as 59.59% for OLS. Regarding model parsimony, as indicated by the AIC indices, GWR (2984) performs better than OLS (3375).

Table 4.47. Diagnostics of GWR compared to OLS.

	GWR	OLS
Kernel function:	Gauss	Gauss
Adaptive quantile:	0.0044	1.0000
Number of data points:	1363	1363
AIC:	2983.91	3375
Residual sum of squares:	532.4	942.6
Quasi-global R2:	0.7722	0.5959

According to the BFC test (Fotheringham et al., 2002) which evaluates the overall performance of the two models; the GWR model explains the data better than the OLS model ($F = 1.7704$, $df1 = 1359.00$, $df2 = 866.89$, $p\text{-value} < 2.2e-16$).

The distribution of the model coefficients is summarized in the table below. The OLS coefficient values are provided under the "Global" heading on the far right. While the intercept term serves a normalization purpose, the values of the factor scores' coefficients relative to their global values significantly indicate how much the model varies across the geography. This table provides a summary of the results detailed below and illustrates the extent to which local models vary compared to the global model. The values under the minimum and maximum columns in the table are significantly outside the general picture. The results presented below on a covariate basis will allow a clearer understanding of this situation.

Table 4.48. Basic distributional characteristics of GWR model coefficients

	Min	1st Quad	Median	3rd Quad	Max	Global
X.Intercept.	-5.131	-1.313	-0.616	0.071	10.962	-0.692
AHP_Factor1	-3.099	0.033	0.246	0.469	1.869	0.229
AHP_Factor2	-1.302	0.145	0.355	0.536	3.295	0.367

	Min	1 st Quad	Median	3 rd Quad	Max	Global
AHP_Factor3	-1.364	0.329	0.565	0.779	1.530	0.567

The visual representation of the model's explanatory power distribution indicates that, although the average explanatory power of the GWR model is calculated at 77.2%, the locally calculated explanatory power (local R^2) exceeds 90% in coastal areas and large cities, where observation density is relatively higher. Consequently, the model parameters that relate to AHP scores, which are model experts' perspectives and customer opinions, are significantly more reliable in these regions. Conversely, in certain inland and northern areas, the explanatory power of the GWR model is comparable only to that of the OLS model.

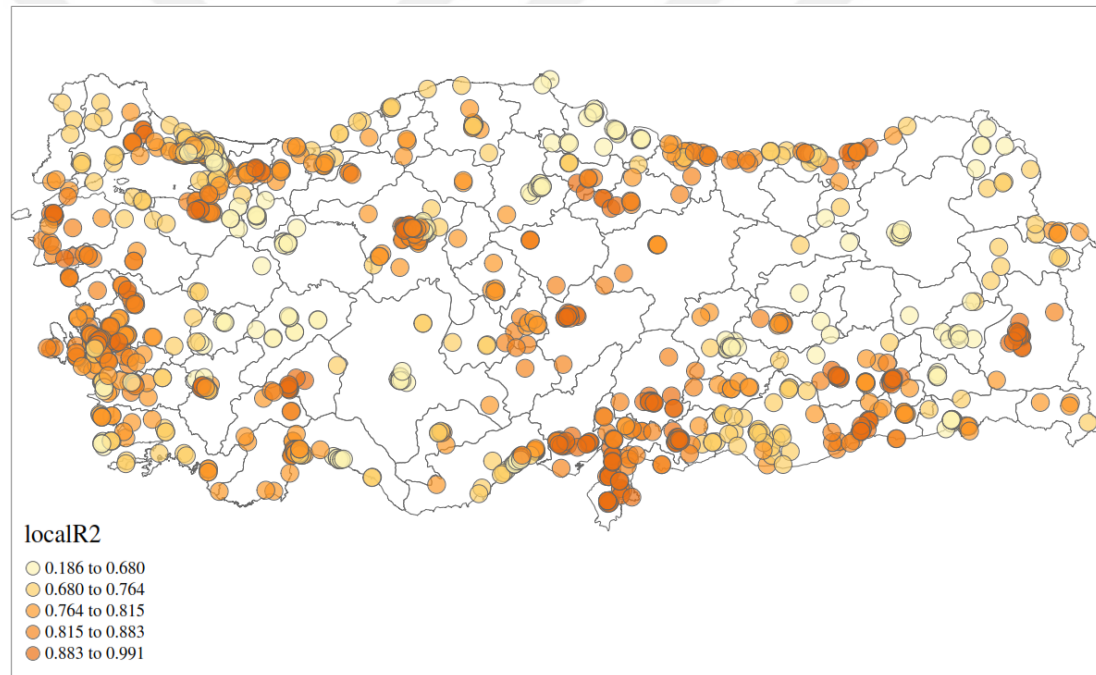


Figure 4.14. Geographic distribution of local R^2 values.

One of the main advantages of GWR models is the informative visual representation of their results. Detailed results for each covariate are presented below. For easier follow-up of the discussion, it is recommended to have access to the colored versions of the images provided below.

The visual representation below illustrates how the effect of Factor 1 on Overall Satisfaction varies across the map. Before examining the visual, note that the numerical ranges represented by the colors are provided in the legend table at the

bottom left corner of the visual (and subsequent visuals). These numerical ranges correspond to 20% percentiles, meaning the number of observations within each color group is equal. The software will utilize intermediate colors if the number of observations within the relevant value ranges is high. This point is crucial for a full understanding of the visuals presented later on. After providing this explanation, we can proceed to the detailed analysis.

The power of the GWR model becomes apparent with the presentation of results related to the first factor. Regions marked in shades of red indicate areas where customers believe Factor 1 negatively impacts Overall Satisfaction. Conversely, areas in shades of green indicate regions where customers believe Factor 1 positively affects Overall Satisfaction. Orange areas include the value zero (0), indicating regions where customers believe Factor 1 has no or negligible effect on Overall Satisfaction. The global value of Factor 1 corresponds to the yellow color in the legend table at the bottom left. Regions in yellow represent the OLS results for the effect of Factor 1 on Overall Satisfaction. Interestingly, yellow is not the predominant color in this visual; it is the least represented color on the graph. However, the OLS result falls within this range.

As we have stressed above, if an abundance of observations falls into that range, the software introduces a breakdown in color tone. But it did not.

When focusing on different regions of the map and conducting detailed examinations, variations can be observed even within neighborhoods of a city, let alone its districts. Investigating the reasons behind these variations falls outside the scope of this thesis, but the primary objective of highlighting the clear heterogeneities that require further research has been successfully achieved.

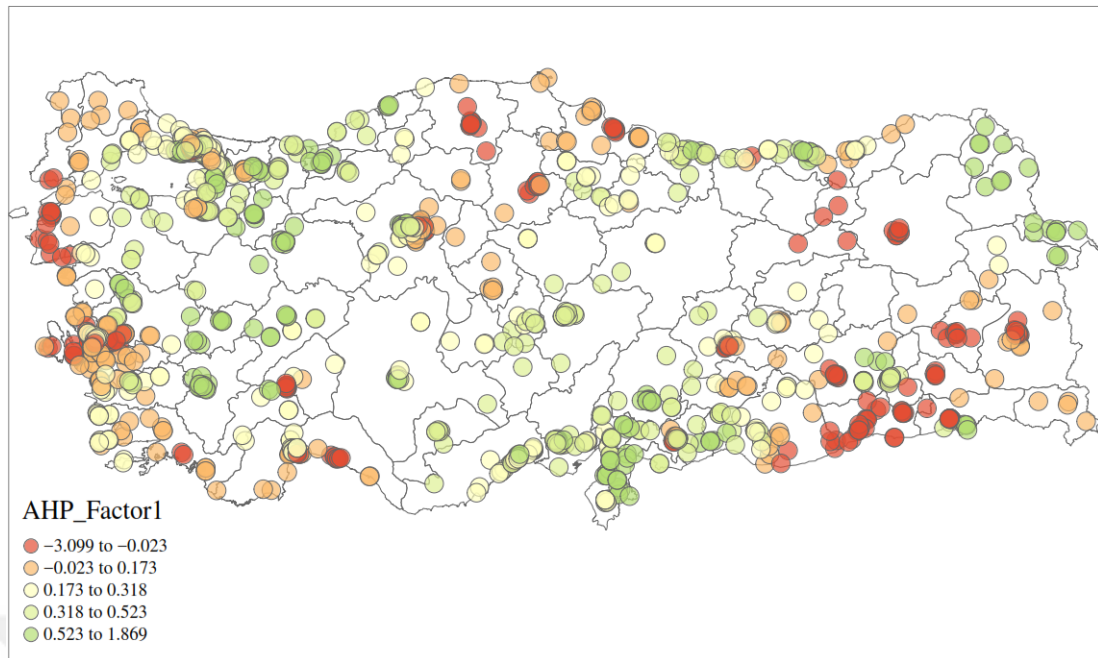


Figure 4.15. Geographic distribution of values of Factor1 coefficient in GWR

In addition to the geographical distribution of the Factor 1 coefficients, the descriptive statistics and distributional characteristics of the coefficients are provided in the table and figure below. The coefficients for Factor 1 range from -3.099 to 1.869, with an average value of 0.255. The coefficients exhibit a distribution that approximately follows the characteristics of a normal distribution. The zero value within the coefficient profile indicates that Factor 1 is not statistically significant in some models. Identifying the areas where Factor 1 is not significant for overall satisfaction could be a standalone sub-analysis, which can be quickly concluded following the GWR application.

Table 4.49. Six-point summary table of the Factor 1 coefficients

Min	1 st Quad	Median	Mean	3 rd Quad	Max
-3.099	0.033	0.246	0.255	0.469	1.869

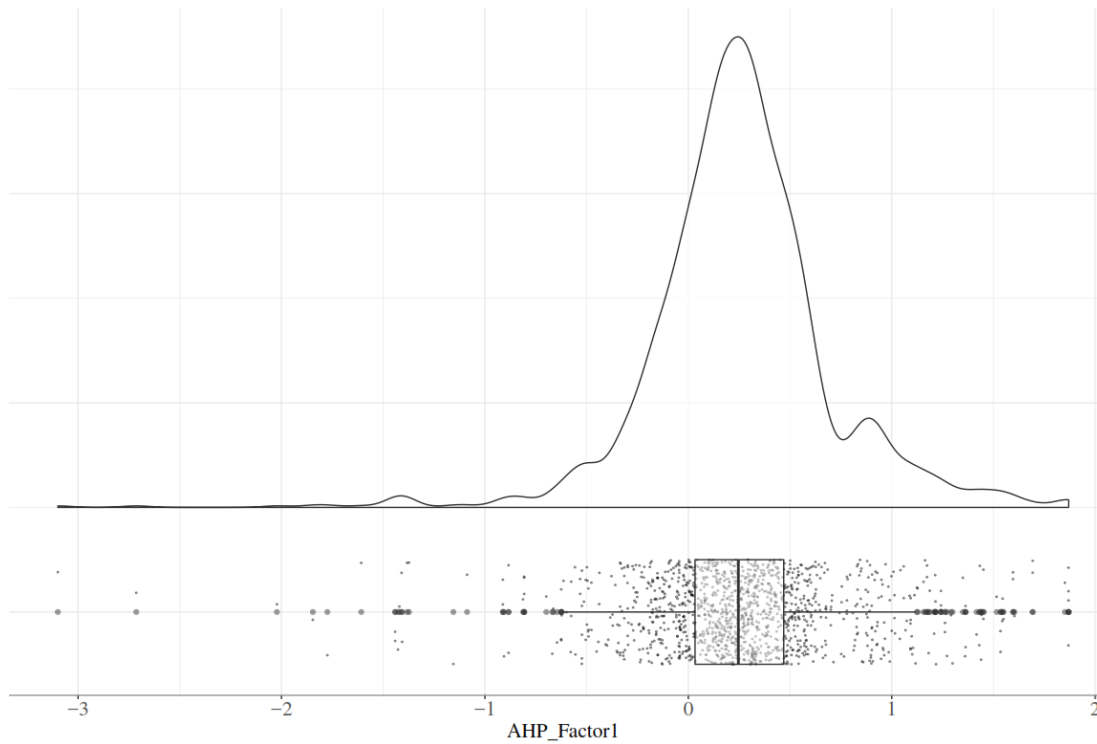


Figure 4.16. Distribution of values of Factor 1 coefficients

The distribution of the variable coefficient shown on the map can be filtered according to different research questions. For example, for the Factor1 coefficient, the OLS standard error was calculated as 0.031, and the 95% confidence interval was determined to be [0.167, 0.29] (Table 4.45).



Figure 4.17. Factor1 coefficients in the 95% confidence interval of OLS

The distribution of regions where the coefficient falls within this interval can also be shown on the map. As a result of this filtering, it is understood that 85% of the observation points are excluded from the visual, and only 211 out of 1363 data points have the Factor1 coefficient within the 95% confidence interval determined by OLS.

There are significant differences in opinion among customers regarding the impact of Factor2 on overall satisfaction. When the map is focused on major cities, it can be seen that both extremes of the coefficient's values are represented. Customers who believe that Factor2 strongly impacts overall satisfaction are predominantly concentrated in the coastal Aegean region, particularly in the cities of İzmir and Aydın. Conversely, customers who hold the opposite view are located in areas of the Marmara region outside Istanbul and in inland areas excluding Sivas and Çorum.

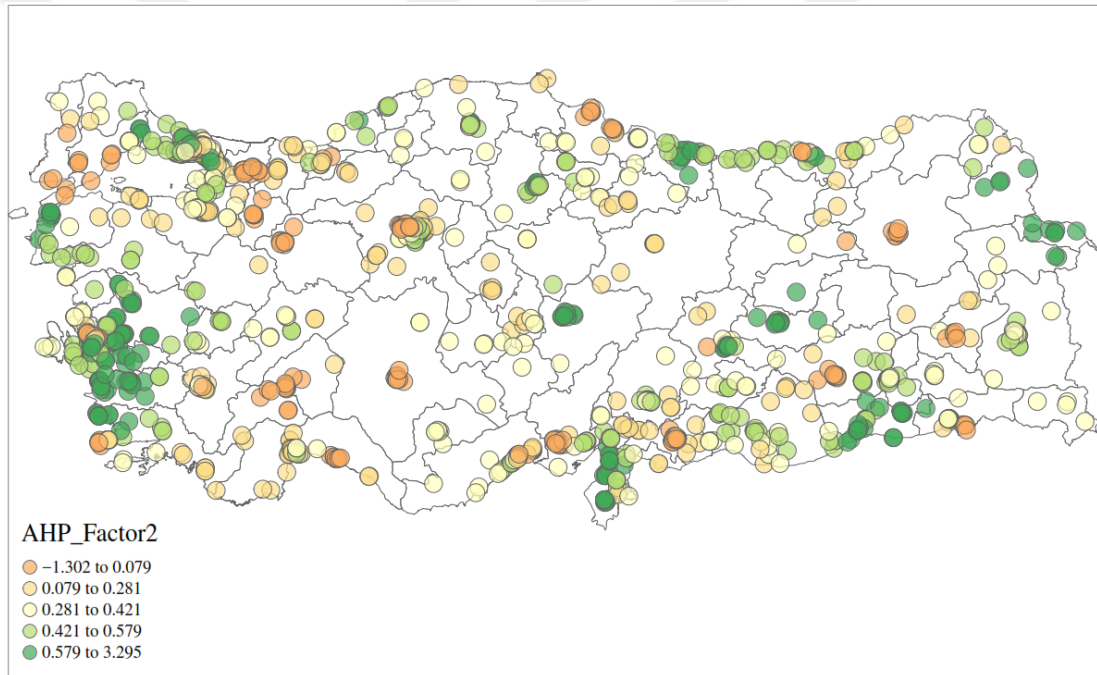


Figure 4.18. Geographic distribution of values of Factor2 coefficient in GWR

The table below and the subsequent figure present the basic descriptive statistics and distributional characteristics of the Factor 2 coefficients. Accordingly, the Factor 2 coefficients range from -1.302 to 3.295. The inclusion of zero within this range indicates that there are customer groups for whom the relationship between Factor 2 and overall satisfaction is insignificant.

Table 4.50. Six-point summary table of the Factor 2 coefficients

Min	1 st Quad	Median	Mean	3 rd Quad	Max
-1.302	0.145	0.355	0.325	0.536	3.295

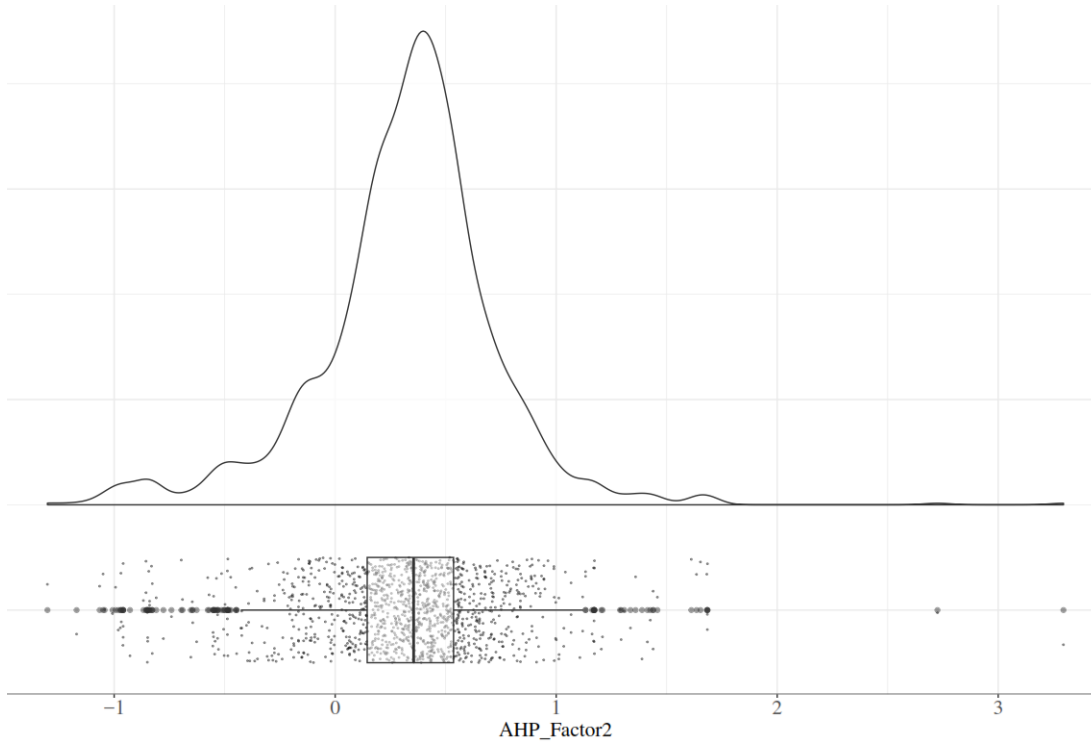


Figure 4.19. Distribution of values of Factor 2 coefficients

According to the OLS model, the impact of Factor 3 on overall satisfaction is as significant as the combined effect of the other two factors. However, according to the GWR model, this is not universally true. Observations depicted in red on the map include those where the coefficient values are close to zero. A closer inspection of these observations can distinguish that Factor 3 strongly negatively impacts overall satisfaction within this subset of customers. While we do not report this sub-analysis here, which we consider to be of managerial importance, we want to emphasize that once the GWR model is established, analyses can increasingly focus on strategic elements in greater detail.

Generally, Factor 3 is depicted in green tones on the map, indicating a consensus that it has a strong positive impact on overall satisfaction. However, exceptions to this consensus include areas such as Istanbul, the eastern Black Sea region, the southern and inland Aegean region, Hatay, and northeastern Anatolia.

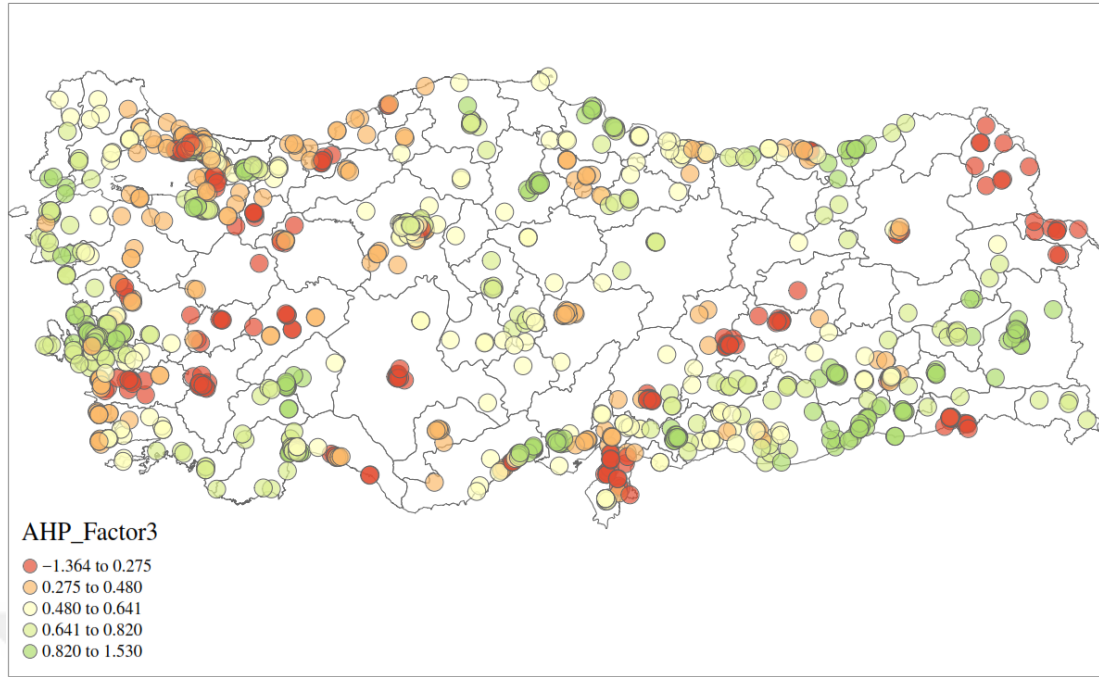


Figure 4.20. Geographic distribution of values of Factor3 coefficient in GWR

The variation of Factor 3 coefficients calculated in locals is summarized in the six-point summary table below, along with the distribution chart beneath the table. Coefficient values vary between -1.364 to 1.531. The zero value within the coefficient profile indicates that Factor 1 is not statistically significant in some models. We want to highlight the proximity of the Factor 3 coefficient's median value to the OLS value of 0.567. The same situation also applies to the coefficients of Factor 1 and Factor 2.

Table 4.51. Six-point summary table of the Factor 3 coefficients

Min	1 st Quad	Median	Mean	3 rd Quad	Max
-1.364	0.329	0.565	0.516	0.779	1.531

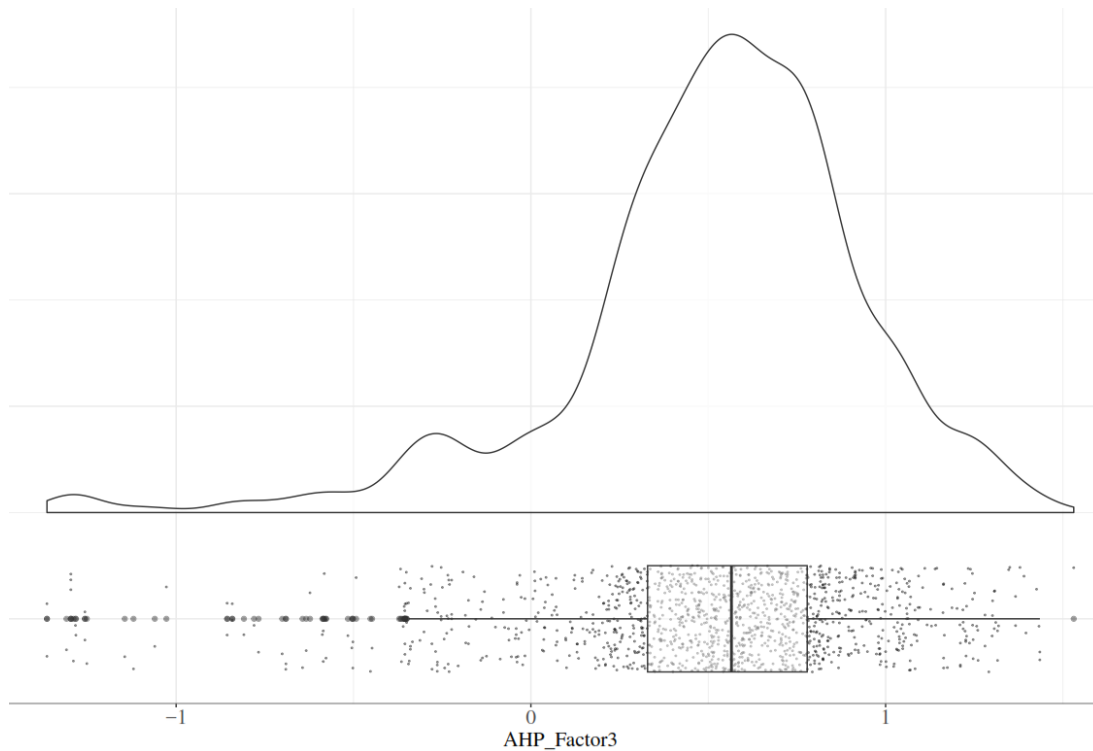


Figure 4.21. Distribution of values of Factor 3 coefficients

4.5.3. Geographically Weighted Regression Installation/Setup Service

The diagnostics for the comparison of OLS and GWR models are provided in the table below (Table 4.52). As a natural consequence of the reduction in the Residual Sum of Squares, the GWR model explains 82.18% of the variation in overall satisfaction, whereas the OLS model accounts for only 57.8%.

Table 4.52. Diagnostics of GWR compared to OLS

	GWR	OLS
Kernel function:	Gauss	Gauss
Adaptive quantile:	0.0044	1.0000
Number of data points:	993	993
AIC:	1171.83	1649.58
Residual sum of squares:	128.1	303.1

	GWR	OLS
Quasi-global R2:	0.8218	0.578

The distribution of GWR model parameters is provided in Table 4.53. The coefficients calculated in the OLS model, which is estimated by pooling the data, are also shown in the last column of the table. The coefficients of covariates in the GWR model vary significantly across the spatial dimension.

Table 4.53. Basic distributional characteristics of GWR model coefficients

	Min	1st Quad	Median	3rd Quad	Max	Global
X.Intercept.	-17.405	-1.049	-0.050	1.052	12.546	0.072
AHP_Factor1	-4.764	-0.158	0.036	0.272	4.571	0.000
AHP_Factor2	-3.400	0.119	0.312	0.598	4.568	0.312
AHP_Factor3	-1.446	0.266	0.575	0.788	1.974	0.661

The values range from -4.8 to 4.6 for AHP_Factor 1. On the other hand, in the OLS model, the AHP_Factor 1 variable is found to be insignificant. If the evaluation had been done using only the OLS model without establishing the GWR model, it would have been concluded that Factor 1 did not affect overall satisfaction. However, according to the GWR model, Factor 1 has a strong negative effect in roughly half of the regional regressions, while it has a strong positive effect in the other half. The OLS model averages out these two strong effects, leading to an incorrect result.

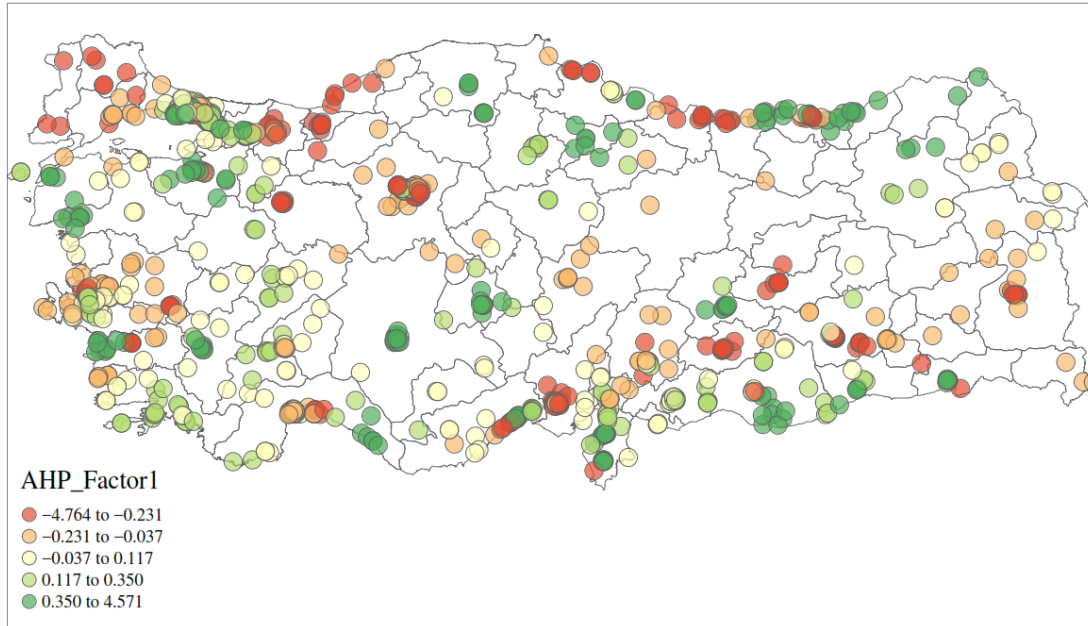


Figure 4.22. Geographic distribution of values of Factor 1 coefficient in GWR

The effect of the AHP_Factor 2 variable on overall satisfaction is observed to be negative across the Mediterranean region and Central Black Sea region, while it has a strong positive effect in İzmir, Central Anatolia, and Trakya (Figure 4.23).

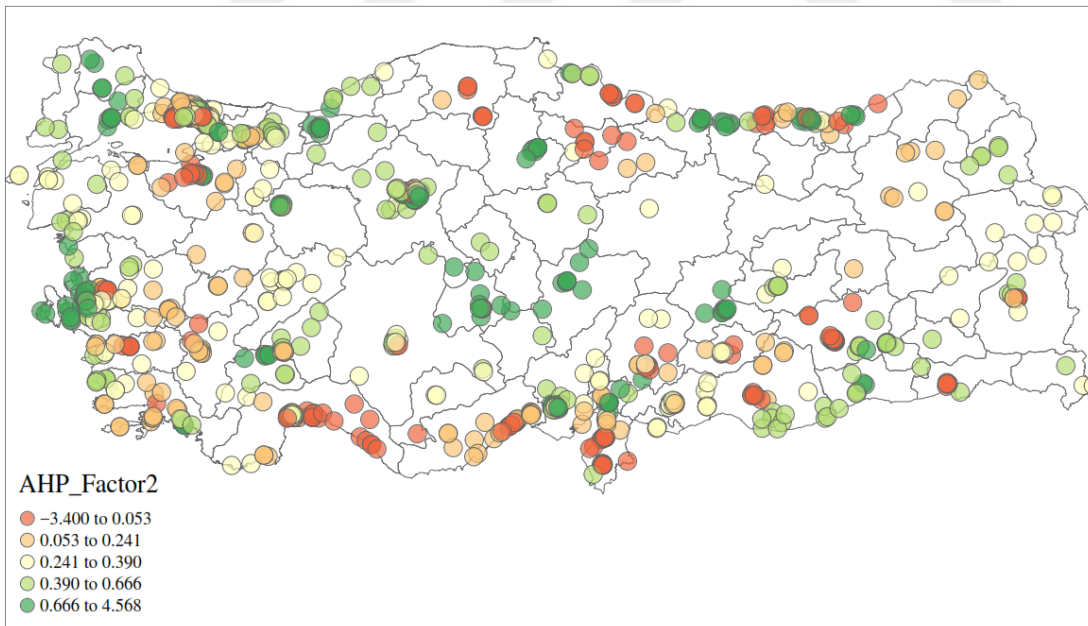


Figure 4.23. Geographic distribution of values of Factor 2 coefficient in GWR

The effect of AHP_Factor 3 on overall satisfaction is in the opposite direction to AHP_Factor 2. In areas where AHP_Factor 3 has a positive effect, AHP_Factor 2 has a negative effect, and in areas where AHP_Factor 3 has a negative effect, AHP_Factor 2 has a positive effect (Figure 4.24)

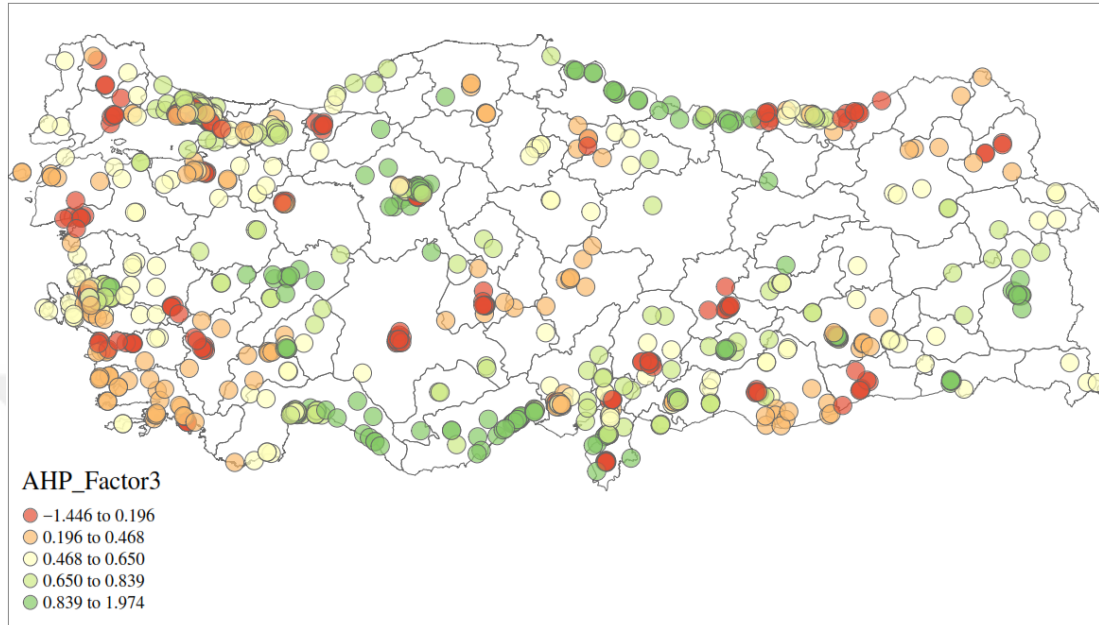


Figure 4.24. Geographic distribution of values of Factor 3 coefficient in GWR.

5. CHAPTER: DISCUSSION AND CONCLUSION

This thesis aims to propose a methodology that can be used by manufacturing companies to measure the quality of after-sales services. The proposed methodology seeks to address questions about the dimensions that need to be considered in evaluating the quality of after-sales services provided by a dealer network spread over a large geographic area, the weight these dimensions should be evaluated with, whether the geographical location of customers evaluating service quality in different dimensions affects this evaluation, and whether this variable can be used to explain service quality.

In this thesis, data from a company producing durable consumer goods and providing after-sales services are obtained by consulting the company. The company's after-sales services are in the form of repair and installation. The analyses were conducted separately since there are differences in evaluating these two service processes. In other words, the methods and measurements proposed in the thesis were applied to two different datasets. This also allows for comparing the results of evaluating these services across two different service types.

Firstly, the appropriateness of the questionnaire used by the company to measure service quality was evaluated. For this evaluation, scales in the literature were examined, and it was found that the scale used generally included the service attributes mentioned in the literature. Although different service attributes have been evaluated using different scales in the literature, it is seen that service quality is considered a multidimensional concept. The scale used by the company was also designed with different dimensions in mind. EFA was used to determine how many dimensions customers evaluate when evaluating the quality of the provided service using the collected data. The EFA results showed that the company could measure service quality in three dimensions for both types of service with the scale used. In the next step, CFA analyses were conducted to determine whether the scale used to measure service quality satisfies construct validity.

Despite obtaining a three-factor structure for both types of services, items representing all service quality dimensions mentioned in SERVQUAL and SERVPERF are included. The Responsiveness, Tangible, and Empathy dimensions merged into a single dimension, while the Reliability and Assurance dimensions remained separate. Although the content of the model dimensions created for both types of services was similar, the Knowledgeable Service, Innovative Service, and Fast Service items were included in the dimension structure for Installation/Setup services. In contrast, these items were not significant for Repair/Maintenance services.

Using EFA and CFA methods, conceptual models validated with data for after-sales service quality were presented. For the following research question of the thesis, the AHP method is used to determine the importance of weights between dimensions. The conceptual structure obtained from the EFA and CFA results is presented and explained to the experts selected for the study. The experts are asked to evaluate the importance of the dimensions and the items under each dimension. The weights obtained from these evaluations are reported, and the evaluation scores for each dimension are calculated as the weighted average of the items under that dimension.

Using the dimension scores calculated with the weights obtained from the AHP method, the question of whether there are geographical differences in evaluating service quality is answered. The relationship between the factor scores calculated from AHP and the Overall Service Satisfaction variable asked in a separate survey question are modeled using OLS in parallel with the literature. The same structure is also enriched with geographical data, and a GWR model is established. Generally, there is a 20% difference in performance between the OLS model, which pools the data, and the GWR models, which are established with geographical localization. The results show significant geographical variability in the impact of factor scores on overall satisfaction. It is observed that in some regions, the coefficients of factor scores even changed signs. In regions where the coefficients of factor scores are negative, it is understood that there was no impact on customers' overall service satisfaction. From a managerial perspective, investments focused on these dimensions and their content would produce more efficient and targeted results in actions to increase customers' overall service satisfaction.

Significant differences are found between the two types of services regarding the impact of factor scores on overall service satisfaction. The factor representing the Assurance dimension was found to have no significant impact on overall service satisfaction in the OLS model calculated with pooled data in the Installation/Setup dataset. On the other hand, it was found that in half of the local GWR models, the assurance dimension had a strong negative impact on overall service satisfaction. In contrast, it had a strong positive impact in the other half. When these two effects are pooled, they cancel each other out, leading to the misleading conclusion that the Assurance dimension does not impact overall service quality nationwide.

One of the prominent results of the GWR models developed for both types of services is that the relative impact of factor scores on overall service satisfaction varies across the spatial dimension. This variability is observed even at the district level in large cities with dense data. For example, in İzmir, it is found that the relative impact of factor scores in Çeşme, Urla, and Foça differs significantly from central districts such as Karşıyaka, Bornova, and Konak, and inland districts such as Ödemiş and Torbalı. This shows that GWR models can capture changes in service quality perception even in different city districts and detect changes across large geographical distances such as the east and west of the country.

After-sales service (ASS) is seen as a key distinguishing factor for companies involved in the sale of long-lasting products such as automobiles, computers, and household appliances. End consumers across several product categories inevitably require ASS at some stage. Consequently, the operations and strategic aims of ASS have become indispensable for manufacturers who produce long-lasting goods. Manufacturers preserve and oversee customer relationships by offering supplementary services. ASS is strategically important as it serves as a feedback channel for customers to provide input on the design, quality, and sales activities of the products.

Evaluating customers' perceptions of service quality and service satisfaction based on regional differences will help the company define geographic customer segments and develop regional strategies accordingly. The differences that need to be considered in evaluating different services, such as installation, maintenance, and repair provided in after-sales services, have not been addressed in the literature before. This thesis has also revealed the differences in evaluation between these two service types, which are

defined as Installation/Setup and Repair/Maintenance in the thesis. Although limited studies emphasize geographical differences in evaluating after-sales service quality, no study evaluates service quality by its dimensions and addresses the geographical differences of these dimensions. The results obtained in this study show how important it is for the company to make this evaluation.

Customers are placing more importance on both the quality of a product and its availability during its entire lifespan, as well as the services supplied throughout this period. In a near future where the linear economy model is no longer viable, and the focus is on the circular economy model, the longevity, repairability, reuse, and recycling of products are vital concerns for society. Manufacturers play a crucial role in the circular economy paradigm. The manufacturer's after-sales services are essential for ensuring the proper utilization, longevity, repairability, availability of spare parts, and the potential reuse or transformation of a product's components when it becomes unusable. For this reason, it is crucial to provide comprehensive and long-term customer services for durable goods. The results obtained in this study presents an approach that allows manufacturers to consider regional differences when planning their after-sales services. In this way, it will be possible to develop strategies specific to regions defined beyond the legally designated geographical boundaries. Consequently, this approach provides the opportunity to allocate limited resources more effectively.

One limitation of the thesis is that the data used is secondary data. The data used for the application in the thesis was already collected by a company providing after-sales services. Therefore, the measurement scale used for the thesis could not be designed. Additionally, the company has already determined the geographic distribution of the sample from which the data is collected. Although the survey design and sampling scheme met the objectives of the thesis, the survey design could be reviewed to highlight the results and increase the discriminant validity of the dimensions. Similarly, the number of observations could be increased for parameter stability, especially in regions with low data density in the GWR model.

The studies conducted in the thesis can be further improved in future studies. First, the CFA analysis can be enhanced to define a second-order overall satisfaction variable. AHP scores can be regressed on this variable to reveal the differences between expert

opinions and customers' perceptions of service quality. Second, the geographical distribution of service dimensions can be examined using the recently developed geographical principal component analysis methodology. However, this would require determining the number of data and sampling schemes from the outset. Third, the methodology proposed in this thesis can also be used to evaluate services provided in different sectors such as hospitality, healthcare, and banking.



6. REFERENCES

- Agnihotri, S., Sivasubramaniam, N., & Simmons, D. (2002). Leveraging technology to improve field service. *International Journal of Service Industry Management*, 13(1), 47–68. <https://doi.org/10.1108/09564230210421155>
- Alam, M. S., & Mondal, M. (2019). Assessment of sanitation service quality in urban slums of Khulna city based on SERVQUAL and AHP model: A case study of railway slum, Khulna, Bangladesh. *Journal of Urban Management*, 8(1), 20–27. <https://doi.org/10.1016/j.jum.2018.08.002>
- Altuntas, S., Dereli, T., & Yilmaz, M. K. (2012). Multi-criteria decision making methods based weighted SERVQUAL scales to measure perceived service quality in hospitals: A case study from Turkey. *Total Quality Management & Business Excellence*, 23(11–12), 1379–1395. <https://doi.org/10.1080/14783363.2012.661136>
- Alzola, L. M., & Robaina, V. P. (2005). Quality Management Journal Servqual: Its Applicability in Electronic Commerce B2C. *Quality Management Journal*, 12(4), 46–57. <https://doi.org/10.1080/10686967.2005.11919270>
- Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations (with discussion). *Journal of the Royal Statistical Society, Series B*, 26, 211–246. <https://ci.nii.ac.jp/naid/10009669212>
- Baglin, J. (2014). Improving your exploratory factor analysis for ordinal data: A demonstration using FACTOR. *Practical Assessment, Research, and Evaluation*, 19(1), 5. <https://doi.org/10.7275/dsep-4220>
- Bartlett, M. S. (1954). A note on the multiplying factors for various χ^2 approximations. *Journal of the Royal Statistical Society. Series B (Methodological)*, 296–298. <https://doi.org/10.1111/j.2517-6161.1954.tb00174.x>

- Bhattacharya, P., Mukhopadhyay, A., Saha, J., Samanta, B., Mondal, M., Bhattacharya, S., & Paul, S. (2023). Perception-satisfaction based quality assessment of tourism and hospitality services in the Himalayan region: An application of AHP-SERVQUAL approach on Sandakphu Trail, West Bengal, India. *International Journal of Geoheritage and Parks*, 11(2), 259–275. <https://doi.org/10.1016/j.ijgeop.2023.04.001>
- Bidanset, P. E., & Lombard, J. R. (2014). The effect of kernel and bandwidth specification in geographically weighted regression models on the accuracy and uniformity of mass real estate appraisal. *Journal of Property Tax Assessment & Administration*, 11(3), 5–14. https://digitalcommons.odu.edu/cgi/viewcontent.cgi?article=1027&context=publicservice_pubs
- Bradlow, E. T., Bronnenberg, B., Russell, G. J., Arora, N., Bell, D. R., Duvvuri, S. D., Hofstede, F. Ter, Sismeiro, C., Thomadsen, R., & Yang, S. (2005). Spatial Models in Marketing. *Marketing Letters*, 16(3–4), 267–278. <https://doi.org/10.1007/s11002-005-5891-3>
- Brogowicz, A. A., Delene, L. M., & Lyth, D. M. (1990). A synthesised service quality model with managerial implications. *International Journal of Service Industry Management*, 1(1), 27–45. <https://doi.org/10.1108/09564239010001640>
- Bronnenberg, B. J., & Mahajan, V. (2001). Unobserved retailer behavior in multimarket data: Joint spatial dependence in market shares and promotion variables. *Marketing Science*, 20(3), 284–299. <https://doi.org/10.1287/mksc.20.3.284.9768>
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. Guilford publications. <https://psycnet.apa.org/record/2015-10560-000>
- Buttle, F. (1996). SERVQUAL: review, critique, research agenda. *European Journal of Marketing*, 30(1), 8–32.
- Carman, J. (1990). Consumer perceptions of service quality: an assessment of the SERVQUAL dimensions. *Journal of Retailing*, 66(1), 33–55. <http://search.proquest.com/openview/9b58a6fed9a03970daaa0c1557ae7bb2/1?pq-origsite=gscholar&cbl=41988>

- Carroll, J. B. (1978). How shall we study individual differences in cognitive abilities?—Methodological and theoretical perspectives. *Intelligence*, 2(2), 87–115. [https://doi.org/10.1016/0160-2896\(78\)90002-8](https://doi.org/10.1016/0160-2896(78)90002-8)
- Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, 1(2), 245–276. https://doi.org/10.1207/s15327906mbr0102_10
- Chang, C.-W., & Zhang, J. Z. (2016). The effects of channel experiences and direct marketing on customer retention in multichannel settings. *Journal of Interactive Marketing*, 36(1), 77–90. <https://doi.org/10.1016/j.intmar.2016.05.002>
- Charlton, M. E., Brunsdon, C. F., & Fotheringham, A. S. (1998). Geographically weighted regression-modelling spatial non-stationarity. *Journal of The Royal Statistical Society, Series D: The Statistician*. <http://dx.doi.org/10.1111/1467-9884.00145>
- Chase, R. B. (1981). The Customer Contact Approach to Services: Theoretical Bases and Practical Extensions, *Operations Research*, 29(4), 698–706. <https://doi.org/10.1287/OPRE.29.4.698>
- Cheung, G. W., Cooper-Thomas, H. D., Lau, R. S., & Wang, L. C. (2024). Reporting reliability, convergent and discriminant validity with structural equation modeling: A review and best-practice recommendations. *Asia Pacific Journal of Management*, 41(2), 745–783. <https://doi.org/10.1007/s10490-023-09871-y>
- Comber, A., Brunsdon, C., Charlton, M., Dong, G., Harris, R., Lu, B., Lü, Y., Murakami, D., Nakaya, T., & Wang, Y. (2023). A route map for successful applications of geographically weighted regression. *Geographical Analysis*, 55(1), 155–178. <https://doi.org/10.1111/gean.12316>
- Crede, M., & Harms, P. (2019). Questionable research practices when using confirmatory factor analysis. *Journal of Managerial Psychology*, 34(1), 18–30. <https://doi.org/10.1108/jmp-06-2018-0272>
- Cronin, J. J., & Taylor, S. (1992). Measuring Service Quality: A Reexamination and Extension. *Journal of Marketing*, 56(3), 55–68. <https://doi.org/10.1177/002224299205600304>
- Cronin, J. J., & Taylor, S. A. (1994). SERVPERF versus SERVQUAL: reconciling performance-based and perceptions-minus-expectations measurement of service

- quality. *Journal of Marketing*, 58(1), 125–131.
<https://doi.org/10.1177/002224299405800110>
- Cudeck, R. (2000). Exploratory factor analysis. In *Handbook of applied multivariate statistics and mathematical modeling* (pp. 265–296). Elsevier.
<https://doi.org/10.1016/b978-012691360-6/50011-2>
- Dabholkar, P. A., Shepherd, C. D., & Thorpe, D. I. (2000). A comprehensive framework for service quality: an investigation of critical conceptual and measurement issues through a longitudinal study. *Journal of Retailing*, 76(2), 139–173. [https://doi.org/10.1016/s0022-4359\(00\)00029-4](https://doi.org/10.1016/s0022-4359(00)00029-4)
- Devaraj, S., Fan, M., & Kohli, R. (2002). Antecedents of B2C channel satisfaction and preference: validating e-commerce metrics. *Information Systems Research*, 13(3), 316–333. <https://doi.org/10.1287/isre.13.3.316.77>
- DeVellis, R. F., & Thorpe, C. T. (2021). *Scale development: Theory and applications*. Sage publications. <https://ci.nii.ac.jp/ncid/BB06453461>
- Dimitriadis, Z. S., & Maroudas, T. S. (2007). Demographic predictors of service satisfaction in Greek public organizations. *Measuring Business Excellence*, 11(2), 32–43. <https://doi.org/10.1108/13683040710752724>
- DiStefano, C. (2002). The impact of categorization with confirmatory factor analysis. *Structural Equation Modeling*, 9(3), 327–346.
https://doi.org/10.1207/s15328007sem0903_2
- DiStefano, C., & Hess, B. (2005). Using confirmatory factor analysis for construct validation: An empirical review. *Journal of Psychoeducational Assessment*, 23(3), 225–241. <https://doi.org/10.1177/073428290502300303>
- Durvasula, S., Lysonski, S., & Mehta, S. C. (1999). Testing the SERVQUAL scale in the business-to-business sector: the case of ocean freight shipping service. *Journal of Services Marketing*, 13(2), 132–150.
<https://doi.org/10.1108/08876049910266040>
- Emrouznejad, A., & Marra, M. (2017). The state of the art development of AHP (1979–2017): A literature review with a social network analysis. *International Journal of Production Research*, 55(22), 6653–6675.
<https://doi.org/10.1080/00207543.2017.1334976>

- Fabrigar, L. R., & Wegener, D. T. (2011). *Exploratory factor analysis*. Oxford University Press.
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4(3), 272. <https://doi.org/10.1037/1082-989x.4.3.272>
- Field, A. P. (2000). *Discovering statistics using SPSS for Windows: Advanced techniques for beginners*. Sage Publications, Inc.
- Fotheringham, A. S., Brunsdon, C., & Charlton, M. E. (2009). Geographically weighted regression. *The Sage Handbook of Spatial Analysis*, 1, 243–254.
- Fotheringham, S., Brundson, C., & Charlton, M. (2002). *Geographically weighted regression & associated techniques*. Wiley.
- Fragoso, J. T., & Espinoza, I. L. (2017). Assessment of banking service quality perception using the SERVPERF model. *Contaduría y Administración*, 62(4), 1294–1316. <https://doi.org/10.1016/j.cya.2017.06.011>
- Franco, M. A. (2019). A system dynamics approach to product design and business model strategies for the circular economy. *Journal of Cleaner Production*, 241, 118327. <https://doi.org/10.1016/j.jclepro.2019.118327>
- Frost, F. A., & Kumar, M. (2000). INTSERVQUAL – an internal adaptation of the GAP model in a large service organisation. *Journal of Services Marketing*, 14(5), 358–377. <https://doi.org/10.1108/08876040010340991>
- Gentry, J. W., Tansuhaj, P., Manzer, L. L., & John, J. (1988). Do geographic subcultures vary culturally? *Advances in Consumer Research*, 15(1). <https://www.acrwebsite.org/volumes/6839/volumes/v15/NA-15>
- Gerbing, D. W., & Anderson, J. C. (1988). An updated paradigm for scale development incorporating unidimensionality and its assessment. *Journal of Marketing Research*, 25(2), 186–192. <https://doi.org/10.1177/002224378802500207>
- Ghotbabadi, A. R., Feiz, S., & Baharun, R. (2015). Service quality measurements: a review. *International Journal of Academic Research in Business and Social Sciences*, 5(2), 267–286. <http://dx.doi.org/10.6007/IJARBSS/v5-i2/1484>

- Govind, R., Chatterjee, R., & Mittal, V. (2017). Segmentation of Spatially Dependent Geographical Units: Model and Application. *Management Science*, August, mns.2016.2699. <https://doi.org/10.1287/mns.2016.2699>
- Govindasamy, P., Isa, N. J. M., Mohamed, N. F., Noor, A. M., Ma, L., Olmos, A., & Green, K. (2024). A systematic review of exploratory factor analysis packages in R software. *Wiley Interdisciplinary Reviews: Computational Statistics*, 16(1), e1630. <https://doi.org/10.1002/WICS.1630>
- Grönroos, C. (1984). A Service Quality Model and its Marketing Implications. *European Journal of Marketing*, 18(4), 36–44. <https://doi.org/10.1108/EUM0000000004784>
- Groskurth, K., Bluemke, M., & Lechner, C. M. (2023). Why we need to abandon fixed cutoffs for goodness-of-fit indices: An extensive simulation and possible solutions. *Behavior Research Methods*, 56(4), 3891–3914. <https://doi.org/10.3758/s13428-023-02193-3>
- Habib, M. M., & Victor, B. (1991). Strategy, structure, and performance of U.S. manufacturing and service MNCs: A comparative analysis. *Strategic Management Journal*, 12(8), 589–606. <https://doi.org/10.1002/SMJ.4250120803>
- Hair, J. F., Black, W. C., & Babin, B. J. (2010). *Multivariate Data Analysis: A Global Perspective (7th ed.)*. Pearson Education.
- Harrell, F. E. (2015). *Regression modeling strategies with applications to linear models, logistic and ordinal regression, and survival analysis*. Springer.
- Hayashi, K., Bentler, P. M., & Yuan, K.-H. (2007). Structural Equation Modeling. *Handbook of Statistics*, 27, 395–428. [https://doi.org/10.1016/S0169-7161\(07\)27013-0](https://doi.org/10.1016/S0169-7161(07)27013-0)
- Haywood-Farmer, J. (1988). A Conceptual Model of Service Quality. *International Journal of Operations & Production Management*, 8(6), 19–29. <https://doi.org/10.1108/eb054839>
- Ho, W. (2008). Integrated analytic hierarchy process and its applications—A literature review. *European Journal of Operational Research*, 186(1), 211–228. <https://doi.org/10.1016/j.ejor.2007.01.004>

- Hofmann, R. J. (1978). Complexity and simplicity as objective indices descriptive of factor solutions. *Multivariate Behavioral Research*, 13(2), 247–250. https://doi.org/10.1207/s15327906mbr1302_9
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30, 179–185. <https://doi.org/10.1007/bf02289447>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Humphreys, L. G., & Montanelli Jr, R. G. (1975). An investigation of the parallel analysis criterion for determining the number of common factors. *Multivariate Behavioral Research*, 10(2), 193–205. https://doi.org/10.1207/s15327906mbr1002_5
- Hussain, M., Javed, A., Khan, S. H., & Yasir, M. (2024). Pillars of Customer Retention in the Services Sector: Understanding the Role of Relationship Marketing, Customer Satisfaction, and Customer Loyalty. *Journal of the Knowledge Economy*, 1–21. <https://doi.org/10.1007/s13132-024-02060-2>
- Hussain, N., Bhatt, W., & Jilani, A. (2011). An Empirical Analysis of After Sales Service and Customer Satisfaction. *Management & Marketing*, 6(4), 561–572. <http://www.managementmarketing.ro/pdf/articole/243.pdf>
- Ishizaka, A., & Labib, A. (2011). Review of the main developments in the analytic hierarchy process. *Expert Systems with Applications*, 38(11), 14336–14345. <https://doi.org/10.1016/j.eswa.2011.04.143>
- Izquierdo, I., Olea, J., & Abad, F. J. (2014). Exploratory factor analysis in validation studies: Uses and recommendations. *Psicothema*, 395–400. <https://doi.org/10.7334/psicothema2013.349>
- Jain, S. K., & Gupta, G. (2004). Measuring service quality: Servqual vs. Servperf scales. *Vikalpa the Journal for Decision Makers*, 29(2), 25–38. <https://doi.org/10.1177/0256090920040203>
- JASP Team. (2024). *JASP (Version 0.18.3)[Computer software]*. <https://jasp-stats.org/>

- Kahle, L. R. (1986). The nine nations of North America and the value basis of geographic segmentation. *Journal of Marketing*, 50(2), 37–47. <https://doi.org/10.2307/1251598>
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39(1), 31–36. <https://doi.org/10.1007/bf02291575>
- Kasper, H., & Lemmink, J. (1989). After sales service quality: Views between industrial customers and service managers. *Industrial Marketing Management*, 18(3), 199–208. [https://doi.org/10.1016/0019-8501\(89\)90036-9](https://doi.org/10.1016/0019-8501(89)90036-9)
- Knecht, T., Leszinski, R., & Weber, F. A. (1993). Making profits after the sale. *The McKinsey Quarterly*, 4, 79.
- Leong, L. Y., Hew, T. S., Lee, V. H., & Ooi, K. B. (2015). An SEM–artificial-neural-network analysis of the relationships between SERVPERF, customer satisfaction and loyalty among low-cost and full-service airline. *Expert Systems with Applications*, 42(19), 6620–6634. <https://doi.org/10.1016/j.eswa.2015.04.043>
- Liu, D., Bishu, R. R., & Najjar, L. (2005). Using the analytical hierarchy process as a tool for assessing service quality. *Industrial Engineering and Management Systems*, 4(2), 129–135.
- Lohr, S. L. (2021). *Sampling: design and analysis*. Chapman and Hall/CRC. <https://doi.org/10.1201/9780429298899>.
- Mardani, A., Jusoh, A., Zavadskas, E. K., Khalifah, Z., & Nor, K. M. D. (2015). Application of multiple-criteria decision-making techniques and approaches to evaluating of service quality: a systematic review of the literature. *Journal of Business Economics and Management*, 16(5), 1034–1068. <https://doi.org/10.3846/16111699.2015.1095233>
- Maydeu-Olivares, A. (2017). Assessing the size of model misfit in structural equation models. *Psychometrika*, 82(3), 533–558. <https://doi.org/10.1007/s11336-016-9552-7>
- McNeish, D. (2023). Generalizability of dynamic fit index, equivalence testing, and Hu & Bentler cutoffs for evaluating fit in factor analysis. *Multivariate Behavioral Research*, 58(1), 195–219. <https://doi.org/10.1080/00273171.2022.2163477>

- Mittal, V., & Kamakura, W. A. (2001). Satisfaction, repurchase intent, and repurchase behavior: Investigating the moderating effect of customer characteristics. *Journal of Marketing Research*, 38(1), 131–142. <https://doi.org/10.1509/jmkr.38.1.131.18832>
- Mittal, V., Kamakura, W. A., & Govind, R. (2004). Geographic patterns in customer service and satisfaction: An empirical investigation. *Journal of Marketing*, 68(3), 48–62. <https://doi.org/10.1509/jmkg.68.3.48.34766>
- Moeller, S. (2010). Characteristics of services - a new approach uncovers their value. *Journal of Services Marketing*, 24(5), 359–368. <https://doi.org/10.1108/08876041011060468>
- Mundfrom, D. J., Shaw, D. G., & Ke, T. L. (2005). Minimum sample size recommendations for conducting factor analyses. *International Journal of Testing*, 5(2), 159–168. https://doi.org/10.1207/s15327574ijt0502_4
- Murali, S., Pugazhendhi, S., & Muralidharan, C. (2016). Modelling and Investigating the relationship of after sales service quality with customer satisfaction, retention and loyalty – A case study of home appliances business. *Journal of Retailing and Consumer Services*, 30, 67–83. <https://doi.org/10.1016/j.jretconser.2016.01.001>
- Nasir, M., Adil, M., & Dhamija, A. (2021). The synergetic effect of after sales service, customer satisfaction, loyalty and repurchase intention on word of mouth. *International Journal of Quality and Service Sciences*, 13(3), 489–505. <https://doi.org/10.1108/ijqss-01-2021-0015>
- Nye, C. D. (2023). Reviewer resources: Confirmatory factor analysis. *Organizational Research Methods*, 26(4), 608–628. <https://doi.org/10.1177/1094428122112054>
- Nye, C. D., & Drasgow, F. (2011). Assessing goodness of fit: Simple rules of thumb simply do not work. *Organizational Research Methods*, 14(3), 548–570. <https://doi.org/10.1177/1094428110368562>
- Olsson, U. (1979). Maximum likelihood estimation of the polychoric correlation coefficient. *Psychometrika*, 44(4), 443–460. <https://doi.org/10.1007/bf02296207>
- Osborne, J. W. (2015). What is rotating in exploratory factor analysis? *Practical Assessment, Research, and Evaluation*, 20(2), 1–7. <https://doi.org/10.7275/hb2g-m060>

- Pakdil, F., & Aydın, Ö. (2007). Expectations and perceptions in airline services: An analysis using weighted SERVQUAL scores. *Journal of Air Transport Management*, 13(4), 229–237. <https://doi.org/10.1016/j.jairtraman.2007.04.001>
- Parasuraman, A., Berry, L. L., & Zeithaml, V. A. (1991). Refinement and reassessment of the SERVQUAL scale. *Journal of Retailing*, 67(4), 420.
- Parasuraman, A., Zeithaml, V. A. and Berry, L. L. (1988). SERVQUAL: A multiple-Item Scale for measuring consumer perceptions of service quality. *Journal of Retailing*, 64(1), 12–40. <https://psycnet.apa.org/record/1992-37673-001>
- Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1985). A Conceptual Model of Service Quality and Its Implications for Future Research. *Journal of Marketing*, 49(4), 41. <https://doi.org/10.2307/1251430>
- Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1994). Reassessment of expectations as a comparison standard in measuring service quality: implications for further research. *Journal of Marketing*, 58(1), 111–124. <https://doi.org/10.2307/1252255>
- Parry, G., Newnes, L., & Huang, X. (2011). Goods, Products and Services. In G. and A. J. Macintyre Mairi and Parry (Ed.), *Service Design and Delivery* (pp. 19–29). Springer US. https://doi.org/10.1007/978-1-4419-8321-3_2
- Philip, G., & Hazlett, S. (1997). The measurement of service quality: a new P-C-P attributes model. *International Journal of Quality & Reliability Management*, 14(3), 260–286. <https://doi.org/10.1108/02656719710165482>
- Polyakova, O., & Mirza, M. (2015). Perceived service quality models: are they still relevant? *The Marketing Review*, 15(1), 59–82. <https://doi.org/10.1362/146934715x14267608178721>
- Qin, H., & Prybutok, V. R. (2009). Service quality, customer satisfaction, and behavioral intentions in fast-food restaurants. *International Journal of Quality and Service Sciences*, 1(1), 78–95. <https://doi.org/10.1108/17566690910945886>
- Rigopoulou, I. D., Chaniotakis, I. E., Lymperopoulos, C., & Siomkos, G. I. (2008). After-sales service quality as an antecedent of customer satisfaction: The case of electronic appliances. *Managing Service Quality*, 18(5), 512–527. <https://doi.org/10.1108/09604520810898866>

- Rogers, P. (2024). Best practices for your confirmatory factor analysis: A JASP and lavaan tutorial. *Behavior Research Methods*, 1–21. <https://doi.org/10.3758/s13428-024-02375-7>
- Roos, J. M., & Bauldry, S. (2021). *Confirmatory factor analysis*. Sage Publications.
- Saaty, T. L. (1977). A scaling method for priorities in hierarchical structures. *Journal of Mathematical Psychology*, 15(3), 234–281. [https://doi.org/10.1016/0022-2496\(77\)90033-5](https://doi.org/10.1016/0022-2496(77)90033-5)
- Saaty, T. L. (2004). Decision making—the analytic hierarchy and network processes (AHP/ANP). *Journal of Systems Science and Systems Engineering*, 13, 1–35. <https://doi.org/10.1007/s11518-006-0151-5>
- Saaty, T. L., & Vargas, L. (2001). Models, Methods, Concepts & Applications of the Analytic Hierarchy Process. In *International series in management science/operations research/International series in operations research & management science*. <https://doi.org/10.1007/978-1-4614-3597-6>
- Saccani, N., Songini, L., & Gaiardelli, P. (2006). The role and performance measurement of after-sales in the durable consumer goods industries: An empirical study. *International Journal of Productivity and Performance Management*, 55(3–4), 259–283. <https://doi.org/10.1108/17410400610653228>
- Sauvé, S., Bernard, S., & Sloan, P. (2016). Environmental sciences, sustainable development and circular economy: Alternative concepts for trans-disciplinary research. *Environmental Development*, 17, 48–56. <https://doi.org/10.1016/j.envdev.2015.09.002>
- Seth, N., Deshmukh, S. G., & Vrat, P. (2005). Service quality models: a review. *International Journal of Quality & Reliability Management*, 22(9), 913–949. DOI 10.1108/02656710510625211
- Shi, D., & Maydeu-Olivares, A. (2020). The effect of estimation methods on SEM fit indices. *Educational and Psychological Measurement*, 80(3), 421–445. <https://doi.org/10.1177/0013164419885164>
- Shi, D., Maydeu-Olivares, A., & Rosseel, Y. (2020). Assessing fit in ordinal factor analysis models: SRMR vs. RMSEA. *Structural Equation Modeling: A*

- Multidisciplinary Journal*, 27(1), 1–15.
<https://doi.org/10.1080/10705511.2019.1611434>
- Simmons, D. (2001). *Field service management: a classification scheme and study of server flexibility*. Binghamton University.
- Spreng, R. A., & Mackoy, R. D. (1996). An empirical examination of a model of perceived service quality and satisfaction. *Journal of Retailing*, 72(2), 201–214.
[https://doi.org/10.1016/s0022-4359\(96\)90014-7](https://doi.org/10.1016/s0022-4359(96)90014-7)
- Tabachnick, B. G., Fidell, L. S., & Ullman, J. B. (2013). *Using multivariate statistics* (Vol. 6). Pearson Boston, MA.
- Teas, R. K. (1993). Expectations, Performance Evaluation, and Consumers' Perceptions of Quality. *Journal of Marketing*, 57(4), 18–34.
<https://doi.org/10.1177/002224299305700402>
- Ter Hofstede, F., Steenkamp, J.-B. E. M., & Wedel, M. (1999). International market segmentation based on consumer–product relations. *Journal of Marketing Research*, 36(1), 1–17. <https://doi.org/10.1177/002224379903600101>
- The jamovi project. (2024). *jamovi (Version 2.5) [Computer Software]*.
<https://www.jamovi.org>
- Trochim, W. M., Donnelly, J. P., & Arora, K. (2016). *Research Methods: The Essential Knowledge*. Cengage Learning Boston, MA.
- Verhoef, P. C., & Lemon, K. N. (2013). Successful customer value management: Key lessons and emerging trends. *European Management Journal*, 31(1), 1–15.
<https://doi.org/10.1016/j.emj.2012.08.001>
- Watkins, M. W. (2018). Exploratory Factor Analysis: A Guide to Best Practice. *Journal of Black Psychology*, 44(3), 219–246.
<https://journals.sagepub.com/doi/10.1177/0095798418771807>
- Westen, D., & Rosenthal, R. (2003). Quantifying construct validity: two simple measures. *Journal of Personality and Social Psychology*, 84(3), 608.
<https://doi.org/10.1037/0022-3514.84.3.608>

- Wind, Y., & Saaty, T. L. (1980). Marketing applications of the analytic hierarchy process. *Management Science*, 26(7), 641–658.
<https://doi.org/10.1287/mnsc.26.7.641>
- Woods, C. M., & Edwards, M. C. (2011). Factor Analysis and Related Methods. In C. R. Rao, J. P. Miller, & D. C. Rao (Eds.), *Essential Statistical Methods for Medical Statistics* (pp. 174–201). North-Holland.
<https://doi.org/https://doi.org/10.1016/B978-0-444-53737-9.50009-8>
- Xia, Y., & Yang, Y. (2019). RMSEA, CFI, and TLI in structural equation modeling with ordered categorical data: The story they tell depends on the estimation methods. *Behavior Research Methods*, 51, 409–428.
<https://doi.org/10.3758/s13428-018-1055-2>
- Yarimoglu, E. K. (2014). A review on dimensions of service quality models. *Journal of Marketing Management*, 2(2), 79–93.
- Zeithaml, V. (1981). How Consumer Evaluation Processes Differ Between Goods and Services. In *Marketing of services* (Vol. 9, pp. 186–190).
- Zeithaml, V. A., Parasuraman, A., & Berry, L. L. (1990). *Delivering quality service: Balancing customer perceptions and expectations*. Simon and Schuster.
- Zhang, S. (2023, November 13). *R Programming for Psychometrics*.
https://bookdown.org/sz_psyc490/r4psychometrics/

APPENDICES

APPENDIX 1. HOAM Questionnaire

COMMON QUESTIONS	1.	Overall, considering the installation or repair service you received from HOAM Customer Services, could you rate your satisfaction?	
	2.	Considering this installation/repair service, could you rate your level of recommendation?	
	3.	Considering this installation service, could you rate your likelihood of receiving service from HOAM Customer Services again?	
	4.	Could you rate your level of agreement with the following statements?	
	4.	1.	The overall image is reputable.
	4.	2.	It is understanding towards its customers.
	4.	3.	It is attentive to its customers.
	4.	4.	The employees are knowledgeable in their respective fields.
	5.	Could you rate your level of agreement?	
	5.	1.	The service I received was worth the money I paid.
	5.	2.	It provides quality service.
	5.	3.	It provides fast service.
	6.	Was a fee charged for the service you received?	

INSTALLATION/SETUP	7.	Were you contacted by the authorized service for an appointment regarding your installation (product setup) process?	
	8.	"Could you rate your satisfaction with the authorized service's adherence to the appointment schedule for the installation (product setup) service?"	
	9.	Did the authorized service technician (personnel) use shoe covers when entering your residence?	
	10.	Considering the installation service you received, could you rate your satisfaction with the following statements?	
	10.	1.	The technician's appearance and clothing being neat and clean upon arriving at the house.
	10.	2.	Your satisfaction with the technician's behavior and communication upon arriving at the house.
	10.	3.	The authorized service technician providing you with information about the product and how to use it.
	10.	4.	The technician leaving the area clean and tidy after the installation.
	10.	5.	Your satisfaction with 'easy communication with the authorized service' during the installation process.
REPAIR /MAINTENANCE	7.	Could you rate your satisfaction with the authorized service's adherence to the appointment schedule for the repair service?	
	8.	Did the authorized service technician (personnel) use shoe covers when entering your residence?	
	9.	Considering the repair service you received, could you rate your satisfaction with the following statements?	

	9.	1.	The technician's appearance and clothing being neat and clean upon arriving at the house.
	9.	2.	Your satisfaction with the technician's behavior and communication upon arriving at the house.
	9.	3.	The authorized service technician providing you with information about the product malfunction and its causes.
	9.	4.	The technician leaving the area clean and tidy after the repair.
	9.	5.	Providing information by the authorized service when necessary, during the repair process.
	9.	6.	Your satisfaction with 'easy communication with the authorized service' during the repair process.
	10.	Was your repair service performed correctly the first time?	

APPENDIX 2. EFA Results for Installation/Setup Service

Following the procedure used in the Repair dataset, all results in here are derived from the polychoric correlation matrix obtained from the Installation/Setup dataset.

Kaiser-Meyer-Olkin Test

	MSA
Overall MSA	0.921
Recommendation	0.886
Opt for Service	0.923
Reputable Image	0.946
Understanding Service	0.872
Attentive Service	0.876
Knowledgeable Service	0.946
Innovative Service	0.938
Fast Service	0.938
Technicians Keeping the Appointment Time	0.935
Technician Appearance	0.929
Technician Behavior and Communication	0.922
Knowledgeable About Product	0.942
Technician Post Work Cleanliness	0.900
Technician Accessibility	0.966

Kaiser-Meyer-Olkin Test

		MSA
Quality Service		0.924

Bartlett test approves applicability of EFA.

Bartlett's Test

X ²	df	p
12661.472	105.000	< .001

According to parallel analysis the suggested number of factors to be extracted from data is 3.

Parallel Analysis

	Real Eigenvalues	Simulated Eigenvalues
Factor 1*	7.520	1.196
Factor 2*	1.831	1.145
Factor 3*	1.164	1.117
Factor 4	0.738	1.087

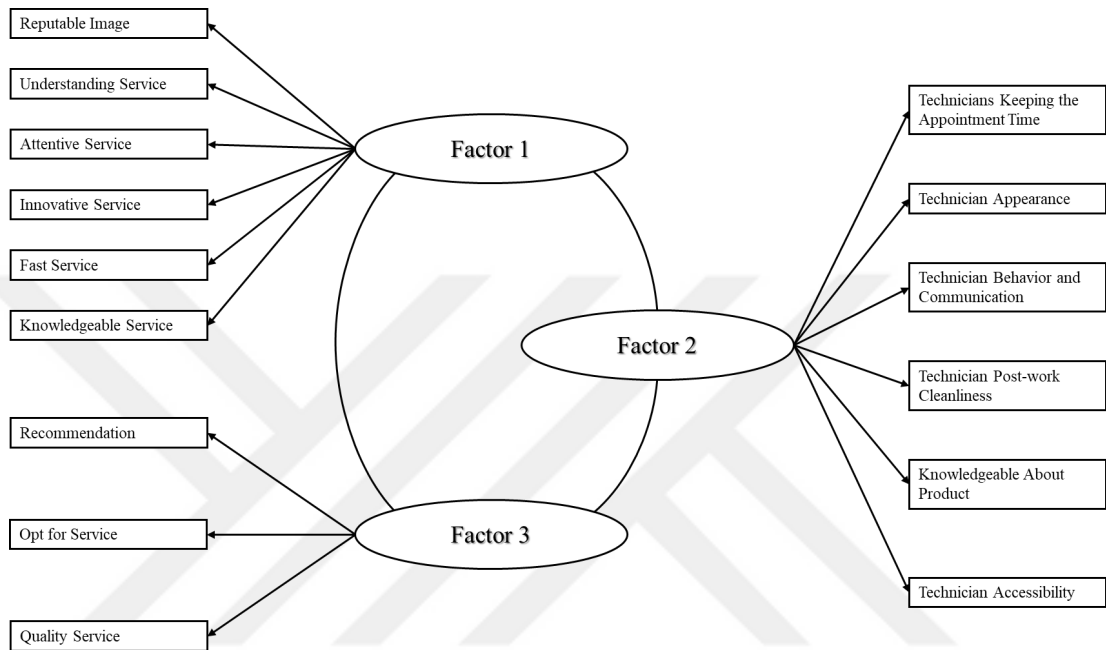
Note. '*' = Factor should be retained. Results from PC-based parallel analysis.

Factor Characteristics

	Unrotated solution				Rotated solution		
	Eigenvalues	SumSq. Loadings	Proportion var.	Cumulative	SumSq. Loadings	Proportion var.	Cumulative
Factor 1	7.520	7.176	0.478	0.478	3.942	0.263	0.263
Factor 2	1.831	1.481	0.099	0.577	3.243	0.216	0.479
Factor 3	1.164	0.849	0.057	0.634	2.321	0.155	0.634

APPENDIX 3. CFA Results for Installation/Setup Service

Based on the results of the EFA, the schematic representation of the CFA model is shown in the graph below. As can be seen from the presented tables, the model met expectations and fulfilled the conditions on the first attempt.



The chi-square table of CFA model does not validate the model (as expected since it never does when the sample size is high).

Model tests

Label	X ²	df	p
User Model	293	87	< .001
Baseline Model	33786	105	< .001

In the text CFI and TLI indexes that are based on chi-square test, developed to overcome its vulnerability to sample size have been discussed. Table below presents additional derived statistics. All indexes have the following cut-off criteria:

- **0.90 and above:** Excellent fit
- **0.80 - 0.89:** Good fit
- **0.70 - 0.79:** Acceptable fit
- **Below 0.70:** Poor fit

User model versus baseline model

	Model
CFI	0.994
TLI	0.993
NNFI	0.993
RNI	0.994
NFI	0.991
RFI	0.990
IFI	0.994
PNFI	0.821

In the CFA and SEM literature, the most used fit indices, which are also discussed in the text, are undoubtedly RMSEA and SRMR.

Fit indices

95% Confidence Intervals				
SRMR	RMSEA	Lower	Upper	RMSEA p
0.052	0.049	0.043	0.055	0.610

The following results were prepared to demonstrate the construct validity of the CFA model.

Construct validity has two components. The first component is convergent validity, and we report four tables to demonstrate that CFA model satisfies convergent validity.

The first table related to convergent validity is AVE table. Each factor's AVE value is expected to be greater than 0.50.

Factor	AVE
Factor 1	0.655
Factor 2	0.542
Factor 3	0.745

The second table presents the factor loadings, standard errors, 95% confidence intervals, and significance test scores of the CFA. As can be seen from the table, all parameters are statistically significant. For convergent validity, all loadings are expected to be greater than 0.7.

Latent	Observed	Estimate	SE	95% Confidence Intervals		z	p
				Lower	Upper		
Factor1	Reputable Image	0.796	0.0122	0.772	0.820	65.5	< .001
	Understanding Service	0.932	0.0112	0.910	0.954	83.3	< .001
	Attentive Service	0.931	0.0112	0.909	0.953	83.2	< .001
	Knowledgeable Service	0.778	0.0130	0.752	0.803	59.7	< .001
	Innovative Service	0.700	0.0132	0.674	0.726	53.1	< .001
	Fast Service	0.682	0.0143	0.654	0.710	47.9	< .001
Factor2	Technicians Keeping the Appointment Time	0.607	0.0170	0.574	0.640	35.8	< .001
	Technician Appearance	0.657	0.0147	0.628	0.686	44.6	< .001
	Technician Behavior and Communication	0.862	0.0144	0.834	0.890	59.9	< .001
	Knowledgeable About Product	0.795	0.0141	0.767	0.823	56.3	< .001
	Technician Post Work Cleanliness	0.637	0.0148	0.608	0.666	43.0	< .001
	Technician Accessibility	0.819	0.0146	0.790	0.848	55.9	< .001

Latent	Observed	Estimate	SE	95% Confidence Intervals		z	p
				Lower	Upper		
Factor3	Recommendation	0.856	0.0130	0.831	0.882	65.8	< .001
	Opt for Service	0.865	0.0139	0.837	0.892	62.2	< .001
	Quality Service	0.869	0.0136	0.842	0.895	63.9	< .001

The third table used for convergence validity is the table of R^2 values. R^2 values indicate the extent to which each measured variable is explained by the CFA model. Values are expected to be higher than 0.40.

R^2	
Variable	R^2
Reputable Image	0.634
Understanding Service	0.869
Attentive Service	0.867
Knowledgeable Service	0.605
Innovative Service	0.490
Fast Service	0.465
Technicians Keeping the Appointment Time	0.368
Technician Appearance	0.432
Technician Behavior and Communication	0.743
Knowledgeable About Product	0.632
Technician Post Work Cleanliness	0.406
Technician Accessibility	0.671
Recommendation	0.733
Opt for Service	0.748

R²

Variable	R²
Quality Service	0.755

The fourth and the final table for convergent validity is Cronbach's alpha and McDonald's omega.

Variable	α	Ordinal α	ω_1
Factor1	0.875	0.912	0.863
Factor2	0.777	0.870	0.798
Factor3	0.864	0.894	0.828

Factor Covariance Table

Variable 1	Variable 2	Estimate	SE	95% Confidence Intervals		z	p
				Lower	Upper		
Factor1	Factor2	0.632	0.0119	0.609	0.656	53.1	< .001
Factor1	Factor3	0.650	0.0127	0.625	0.675	51.3	< .001
Factor2	Factor3	0.703	0.0145	0.674	0.731	48.4	< .001

	AVE	Square of Factor Correlations
Factor 1	0.655	0.399
Factor 2	0.542	
Factor 1	0.655	0.423
Factor 3	0.745	
Factor 2	0.542	0.494
Factor 3	0.745	

	Index	Final Model	Constrained Model
	CFI	0.994	0.856
	TLI	0.993	0.832
	RMSEA	0.049	0.197
	RMSEA-90%-CI-lower-bound	0.045	0.192
	RMSEA-90%-CI-upper-bound	0.057	0.203
	RMSEA-p-value	0.419	0.000
	SRMR	0.052	0.173

APPENDIX 4. AHP Results

Pair-wise Matrix	Fast Service	Reputable Image	Understanding Service	Attentive Service	Knowledgeable Service	Innovative Service
Fast Service	1.00	1.25	1.40	1.09	0.86	1.65
Reputable Image	0.80	1.00	1.12	0.87	0.69	1.32
Understanding Service	0.71	0.89	1.00	0.78	0.61	1.18
Attentive Service	0.92	1.15	1.29	1.00	0.79	1.52
Knowledgeable Service	1.16	1.45	1.63	1.26	1.00	1.91
Innovative Service	0.61	0.76	0.85	0.66	0.52	1.00

Pair-wise Matrix	Technician Appearance	Technician Behavior and Communication	Technician Knowledge about Product	Technician Post-work Cleanliness	Technician Accessibility	Appointment Time
Technician Appearance	1.00	1.06	0.77	1.12	1.29	0.80
Technician Behavior and	0.94	1.00	0.72	1.06	1.22	0.75

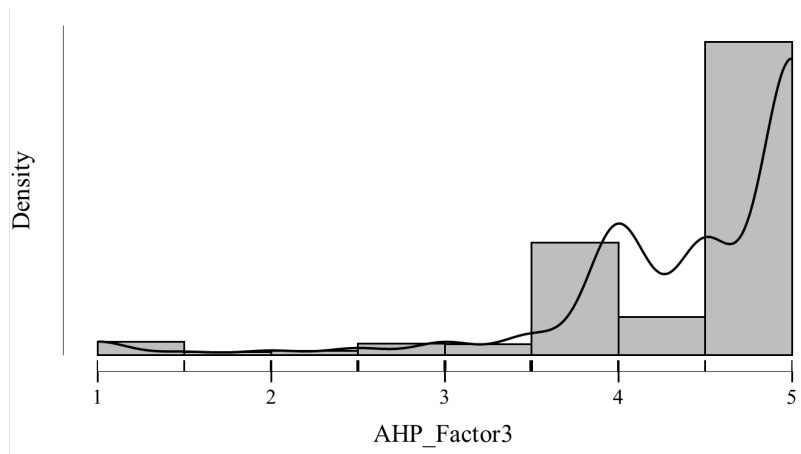
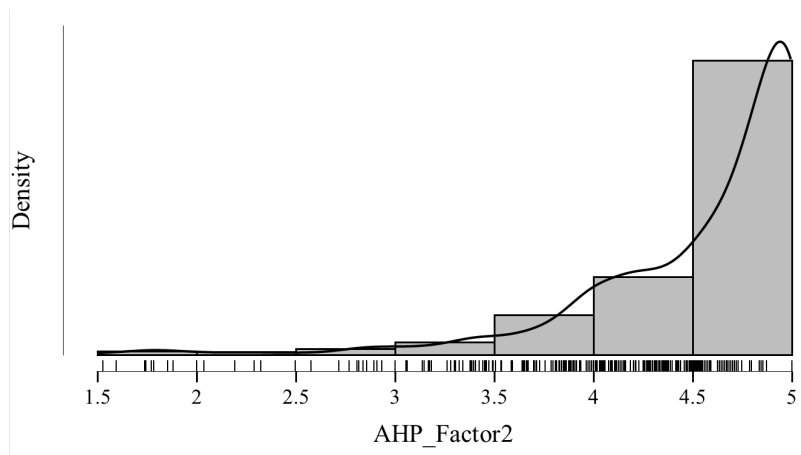
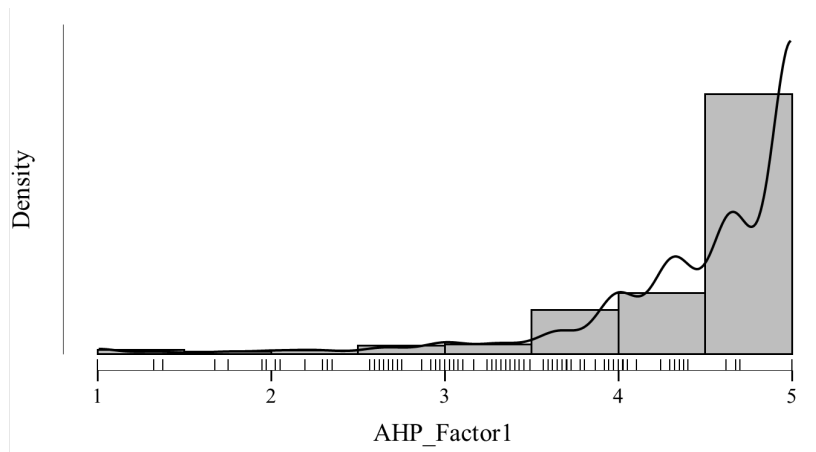
Communication						
Technician Knowledge about Product	1.30	1.38	1.00	1.46	1.68	1.04
Technician Post-work Cleanliness	0.89	0.94	0.68	1.00	1.15	0.71
Technician Accessibility	0.78	0.82	0.60	0.87	1.00	0.62
Appointment Time	1.26	1.33	0.96	1.41	1.62	1.00

Pair-wise Matrix	Quality Service	Recommendation	Opt for Service
Quality Service	1.00	0.70	0.93
Recommendation	1.42	1.00	1.32
Opt for Service	1.08	0.76	1.00

APPENDIX 5. GWR Results for Installation/Setup Service

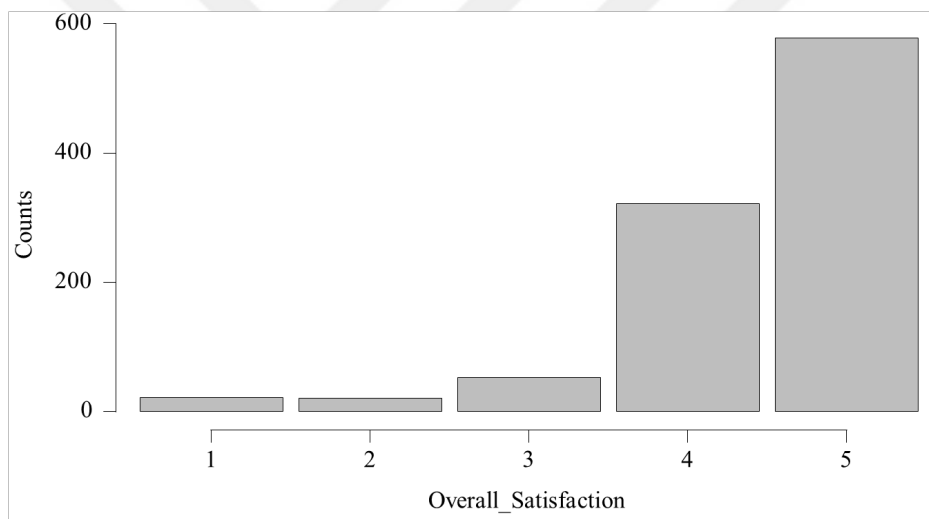
Descriptive Statistics

	AHP_Factor1	AHP_Factor2	AHP_Factor3
Valid	993	993	993
Missing	0	0	0
Median	4.676	4.789	4.502
Mean	4.530	4.555	4.441
Std. Deviation	0.681	0.568	0.813
Skewness	-2.416	-2.131	-2.310
Std. Error of Skewness	0.078	0.078	0.078
Kurtosis	7.429	5.950	6.360
Std. Error of Kurtosis	0.155	0.155	0.155
Shapiro-Wilk	0.705	0.763	0.686
P-value of Shapiro-Wilk	< .001	< .001	< .001



Descriptive Statistics	
Valid	993
Missing	0
Mode	5
Median	5
Mean	4.426
Std. Deviation	1.31

Overall Service Satisfaction	Frequency	%	Cumulative %
1	21	2.115	2.115
2	20	2.014	4.129
3	52	5.237	9.366
4	322	32.427	41.793
5	578	58.207	100



	Sum of Squares	df	Mean Square	F-stat	p-value
Regression	415.731	3	138.537	452.202	< 2.2e-16
Residual	303.079	989	0.306		
Total	718.810	992			

Model Summary - Overall_Satisfaction

Model	R	R ²	Adjusted R ²	RMSE
H ₀	0.761	0.578	0.577	0.554

Note. Null model includes AHP_Factor1, AHP_Factor2, AHP_Factor3



Coefficients

Model		Unstandardized	Standard Error	Standardized	t	p	Collinearity Statistics	
							Tolerance	VIF
Ho	(Intercept)	0.072	0.153		0.473	0.637		
	AHP_Factor1	2.471×10 ⁻⁴	0.031	1.977×10 ⁻⁴	0.008	0.994	0.690	1.449
	AHP_Factor2	0.312	0.038	0.208	8.133	< .001	0.653	1.531
	AHP_Factor3	0.661	0.027	0.631	24.834	< .001	0.660	1.514

