Improved importance–performance analysis: New ranking identification method for passenger satisfaction of urban rail transit

Zhu Shunying, Yu Cai (corresponding author), Wang Hong, Xu Wangtu

Commented [Ed16521]: Check for accuracy.

Abstract

Improvements in service quality and passenger satisfaction are important in enhancing the attractiveness of city rail transit and loyalty of passengers. Satisfaction improvement decisions and methods include the traditional importance–performance analysis (IPA) and the newly added three-factor method. However, these techniques can only determine the category precedence and do not rank all the elements in the category. The threshold value of the category division also remains debated. We keep the advantages of IPA method as consideration of index importance, short-plate theory, and classified sequence. We then introduce the "potential performance" (PP) concept and models, and a ranking method that is based on PP. Index ranking and classification can be implemented on the basis of the performance value and curve according to this method. The PP ranking method considers the potential, difficulty, and uncertainty of index improvement and simplifies the required multidimensional ranking into one-dimensional sequencing. Finally, the improvement strategies of Wuhan City for rail transit satisfaction in 2017 are regarded as an example. The improvement precedence of the satisfaction index is divided into three bands through PP: improvement items in dire need, secondary need, and with no need, all of which should be compared and analyzed with IPA results. Discovery of new methods can sequence and classify the indexes and PP value and indicate several differences with IPA results. Such disparitiesresult from the substantial consideration of ranking influential factors, which yields credible, explanatory classification and ranking results.

Commented [Ed16522]: The en dash (–) is used to link words that represent items of equal rank. To ensure consistency, this rule was applied throughout the file.

Commented [Ed16523]: By default, the language setting of the entire file was set to English (United States), and the conventions in editing a US English file were followed.

Key Words: Rail transit; service quality; passenger satisfaction; improvement strategy sequencing; potential performance; structural equation modeling (SEM); importance–performance analysis (IPA); three-factor theory

Commented [Ed17304]: Adjectival nouns (those that are used to modify other nouns) are commonly written in singular form.

We searched Google and Google Scholar with the keywords *improvement strategies sequencing* and *improvement strategy sequencing* and failed to obtain a single match.

Commented [Ed16525]: The common format of key words is as follows: aa, bb, cc (lowercase except for proper nouns). Verify the preferred format of your target journal.

1. Introduction

Public transportation plays a significant role in alleviating traffic congestion and reducing carbon emissions. To improve the attractiveness of public transportation and maintain passenger capacity, transportation departments should study passenger satisfaction and improve public transportation services. Research on passenger feedback regarding the entire service experience is increasingly important in assessing, measuring, and improving the service quality of public transportation and is emphasized by the public transportation service industry. Academic literature on passenger satisfaction mainly focuses on the following aspects: (1) positioning of passenger demands and expectations through passenger satisfaction (Shaaban and

Kim 2016), (2) compilation of service benchmark satisfied by passengers (Freitas 2013), (3) tracing of changes in passenger demands (De Vos and Witlox 2017), (4) identification of key indicators affecting passenger satisfaction (Mouwen 2015; Susilo and Cats 2014), (5) update of passenger satisfaction measurement methods (Nassereddine and Eskandari 2017; Rahman et al. 2016; Stradling et al. 2007), and (6) determination of indicator performances of satisfaction factors and research on improvement strategies (Wu et al. 2018).

Importance–performance analysis (IPA) is a dominant method in determining the precedence of service improvements in customer satisfaction research due to its conciseness and availability. Many scholars have used the IPA method to analyze survey results of passenger satisfaction and determine improvement directions. However, the IPA method also has limitations. One outstanding problem is that it is based on the assumption that the relationship between the behavior of individual properties and overall satisfaction is linear and symmetrical. These limitations can be solved by three-factor theory, which classifies the perception and transmission property of service quality into basic, performance, and exciting factors. Three-factory theory can improve the shortcomings of the IPA method in classification to a certain degree. However, few factors are considered in the classification process by the IPA and three-factor methods. The classification threshold value is still debated, and class elements still cannot determine the precedence.

This passage creates a ranking method that is based on the “potential performance” (PP) to determine the precedence of the service improvement indicators in passenger satisfaction and overcome the limitations of the IPA and three-factor methods. PP considers the main decision factors, especially uncertainty, to improve performance (represented by the variation coefficient of an indicator) through the general consideration of passenger review and structural equation modeling (SEM), and the model initiated in this research can reveal the PP value of each indicator. Precedence ranking by PP value, a one-dimensional variable, renders the method simple and easily operatable.

The rest of this paper is structured as follows. Chapter 2 reviews the analysis method of the indicator performance of passenger satisfaction factors. Chapter 3 introduces the data sources and analysis method. Chapter 4 analyzes and discusses the research results. Chapter 5 presents the key research conclusion and directions for future research.

2. Literature review

In the research related to passenger satisfaction, improvement strategies of passenger satisfaction are a crucial objective, and ranking the improvement precedence of each factor indicator/strategy is always the focus of studies on such strategies. An early ranking method is the linear weighting method. This technique can indicate indicator importance by the analytic hierarchy process (AHP) (Golden et al. 1989), through which each indicator’s weight can be acquired. Indicator importance can be quantitatively calculated through the hierarchy and quantification of each indicator according to mindset and psychological rules. AHP is mainly divided into five steps: (1) establishment of a hierarchical structural model; (2) construction of a judgment matrix; (3) single hierarchical arrangement; (4) consistency inspection; and (5) total hierarchical arrangement. The three final steps should be implemented in order in the whole process.

The relative importance of each indicator is determined by the expert score in AHP; hence, experts should judge the importance of the indicators according to their personal experiences. When the indicators are numerous and determining the relative importance is consequenctly difficult, guaranteeing the consistency of each hierarchy's ranking and the total ranking is challenging. The weight yielded by the AHP method is the experts' judgment of each indicator's importance. Although the method is systematic, experts' judgment cannot represent the "passenger-oriented" service concept due to the limited number of experts.

The IPA method has been widely adopted in all service sectors due to its intuitiveness and operability since it was initiated in 1977 by Martilla and James (Martilla and James 1977). It covers the service sector's satisfaction, product performance, macro-tourism policy formulation, and education industry, among others. Ennew et al. (Ennew et al. 1993) created an improvement strategy of service quality for small companies of English banks with the aid of the IPA method. Hsu et al. (Hsu et al. 1998) developed an improvement strategy for Korean university students' satisfaction with fast food services, family restaurants, and fancy restaurants. Evans and Chon (Evans and Chon 1989) discussed the key elements of IPA technology in tourism policy formulation and assessment and explored the method's applicability through two cases of adults and teenagers. O Neill and Palmer (O Neill and Palmer 2004) stated the service quality assessment problems of higher education departments and explained that IPA technology can determine specific defects and evaluate its importance to the quality improvement plan.

IPA constructs four quadrants on the basis of the importance and passenger rating of each factor indicator. All the indicators are distributed among the four quadrants with different improvement precedence. The first quadrant is the "advantage-impacted" region, in which each indicator is of great importance and has good service performance. The service indicators in this region need to be maintained by a company. The second quadrant is the "status quo-held" region, in which each indicator is of low importance and has good service performance. Added energy is not needed in this region. The third quadrant is the "second improvement" region, in which each indicator is of low importance and has bad service performance. The indicators in this region need to be improved but not urgently. The fourth quadrant is the "prior improvement" region, in which each indicator is of high importance and has bad service performance. The indicators in this region need to be emphasized in the service quality improvement.

Many scholars have applied the IPA method into research cases about transportation service satisfaction. Stradling et al. (Stradling et al. 2007) used the

Commented [Ed16526]: Check for accuracy.

Commented [Ed16527]: Do not use ampersands (&) in the text, except if they are part of a proper name.

Commented [Ed17308]: Verify if *small English banks* will do.

Commented [Ed16529]: "Respectively" shows parallelism; thus, it should never be used if there is no one-to-one correspondence.

Commented [Ed165210]: Verify if the revision expresses your intended meaning accurately.

Commented [Ed173011]: Check for accuracy.

six-step IPA method that they proposed to analyze the focusing differences of bus passengers, car passengers, and footers in the road transportation process. Zhang et al. used examples of bus, bus rapid transit, and van riders in Indore, India, to determine the focusing differences of the passengers of the three transportation vehicles to enable improvement directions to be determined to improve public transportation quality.

Commented [Ed173012]: Add citation.

The importance of factor indicators in the IPA method can be determined in many ways. Weight in the early IPA method is generally determined by multiple regression or correlation methods. Weinstein (Weinstein 2000) regarded the analysis results of Pearson multiple regression and correlation method as the importance of factor indicators in the IPA method and analyzed passenger satisfaction data in the San Francisco Bay Area Rapid Transit District to compare the two results. As satisfaction research increasingly deepens, the measurement indicator system gradually expands. Indicators begin to affect one another, thereby hardly satisfying the independence requirements. In addition, multiple-indicator regression can barely pass statistics inspection, and multiple regression method is being gradually replaced by SEM (Stuart et al. 2000). SEM originated from the traditional factor analysis and is based on the variable covariance matrix that analyzes variable relationships. A measurable variable is used to indirectly measure certain variables that cannot be directly measured, which are called potential variables. Traditional factor analysis requires that the potential factors should be independent of one another. However, in actual problems, some potential factors have independent or dependent relationships. SEM can describe such relationships well and thus make the model realistic through the introduction of exogenous variables, endogenous variables, and error terms. Chou et al. (Chou et al. 2011) used SEM to determine the importance of service indicators and adopted the IPA method to analyze and assess the service quality and improvement strategies of high-speed railways in China, Taiwan, and Korea. The weight of factor indicators determined by SEM renders the IPA method objective and perfect. Comparatively complete methods and systems are already formed in actual applications. Shen et al. (Shen et al. 2016) adopted SEM to analyze a passenger satisfaction survey of the city railway between Suzhou and Shanghai and used the IPA method to analyze improvement directions.

Commented [Ed165213]: Check if this is what you mean.

Commented [Ed173014]: You may have to specify which Korea.

The IPA method has two presuppositions. (1) The importance of each indicator must be independent of its satisfaction evaluation score, and (2) the satisfaction scores of all sub-indicators must be linearly independent of the overall satisfaction. However, due to the limitations of survey sampling and implementation, the two premises can hardly be fulfilled. In a post-study research, Kano et al. (Kano 1984) stated that the relationship between indicators' satisfaction scores and overall satisfaction is sometimes asymmetric. The researchers presented a query to customers about the practical evaluation of each indicator by designing the "Kano model survey." Customer demands of each indicator were analyzed and then divided into five types. Subsequently, many researchers (Anderson and Mittal 2000; Busacca and Padula 2005; Matzler et al. 2004) have applied the Kano model to analyze customer satisfaction. Such research contributed to the emergence of the three-factor theory, which simplifies the five-type satisfaction indicators in the Kano model into three-type indicators: basic, performance, and exciting factors. Basic factors remarkably influence the overall satisfaction when the satisfaction level is low, but they influence the overall satisfaction only slightly when the satisfaction level is high. Performance factors pose the same influence on the overall satisfaction when the satisfaction level is either high or low. Exciting factors affect the overall satisfaction only slightly when the satisfaction level is low and substantially when the satisfaction level is high (Wu et al. 2018). Three-factor theory adds the perspective of the property features of the factor indicators to the IPA category sequencing, which has a certain practical value and is mutually complementary and supported by the IPA method. However, the three-factor method considers only the importance of factors trending from low satisfaction to high satisfaction and average passenger evaluation values of different indicators. Few factors are taken into account, and the category classification threshold value is debated because the IPA method encounters difficulty in sequencing all the factors in one category. This technique is therefore limited.

Commented [Ed173015]: Check for accuracy.

The IPA method and three-factor theory are widely used in the customer satisfaction field and state the improvement directions of the service quality concerning the key factor classification. However, the IPA method still has limitations that are hard to overcome. Enterprise managers and policy makers can easily distinguish the indicators in different quadrants through the IPA method and then formulate development directions and strategies through the classification precedence of each indicator. However, indicator precedence in the same quadrant cannot be distinguished. In addition, quadrant boundaries are divided according to the average value of satisfaction and importance. Although it is arbitrary, the theoretical basis is insufficient. Thus, the current work determines another means of overcoming the inherent defects of the IPA method, adds and considers the influential improvement factors, develops a PP-based ranking method of improvement strategies, analyzes the results of a 2017 passenger satisfaction survey concerning Wuhan rail transit, and formulates a new ranking method for improvement strategy precedence.

Commented [Ed165216]: Specify what “it” is referring to.

3. Methodology

3.1. Data and variables

Data in this passage were collected from a survey executed in July 2017 about passenger satisfaction with Wuhan rail transit. Simple random sampling was implemented among the passengers in each transit line from the five open rail transit lines (metro lines 1, 2, 3, 4, and 6 of Wuhan) by the research team along with Wuhan Railway Operation Company. The selected passengers were presented with a paper questionnaire through one-to-one face-to-face communication. The survey was administered for two weeks, including weekdays and weekends. The daily survey time covered the rail operation times.

Commented [Ed165217]: We used the past tense in presenting result sentences in the paper for conciseness.

A total of 25,000 questionnaires were distributed in this survey, of which 23,180 were valid (Line 1: 4,238, Line 2: 5,308, Line 3: 4,267, Line 4: 5,156, and Line 6:

4,211), and the effective questionnaire recovery rate was 92.7%. The questionnaire included numerous passenger features that distinguish different passenger groups. The passenger group sample that accepted the survey is shown in Table 1. The questionnaire also contained 12 investigation attributes, including one overall satisfaction indicator and 11 factor indicators. The 11 factor indicators covered safety, comfort, convenience, reliability, and rapidity. Moreover, the 11 factor indicators were the service quality component elements in the SERVQUAL (Cavana et al. 2007) model, combined related opinion from industrial and management experts and staff from Wuhan rail transit, and were determined after small-scale adjustments and modifications. Factor indicators that belong to service quality but do not belong to the service provided by the operation enterprises were deleted, such as prices, expectations, and ideals, which are managed by government departments. Such simplification poses no influence on the research of ranking methods.

Commented [Ed173018]: Check for accuracy.

In the survey, the interviewees needed to evaluate the 12 investigation indicators and select one out of five satisfaction levels ("very satisfied," "satisfied," "basically satisfied," "dissatisfied," and "very disappointed"). Indicator quantization, which was conducted after the questionnaires were collected, used a universal five-point Likert scale ("very disappointed" = 1, "dissatisfied" = 2, "basically satisfied" = 3, "satisfied" = 4, and "very satisfied" = 5). The higher the value, the higher the passenger satisfaction.

Table 1 Statistical features of passenger interviewees

Classification standard		Number	Percentage (%)
Gender	Male	12902	55.66
	Female	10278	44.34
Age	Under 18	2568	11.08
	18–35 years old	13849	59.75
	36–50 years old	5041	21.75
	51–60 years old	896	3.87
	61–65 years old	420	1.81
	Above 65	406	1.75
Permanent	Yes	16681	71.96

resident	No	6499	28.04
Occupation	Student	4896	21.12
	Civil servant	1660	7.16
	Enterprise and public institution employee	9052	39.05
	Retired staff	1504	6.49
	Freelance worker	4181	18.04
	Others	1887	8.14
Trip purpose	Commute	5302	22.87
	Going to and from school	2765	11.93
	Public affairs	4349	18.76
	Shopping and recreation	5024	21.67
	Others	5740	24.76
Revenue	Below 2000 yuan	4774	20.60
	2000–2999 yuan	4799	20.70
	3000–3999 yuan	4496	19.40
	4000–4999 yuan	3932	16.96
	Above 5000 yuan	5179	22.34

3.2. Modeling approach

Questionnaire credibility and validity were inspected through SPSS 24.0 before data analysis. The Cronbach’s α of the questionnaire is 0.900, which indicates that the inner consistency of all indicators in the survey is relatively good.

This research first calibrated each indicator of the survey results through SEM. Path coefficient in the SEM analysis result was regarded as the weight of each indicator. IPA and PP analysis were implemented on the basis of the weight and score of each indicator, respectively. Differences of the two results were analyzed and noted.

3.2.1. SEM

SEM is a common method that deals with interactive causal relationships and influence degrees between multiple variables that cannot be directly measured. It is composed of a structural model (1), a measurement model (2), and (3).

Commented [Ed165219]: If possible, specify what the third model is.

$$\eta = B\eta + \Gamma\xi + \zeta \tag{1}$$

$$X = \Lambda_x \xi + \delta \tag{2}$$

$$Y = \Lambda_y \eta + \varepsilon \tag{3}$$

In the equation, η is an endogenous potential variable; ξ is an exogenous potential variable; ζ is a random distracter, reflecting the part that η cannot explain in Equation (1); B is the coefficient matrix of η, describing the influence of η on another variable; Γ is the coefficient matrix of ξ, describing the influence of exogenous potential variable η on one another; X is the observational indicator of ξ; Y is the observational indicator of η; δ is the measurement error of X; ε is the measurement error of Y; Λ_x is the factor loading of X; and Λ_y is the factor loading of Y.

Commented [Ed165220]: One-letter symbols (normal, in subscript, or superscript) representing numerical values (mathematical or statistical variables) are generally italicized.

Commented [Ed173021]: Double-check.

On the basis of the American Customer Satisfaction Index (Angelova and Zekiri 2011), together with the operation service features of urban rail transit companies, six potential variables were classified in this work, which were concluded from 12 factors, including overall satisfaction, in which security sense is the exogenous variable and quality awareness, convenience, comfort, reliability, and quickness were the endogenous variables. The classification of the measurable and potential variables and the satisfaction score of each indicator are shown in Table 2.

Table 2 Corresponding measurable and potential variables in SEM

Potential variable	Observational variable	Score	Score ranking	Weight
ξ1: Security sense	x1: Safety	3.77	1	0.195
η1: Quality awareness	y1: Overall satisfaction	3.53	/	/
η2: Convenience	y2: Transfer shuttle	3.54	8	0.068
	y3: Ticketing system	3.54	9	0.078
η3:	y4: Train	3.61	5	0.046

Comfortable sense	comfortability			
	y5: Platform comfortability of the station hall	3.60	6	0.047
	y6: Personnel service	3.64	4	0.039
	y7: Passenger civilization	3.34	10	0.032
η4: Reliability	y8: Service facility	3.58	7	0.188
	y9: Train information	3.66	3	0.181
η5: Quickness	y10: Peak waiting time	3.24	11	0.057
	y11: Normal waiting time	3.68	2	0.068

3.2.2 Ranking method based on PP

The PP-based importance ranking method was created to compensate for the defects of the IPA method, which can only classify and sequence each improvement strategy but cannot easily determine the shortcomings of classification boundaries. This study assigns a variable as the PP, thereby maintaining the inherent advantages of barrel theory (also called short-plate theory) in the IPA method concerning the consideration of indicator importance and improvement urgency and adding consideration indicators and influential improvement factors to represent the precedence of each improvement strategy. The model is shown in Equations (4) and (5).

$$E_k = \frac{I_k \cdot (L_{max} - P_k)}{d}, \qquad (4)$$

$$\mathrm{d} = \frac{\sigma_k}{P_k}, \qquad (5)$$

Commented [Ed165222]: Equations must also be punctuated in the same way as ordinary text. Always place a comma after an equation followed by “where” (introduction of variables).

where E_k is the PP of the improvement indicator (strategy). The larger the E_k value, the greater the potential improvement performance of the corresponding strategy and the higher the precedence. I_k is the function coefficient of each indicator, including direct and indirect functions, which can be gained from SEM. The larger the function coefficient, the greater the indicator importance. P_k is the average value of each indicator’s passenger satisfaction, ranging from 1 to L_{max}, in which L_{max} is the maximum value of the Likert scale. L_{max}–P_k denotes the improvement potential, urgency, and difficulty. According to barrel theory, the larger the value, the shorter the short plate. Thus, improvement is urgent. The larger the value, the higher the potential.

A small value indicates that the indicator has a high score, is close to the maximum value, and has substantial improvement difficulty. The higher the passenger's score on the indicators, which is close to the full score, the more difficult the improvement of the score. *d* represents the improvement uncertainty in the form of the indicator's variation coefficient, which is the ratio of the standard deviation σ_k to the average satisfaction score P_k of the indicator's service quality perception. The larger the variation coefficient, the larger the uncertainty of the improvement effect. σ_k is the discrete degree of the passengers' subjective understanding. A large value means that passengers do not reach a consensus on such indicator's satisfaction and the group's uncertainty is significant. A denominator is used to eliminate the differences between and increase the comparability of various indicators.

3.2.3 Comparative analysis between IPA and PP methods

Comparative analysis was performed between the IPA and PP methods to identify their advantages and shortcomings and verify whether the PP method overcomes the defects of IPA. The adopted analysis method was oriented by the model result differences of the same data. The analysis data used the 2017 satisfaction data of Wuhan rail transit. The factor indicators' weights under the PP and IPA methods were mainly determined by SEM.

Model results adopted the performance value predicted by SEM and the PP value established in this research.

4. Results and analysis

4.1. Factor weights

SEM was established on the basis of the relationship of each indicator and the corresponding potential variables presented in Table 2. Partial least squares estimators were adopted, and parameter estimation and effect inspection were implemented through the software IBM SPSS Amos. SEM standardized estimation was then executed, as shown in Table 1.

In the SEM, the importance of each factor indicator was concluded through the combination and normalization of the relationship between each factor and overall passenger satisfaction and the consideration of direct and indirect impacts.

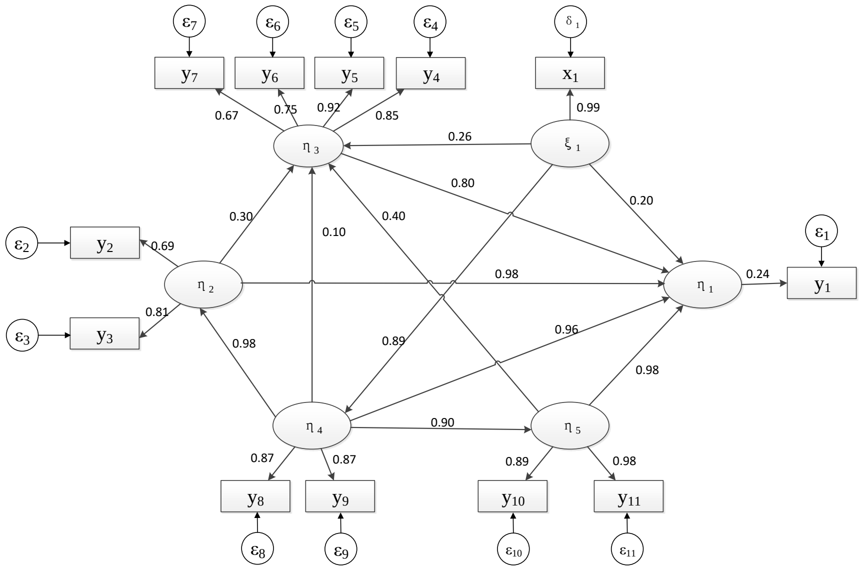

Table 1 SEM parameter estimation results

The matrix form of the SEM path graph can be represented as follows:

$$\begin{matrix}\eta_1\\ \eta_2\\ \eta_3\\ \eta_4\\ \eta_5\end{matrix} = \begin{bmatrix} & 0.98 & 0.8 & 0.96 & 0.98 \\ & & & 0.98 & \\ & 0.3 & & 0.1 & 0.4 \\ & & & & \\ & & & 0.9 & \end{bmatrix}\begin{bmatrix}\eta_1\\ \eta_2\\ \eta_3\\ \eta_4\\ \eta_5\end{bmatrix} + \begin{bmatrix}0.2\\ \\ 0.26\\ 0.89\\ \\ \end{bmatrix}\xi_1 + \begin{bmatrix}\zeta_1\\ \zeta_2\\ \zeta_3\\ \zeta_4\\ \zeta_5\end{bmatrix} \quad , \tag{6}$$

$$x_1 = 0.99\xi_1 + \delta_1, \tag{7}$$

$$\begin{matrix}y_1\\ y_2\\ y_3\\ y_4\\ y_5\\ y_6\\ y_7\\ y_8\\ y_9\\ y_{10}\\ y_{11}\end{matrix} = \begin{bmatrix}0.24 & & & & \\ & 0.69 & & & \\ & 0.81 & & & \\ & & 0.85 & & \\ & & 0.92 & & \\ & & 0.75 & & \\ & & 0.67 & & \\ & & & 0.87 & \\ & & & 0.87 & \\ & & & & 0.89 \\ & & & & 0.98\end{bmatrix}\begin{bmatrix}\eta_1\\ \eta_2\\ \eta_3\\ \eta_4\\ \eta_5\end{bmatrix} + \begin{bmatrix}\varepsilon_1\\ \varepsilon_2\\ \varepsilon_3\\ \varepsilon_4\\ \varepsilon_5\\ \varepsilon_6\\ \varepsilon_7\\ \varepsilon_8\\ \varepsilon_9\\ \varepsilon_{10}\\ \varepsilon_{11}\end{bmatrix}. \tag{8}$$

Through the combination of Equations (6), (7), and (8) without consideration for errors, the overall satisfaction y_1 can be represented by x_1 and y_i in many expressions, among which those with numerous observational variables are expressed in Equation (9).

$$y_1 = 0.24(1.29y_2 + 0.94y_4 + 1.1y_8 + 1.1y_{10} + 0.2x_1) \tag{9}$$

The fit index of SEM is an important indicator in evaluating the fitness of a model and data. Five indicators, namely, goodness-of-fit index (GFI), root mean square residual (RMR), adjusted goodness-of-fit index (AGFI), normed fit index (NFI), and relative fit index, were adopted to evaluate the model fitness with reference to relevant research (Browne and Cudeck 1993; Hooper et al. 2008; Steiger 1990; Vandenberg and Lance 2000). The results are shown in Table 3.

Commented [Ed165223]: Check for correctness.

Table 3 SEM model fitness indicator and inspection

Commented [Ed165224]: We followed the dominant format used by the authors in presenting tables in the paper.

Indicator	GFI	RMR	AGFI	NFI	IFI	CFI
Results	0.965	0.020	0.948	0.941	0.942	0.942
Judging standards	>0.90	<0.05	>0.90	>0.90	>0.90	>0.90

As shown in Table 3, the five model fitness indicators all satisfy the evaluation standards. Therefore, such model and observational data are important.

4.2. Strategic ranking and analysis based on PP

Commented [Ed165225]: We observed inconsistencies in the format of your headings/subheadings. We followed the dominant format. Verify with the style guide of your target journal.

In the IPA method, each satisfaction indicator is distributed in different quadrants, which can serve as the classification basis of service quality improvement strategies. However, the improvement precedence of each indicator in the same quadrant cannot be evaluated through the IPA method. The PP-based improvement strategy ranking method proposed in this work is based on the average evaluation value of the satisfaction of each indicator and proposes the PP concept. The PP value is determined through the importance and correlations of each indicator. The evaluation standards of a single composite variable are also determined. The PP analysis figure shown in Figure 4 was established with the PP value of each satisfaction indicator ranking from the highest to the lowest for an enhanced comparison with the IPA results. We divided the indicator into three sections through the inflexion of the potential utility value curve.∗

Sections I (three indicators included), II (four indicators included), and III were named “improvements in dire need,” ”improvements in secondary need,” and ”improvements with no need.”

Table 4 PP calculation of satisfaction indicators

Satisfaction indicator	Importance	Score	Standard deviation	Variable coefficients	PP

Service facility	0.188	3.58	0.721	0.201	1.325
Safety	0.195	3.77	0.726	0.193	1.241
Train information	0.181	3.66	0.750	0.205	1.182
Ticketing system	0.078	3.54	0.750	0.212	0.541
Transfer shuttle	0.068	3.54	0.745	0.211	0.469
Normal waiting time	0.068	3.68	0.788	0.214	0.423
Peak waiting time	0.057	3.24	0.865	0.267	0.378
Passenger comfort in train	0.046	3.61	0.696	0.193	0.332
Passenger comfort in station hall platform and passage	0.047	3.60	0.740	0.206	0.321
Personnel service	0.039	3.64	0.724	0.199	0.269
Passenger civilization	0.032	3.34	0.798	0.239	0.224

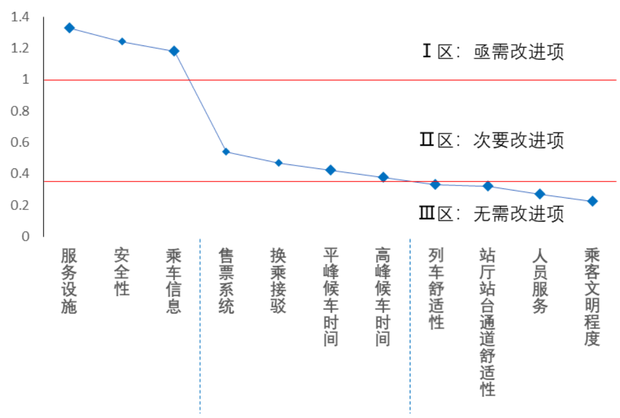

Figure 2 PP analysis of rail transit improvement strategies

4.2. Results and analysis of classification for IPA

The IPA figure was drawn through the general consideration of the average score value of each indicator and the weight gained from SEM, as shown in Figure 3.

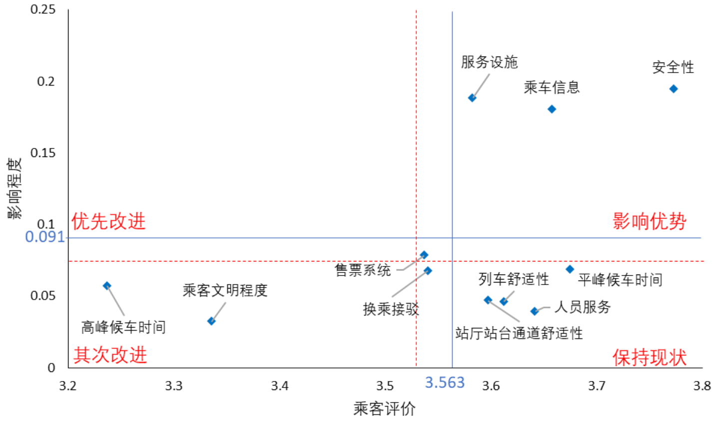

Figure 3 IPA of rail transit improvement strategies

In the IPA figure, 11 satisfaction indicators are divided into four quadrants, with the boundaries being the average score of the indicators (3.563) and the average importance of the indicators (0.091). Three indicators are in the advantage-impacted quadrant, and four indicators are in the status quo-held and secondary improvement quadrants. No indicator is in the prior improvement quadrant.

The IPA method establishes a reference system according to the indicator scores and weight and divides the quadrants according to the average values in a simple and clear manner. However, imprecise loopholes are present, wherein indicators in the same quadrant cannot be ranked and the average value serves as the boundaries.

First, indicators in the same quadrant cannot be ranked due to the unclear indicator scores and weight precedence without composite rules. This situation is explained by the “secondary improvement” in Figure 3. This quadrant has four indicators, namely, “peak waiting time,” “passenger civilization,” “ticketing system,” and “transfer shuttle,” all of which need to be improved in the services of rail operation units. However, with limited funds and resources that can improve only part of the indicators, enhancing overall service quality to satisfy passengers as much as possible while minimizing spending is a challenge. Among the four indicators, whether the utmost improvement precedence should emphasize “passenger civilization” on the basis of service score (lowest service score) or “ticketing system” on the basis of indicator weight (highest indicator weight) Must be decided. An arbitrary assignment of the improvement precedence to one out of the four indicators is impossible; others can be temporarily ignored in this process because such behavior is neither precise nor responsible.

Another shortcoming of IPA is the use of the average value as the boundaries. The division of the quadrants by the average value is subjective and lacks persuasiveness. The adoption of the average value as the boundaries raises the question of whether the median, geometric mean, harmonic mean, or any other one

Commented [Ed173026]: A substantial revision was made here. Check for accuracy.

Commented [Ed173027]: Specify what *others* refers to.

Commented [Ed173028]: *Responsible* is not commonly used to describe a process.

Commented [Ed173029]: Replace with the referent for clarity and formality.

can also become the basis for boundaries. Such division basis will lead to an entirely different classification, and the accuracy of such division basis is undetermined. With Figure 3 as an example, we slightly lower the boundaries of the indicator scores and indicator weights. “Ticketing system,” which originally belonged to the indicator needing to be improved, is now under the “advantage-impacted” indicator (red dash in Figure 3).

4.4 Comparative analysis between PP and IPA

As shown in Figure 2, the 11 factor indicators can be divided into three sections according to their respective inflection points of the PP curve. Each factor has its own PP value ranking from the largest to the smallest, thereby overcoming the shortcomings stated in Chapter 4.3. The result differences of the two methods are further analyzed in the following sections.

The first difference between the IPA and PP method results is that the three indicators with the utmost improvement precedence in the latter method, namely, “service facility,” “safety,” and “train information,” are classified by IPA in the “advantage-impacted” quadrant.

This disparity and the classification division are analyzed in the following section. In China, satisfaction surveys are administered every day to diagnose service levels. A short survey analysis period leads to small improvement effects on passenger satisfaction. In the short term, the interactive relationship between one indicator’s importance and its score can be unchanged. Suppose that the improvement goal is to increase satisfaction, which is analyzed with the examples of “safety” and “ticketing system” indicators. Suppose that the PPs of improvement are E_1 and E_2, and set ΔP_1, $\Delta P_2>0$ to represent their increase impacts. Assume that the improvement impacts ΔP_1 and ΔP_2 are few and equal.

According to Equation (9), the overall impacts after improvement $I_1 \cdot (P_1 + \Delta P_1)$ and $I_2 \cdot (P_2 + \Delta P_2)$ can be compared.

The following can be gained through data substitution:

$$I_1 \cdot (P_1 + \Delta P_1) - I_2 \cdot (P_2 + \Delta P_2) = 0.459 + 0.117\Delta P_1. \qquad (10)$$

The satisfaction of improving “safety” is higher than that of improving “ticketing system.” Therefore, “safety” cannot be included in the “advantage-impacted” category.

According to Equations (4) and (5), the PP values are further compared.

$$E_1\text{-}E_2 = \frac{0.195(5\text{-}3.77\text{-}\Delta P_1)}{0.193} - \frac{0.078(5\text{-}3.54\text{-}\Delta P_1)}{0.212} = 0.7055 - 0.6425\Delta P_1 \qquad (11)$$

When $\Delta P_1 < 1.0981$ (Setting the improvement value to be over 1 is generally difficult.), $E_1-E_2>0$ and the PP value of “safety” is higher than that of “ticketing system.” Correspondingly, the improvement precedence is increased, and “safety” should be concluded in the “prior improvement” category. Therefore, division according to the PP-based method is more reasonable than that according to IPA.

According to the regulations in IPA theory, the “secondary improvement” quadrant is of relatively higher improvement precedence than the “secondary improvement” and “status quo-held” quadrants, which both have relative weights, due to the lower indicator scores of the former. In the PP-based method, “ticketing system,” “transfer shuttle,” and “peak waiting time,” which are in the “secondary improvement” quadrant in IPA, are of higher improvement precedence than “passenger comfort in train,” “passenger comfort in station hall platform and passage,” and “personnel service,” which are in the “status quo-held” quadrant in IPA. These findings are consistent between the two methods.

In IPA, the “passenger civilization” indicator is in the secondary improvement quadrant and “normal waiting time” indicator is in the status quo-held quadrant. The former is of higher precedence than the latter according to IPA regulations. By contrast, in the PP-based method, the improvement precedence is reversed, which is the second difference between the new method and the IPA method. Similar to the above analysis on “safety” and “ticketing system,” the classification of “passenger civilization” and “normal waiting time” is more reasonable in the PP-based method than in IPA. According to the PP definition model, such difference can be well explained. Compared with “passenger civilization,” “normal waiting time” has a lower indicator variable coefficient, which can mean that passengers reach a higher consensus on improvement of “normal waiting time” with the small uncertainty of improvement impacts. In addition, unlike the improvement of “normal waiting time,” that of “passenger civilization” should be realized through joint efforts of the operation party and passengers, which is difficult to achieve.

5. Conclusions

The IPA method, which is commonly used in the service sector for determining improvement strategies for customer satisfaction, has the following inherent shortcomings. The technique determines the category precedence but not the element precedence in the category; category classification has only a few consideration elements; and the threshold value concerning the category classification is uncertain. To overcome these limitations, we propose the PP concept and model, research on the improvement strategies and decisions of rail transit passenger satisfaction, and compare the proposed technique with the IPA method. The PP-based method not only overcomes the inherent shortcomings of the IPA method but also adds other information needed for decisions. It remains strong explanation and rationality on the result inconsistencies with the IPA method. The ranking results obtained for the 2017 Wuhan rail transit improvement strategies can serve as reference for operation companies in enhancing service quality.

Determining the improvement strategies’ precedence of passenger satisfaction

Commented [Ed173030]: We added this to support the comparison introduced by *more*. Check for accuracy.

Commented [Ed173031]: Check the revised version carefully.

Commented [Ed173032]: The meaning is not clear. Suggested revision:

Discrepancies between the results of the proposed and IPA methods show that the former maintains strong explanation and rationality.

Strong explanation is vague in meaning as well. Consider rewording.

and understanding the potential impacts of improvements are the main components of improvement decisions. Numerous factors influence satisfaction. Improving all factors within a short period is impossible. Limited funds, human resources, and capital should be used widely. A decision maker should not only be aware of each improvement strategy's precedence before decision but also acknowledge the potential impacts of each improvement strategy and their differences. The traditional IPA method and the new three-factor method can show decision makers the precedence of strategy categories but not that of all strategies in a category. The PP method can not only directly show the precedence of each category and factor indicator but also acquire an understanding of the differences between the potential improvement impacts of each indicator and provide necessary information for the decision makers, as shown in Figure 4.

A decision maker should be comprehensively considered as the main factor for strategy improvement to avoid mistakes. The IPA and three-factor methods consider the indicators' passenger evaluation scores and weights. However, the PP-based method adds the uncertainty of the indicators' improvement impacts, theory improvement difficulty, and improvement potential on the basis of indicator score and importance in the IPA method. The uncertainty of improvement impacts considers the features of passengers' perception of the service level. Improvement strategy is easily perceived due to small perception variances and high consensus levels. The potential impacts of improvements have small uncertainty, such as the "passenger civilization" indicator.

A decision maker hopes that the determined method of improvement strategy is simple and easy to understand. Although the IPA method is simple, the threshold value basis of category classification is still debated. Different determining methods of threshold value can lead to varying category classifications. A strategy can be classified in this category or in other categories according to different threshold values. For example, the "ticketing system" indicator can be classified in the "prior improvement" and "secondary improvement" categories due to variations in the threshold values. The improvement precedence of different categories is also varied. The decision maker then perceives the decision to be highly random and hard to understand and consequently hesitate, thereby reducing decision confidence. The PP-based method considers more factors than does the IPA method, and the relationships between different factors can be compounded according to physical meanings from multiple dimensions to one dimension, thus avoiding the problems of factor importance determination in multi-property decisions and becoming easy to understand. In addition, the "advantage-impacted" category in the IPA method is hard to understand because it can mean the improvement precedence or status quo maintenance. The PP-based method finds that "advantage-impacted" indicators can be ambiguously regarded as prior improvement indicators, such as "service facility," "safety," and "train information," by calculating the improvements, which is not the case in the PP method. Category differences of factor variations are caused by the threshold value of category classification, which also exists in the three-factor

Commented [Ed173033]: Check for possible contradictions (*PP method finds that…which is not the case in PP method.*).

Commented [Ed173034]: Place this clause beside the noun it modifies (Category differences, factor differences, threshold value, or category classification?).

method.

Multi-attribute indicator decisions can be compounded into one concept or index according to physical meanings, such as PP. The determination of concept connotations is simple, clear, and easily comprehensible. Reducing dimensionality according to causality rather than correlations in the big data era must be emphasized in future decision research. In addition, a single variable hosts the multiple functions, such as the "L_{max}–P_k" PP model, which reflects the improvement potential and improvement urgency and difficulty and thus needs to be studied from the perspective of decision-simplified research.

In conclusion, the PP-based method is better than the traditional IPA method in terms of concept understanding, information provided to decision makers and considered influential factors, and strong precedence result determination. However, many aspects still need to be further studied. First, the difficulty of indicator improvement, which is considered by the proposed method, is merely a theoretical difficulty that is defined by passengers' evaluation scores. Increasing the scores of indicators that already have high scores is difficult and increases the marginal effort cost. The real cost and implementation difficulty entailed by the improvement of service quality by operators are not considered. These variables in such model have clear physical meanings and are thus of equal importance. The accurate function relationship between these variables can be further studied along with the features of research problems. The newly added three-factor method considers the importance and perception characteristics of factor indicators. The evaluation of this new perspective in the PP-based method also needs to be studied.

Commented [Ed173035]: Check for accuracy.

Commented [Ed173036]: This sounds empty because *features of research problem* has no clear meaning.

References:

Anderson, E. W., and Mittal, V. (2000). "Strengthening the satisfaction-profit chain." *Journal of Service research*, 3(2), 107-120.

Angelova, B., and Zekiri, J. (2011). "Measuring customer satisfaction with service quality using American Customer Satisfaction Model (ACSI Model)." *International Journal of Academic Research in Business and Social Sciences*, 1(3), 232.

Browne, M. W., and Cudeck, R. (1993). "Alternative ways of assessing model fit." *Sage focus editions*, 154, 136-136.

Busacca, B., and Padula, G. (2005). "Understanding the relationship between attribute performance and overall satisfaction: Theory, measurement and implications." *Marketing Intelligence & Planning*, 23(6), 543-561.

Cavana, R. Y., Corbett, L. M., and Lo, Y. L. (2007). "Developing zones of tolerance for managing passenger rail service quality." *International Journal of Quality & Reliability Management*, 24(1), 7-31.

Chou, J., Kim, C., Kuo, Y., and Ou, N. (2011). "Deploying effective service strategy in the operations stage of high-speed rail." *Transportation Research Part E: Logistics and Transportation Review*, 47(4), 507-519.

De Vos, J., and Witlox, F. (2017). "Travel satisfaction revisited. On the pivotal role of travel satisfaction in conceptualising a travel behaviour process." *Transportation Research Part A: Policy and Practice*, 106, 364-373.

Ennew, C. T., Reed, G. V., and Binks, M. R. (1993). "Importance-performance analysis and the measurement of service quality." *European journal of marketing*, 27(2), 59-70.

Evans, M. R., and Chon, K. (1989). "Formulating and Evaluating Tou Rism Policy Using Importance-Performance Analysis." *Journal of Hospitality & Tourism Research*, 13(3), 203-213.

Freitas, A. L. P. (2013). "Assessing the quality of intercity road transportation of passengers: An exploratory study in Brazil." *Transportation Research Part A: Policy and Practice*, 49, 379-392.

Golden, B. L., Wasil, E. A., and Harker, P. T. (1989). "The analytic hierarchy process." *Applications and Studies, Berlin, Heidelberg*.

Hooper, D., Coughlan, J., and Mullen, M. (2008). "Structural equation modelling: Guidelines for determining model fit." *Articles*, 2.

Hsu, C. H., Byun, S., and Yang, I. (1998). "Attitudes of Korean college students towards quick-service, family-style, and fine dining restaurants." *Journal of Restaurant & Foodservice Marketing*, 2(4), 65-85.

Kano, N. (1984). "Attractive quality and must-be quality." *Hinshitsu (Quality, The Journal of Japanese Society for Quality Control)*, 14, 39-48.

Martilla, J. A., and James, J. C. (1977). "Importance-performance analysis." *The journal of marketing*, 77-79.

Matzler, K., Bailom, F., Hinterhuber, H. H., Renzl, B., and Pichler, J. (2004). "The asymmetric relationship between attribute-level performance and overall customer satisfaction: a reconsideration of the importance–performance analysis." *Industrial Marketing Management*, 33(4), 271-277.

Mouwen, A. (2015). "Drivers of customer satisfaction with public transport services." *Transportation Research Part A: Policy and Practice*, 78, 1-20.

Nassereddine, M., and Eskandari, H. (2017). "An integrated MCDM approach to evaluate public transportation systems in Tehran." *Transportation Research Part A: Policy and Practice*, 106, 427-439.

O Neill, M. A., and Palmer, A. (2004). "Importance-performance analysis: a useful tool for directing continuous quality improvement in higher education." *Quality assurance in education*, 12(1), 39-52.

Rahman, F., Das, T., Hadiuzzaman, M., and Hossain, S. (2016). "Perceived service quality of paratransit in developing countries: A structural equation approach." *Transportation Research Part A: Policy and Practice*, 93, 23-38.

Shaaban, K., and Kim, I. (2016). "Assessment of the taxi service in Doha." *Transportation Research Part A: Policy and Practice*, 88, 223-235.

Shen, W., Xiao, W., and Wang, X. (2016). "Passenger satisfaction evaluation model for Urban rail transit: A structural equation modeling based on partial least squares." *Transport Policy*, 46, 20-31.

Steiger, J. H. (1990). "Structural model evaluation and modification: An interval estimation approach." *Multivariate behavioral research*, 25(2), 173-180.

Stradling, S. G., Anable, J., and Carreno, M. (2007). "Performance, importance and user disgruntlement: A six-step method for measuring satisfaction with travel modes." *Transportation Research Part A: Policy and Practice*, 41(1), 98-106.

Stuart, K., Mednick, M., and Bockman, J. (2000). "Structural equation model of customer satisfaction for the New York City subway system." *Transportation Research Record: Journal of the Transportation Research Board*(1735), 133-137.

Susilo, Y. O., and Cats, O. (2014). "Exploring key determinants of travel satisfaction for multi-modal trips by different traveler groups." *Transportation Research Part A: Policy and Practice*, 67, 366-380.

Vandenberg, R. J., and Lance, C. E. (2000). "A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research." *Organizational research methods*, 3(1), 4-70.

Weinstein, A. (2000). "Customer satisfaction among transit riders: How customers rank the relative importance of various service attributes." *Transportation Research Record: Journal of the Transportation Research Board*(1735), 123-132.

Wu, X., Cao, J., and Huting, J. (2018). "Using three-factor theory to identify improvement priorities for express and local bus services: An application of regression with dummy variables in the Twin Cities." *Transportation Research Part A: Policy and Practice*, 113, 184-196.

Commented [Ed165237]: References were not edited. Check for possible inconsistencies with the citation style prescribed by your target journal.