ICIC Express Letters
Volume **13**, Number **11**, November **2019**

 ISSN 1881-803X
pp. **1049–1056**

THE INFLUENCE OF PASSENGER BEHAVIOR AND PERFORMANCE SATISFACTION ON THE MINI BUS MODE CHOICE ROUTE OF MALANG-SURABAYA

Sobri Abusini and Rifky Aisyatul Faroh

Department of Mathematics
University of Brawijaya
Jl. Veteran, Malang 65145, East Java, Indonesia
sobri@ub.ac.id; rifkyaisya99@gmail.com

Received May 2019; accepted August 2019

Abstract. *Transportation problems in some regions have been at an alarming stage, including congestion, accidents, delays, and pollution (noise and air). These problems can be controlled if the transportation system in the area runs well. Transportation modeling planning is an effort to improve the transportation system. The planning model used in this study is the mode choice. The research data used is the traveller of "Liza Tour and Travel Malang" who traveled from Malang to Surabaya. The result states that passenger behavior and performance satisfaction influence a person in choosing the mini bus mode. The model produced using structural equation model method is very good. The model has an R^2 value of 0.968.*

Keywords: Transportation, Mode choice, Passenger behavior, Performance satisfaction, Structural equation model

1. **Introduction.** Transportation is the movement of people or goods using tools or vehicles from original location to destination location that are geographically separated [1]. The transportation planning model is needed to reduce or overcome problems that might occur in the transportation system. Mode choice is a very important contribution for users of public transportation, especially the people of Malang who want to go to Surabaya, for example for workers who choose to stay in Malang even though they work in Surabaya. Besides that, it also plays an important role for Malang people who will go to Surabaya for the purpose of traveling or with other purposes.

Over time, more and more travelers will also have more public transportation to provide. However, this can also have an impact on traffic problems, namely the more public transportation works, the more likely the congestion will be. The good mode choice will help travelers to arrive quickly and safely where they are going. Therefore, an analysis is needed to determine the factors that influence the travelers from Malang to Surabaya.

Structural Equation Model (SEM) is compilation from the relationship between measurement model and structural model based on the assumptions. It is a combination of factor analysis and linear regression [2]. In 2008, Shiftan et al. conducted research on the causal relationship between latent variables and their relationship to measured variables while categorizing transportation markets based on these latent variables [3]. Deutsch et al. in 2013 analyzed the relationship between community travel behavior and their perceptions of the surrounding environment using SEM based on 719 sample data from surveys in Barbara and California, USA [4]. In 2017, Chen and Li expanded the model formulation to develop an ICLV model on public transportation that combines a set of

DOI: 10.24507/icicel.13.11.1049

latent variables related to comfort, personal safety, pleasure, service, and feelings of waiting that are more sensitive to users of public transportation according to data carried out in Chengdu, China [5].

In 2019, Faroh et al. conducted a study of the bus and train mode choice with three latent variables, namely passenger behavior, performance satisfaction, and mode choice. The study shows that indicators of passenger behavior and performance satisfaction influence mode choice indicators. However, the R^2 value is not good enough, which is less than 75% [6]. Therefore, this study was carried out in mini-bus mode using 210 passengers of "Liza Tour and Travel Malang" route Malang-Surabaya. The summary of this paper is structured as follows. In Section 2, we present some basic theories about transportation definition and SEM method. In Section 3, we explain the results obtained in accordance with the stages of SEM method. In the last section, we summarize the conclusion and give suggestions for future research.

2. **Materials.** Definition of transportation is the activity of moving or carrying something from one place to another [7]. Transportation is the movement of goods or passengers from one place to another, where the product is moved to the destination location. And generally, transportation is an activity to move something from original location to destination location, either by tool or not [8]. The mode of public transportation is a means of transportation whose services are intended for a number of people together. At present, public transportation modes, especially land public transportation modes are very widely used for travelers to fulfill their activity needs.

The relationship between complex variables is increasingly developing along with the development of science. One of the objectives of choosing an analytical tool is to find a device that has high explanatory power while being efficient. Some analytical tools can be used, for example, multiple regression and factor analysis. These tools have an adequate role in answering managerial questions both practically and theoretically. However, it also has drawbacks, one of which is only being able to test a single relationship at a time. To overcome the problem of research conducted in a single and one time manner, one of the appropriate analysis tools is Structural Equation Model (SEM).

SEM is an analysis technique which consists of combination of multivariate statistical techniques and which is used commonly by the scientists who are engaged in social sciences, like education, economists, and marketing researchers. SEM is an effective model testing and improving method that enables theoretical models to be tested as a whole and that can explain the cause and effect relationship of the variables in mixed hypotheses which are related to the models based on statistical dependence [9]. In general, SEM characteristics are described in the components of the SEM model consisting of:

1) 2 types of variables: Latent Variables and Observed/Measured/Manifest Variables;
2) 2 types of models: Structural Model and Measurement Model;
3) 2 types of errors: Structural Errors and Measurement Errors.

The SEM consists of two parts [10].

a. Structural model

Structural models are models that describe relationships which occur between latent variables. The general form of the structural model is as follows [11]:

$$\eta_{m\times 1} = B_{m\times m} * \eta_{m\times 1} + \Gamma_{m\times n} * \xi_{n\times 1} + \zeta_{m\times 1} \tag{1}$$

where η is endogen latent variable, ξ is exogen latent variable, B is path coefficient matrix for relationships between endogenous latent variables, Γ is path coefficient matrix for the relationship of endogenous latent variables and variables latent exogenous and ζ is structural error.

b. Measurement model

Measurement models are models that describe relationship which occur between latent variable and their indicators (observed variable). Measurement equation models for Y and X are [11]:

$$Y_{p\times 1} = \Lambda_{y_{p\times m}} * \eta_{m\times 1} + \varepsilon_{p\times 1} \tag{2}$$

$$X_{q\times 1} = \Lambda_{x_{q\times n}} * \xi_{n\times 1} + \delta_{q\times 1} \tag{3}$$

where Y is manifest variable for variable latent endogeneous, X is manifest variable for variable latent exogeneous, Λ_y is coefficient variable latent endogeneous, Λ_x is coefficient variable latent exogeneous, ε is error measurement that connects to Y, and δ is error measurement that connects to X.

Bollen and Long in 1993 said that the procedure or stage of the SEM in general will contain the following steps [12].

1) Model specification

The first step in SEM analysis is the model specification to be estimated. Model specification involves using all available relevant theory, research and information, and developing a theoretical model [13].

2) Identification

This stage contains a study of the possibility that unique set of parameters will be obtained for each parameter contained in the model and the simultaneous equation is likely to have no solution [14]. Broadly speaking, there are 3 types of identification in a simultaneous equation, namely Under-identified model, Just-identified model, and Over-identified model.

3) Estimates

This estimation phase explains the estimation or estimation of the model which aims to produce several parameter values using one of the available estimation methods. The choice of the estimation method to be used must be determined according to the characteristics or characteristics of the variables to be analyzed.

4) Goodness of Fit

This stage is the stage of testing the similarity between the data and the model. According to Hair et al., in 1998, an assessment of the high compatibility of data with the model was carried out in several stages, namely overall model fit, measurement model fit, and structural model fit [9].

5) Respecification

This stage is a respecification or modification of the model according to the results of the compatibility test in the previous stage.

3. Main Results.

Latent variables in SEM consist of two types, namely exogenous variables and endogenous variables. Exogenous variables in this study are passenger behavior and performance satisfaction. While the endogenous variables used in this study are mode choice. Manifest variables of Passenger Behavior are age, income, family number, vehicle ownership, and distance. Manifest variables of the Performance Satisfaction are safety, accuracy, speed, service, cost, and cleanliness. And the manifest variables of the Mode Choice are accessibility, security, convenience and frequency of travel.

3.1. Normality test.

Requirement for data can be processed using the SEM method of normality. The assumption of normality can be tested with the help of LISREL 9.30 and SPSS software by looking at the statistical value of z skewness and kurtosis. If the value of $z_{count} \leq 0.05$ then it can be said that the data distribution is not normal. Conversely, if the value of $z_{count} > 0.05$ then it can be said that the data distribution is normal.

Description of variable names in Figure 1, Figure 2, and Figure 3 is: variable US is age, PD is income, JK is family number, KK is vehicle ownership, JR is distance, KS is safety, KT is accuracy, KC is speed, PL is service, BE is cost, KB is cleanliness, AK is accessibility, KA is security, KY is convenience, and FR is frequency of travel.

Test of Univariate Normality for Continuous Variables						
	Skewness		Kurtosis		Skewness and Kurtosis	
Variable	Z-Score	P-Value	Z-Score	P-Value	Chi-Square	P-Value
US	0.909	0.363	−0.908	0.364	1.652	0.438
PD	0.195	0.845	−0.557	0.577	0.349	0.840
JK	0.106	0.916	−0.291	0.771	0.096	0.953
KK	0.138	0.890	−0.793	0.428	0.648	0.723
JR	1.417	0.157	0.413	0.679	2.178	0.337
KS	−1.513	0.130	−0.160	0.873	2.316	0.314
KT	−1.298	0.194	0.769	0.442	2.275	0.321
KC	−1.202	0.229	−1.165	0.244	2.804	0.246
PL	−0.410	0.682	−0.631	0.528	0.566	0.754
BE	0.004	0.997	0.362	0.718	0.131	0.937
KB	−1.333	0.183	0.993	0.321	2.763	0.251
AK	−0.042	0.967	−0.647	0.517	0.421	0.810
KA	−1.362	0.173	−1.606	0.108	4.433	0.109
KY	−1.230	0.219	−0.223	0.823	1.564	0.458
FR	−1.051	0.293	−0.859	0.390	1.842	0.398

FIGURE 1. Univariate normality test of mini bus passenger data

From the results above, the univariate normality test shows the test results for each variable normality since the P-value Skewness and Kurtosis each indicator is greater than 0.05.

3.2. **Multicollinearity test.** The absence of correlation between independent variables indicates that the regression model is good. Multicolinearity test can be seen from the correlation value. Variables are said not to have high multicolinearity if the correlation value is $-0.7 < r < 0.7$ then the variables are said not to have high multicollinearity or correlation. Based on the calculation of Pearson correlation it can be seen that between variables do not have multicollinearity or high correlation.

3.3. **SEM analysis 1.**

1) Model specification 1

Model specification is related to the formation of the initial structural model. This initial structural model can be seen in Figure 2.

2) Identification 1

The model requirement that can be analyzed by SEM is over identified model, which is a model whose estimated parameters are smaller than the number of known data ($df > 0$). The degree of freedom from the above model is 87.

3) Estimation 1

The estimation method used in this study is the Maximum Likelihood method with the help of LISREL 9.3 software analysis done with 289 maximum iteration.

4) Goodness of Fit 1

a. Measurement Model Fit 1

• Validity 1

Validity test can be seen from the factor loading and the t factor loading. According to Sharma in 1996, the weakest loading factor that can be accepted is 0.40 [15]. Valid results from the initial structural model can be seen in Table 1.

• Reliability 1

The Construct Reliability (CR) measure is used to measure reliability with SEM. Variables that have a good CR are those that have a value of more than 0.70. The results of reliability in the initial model can be seen in Table 2.

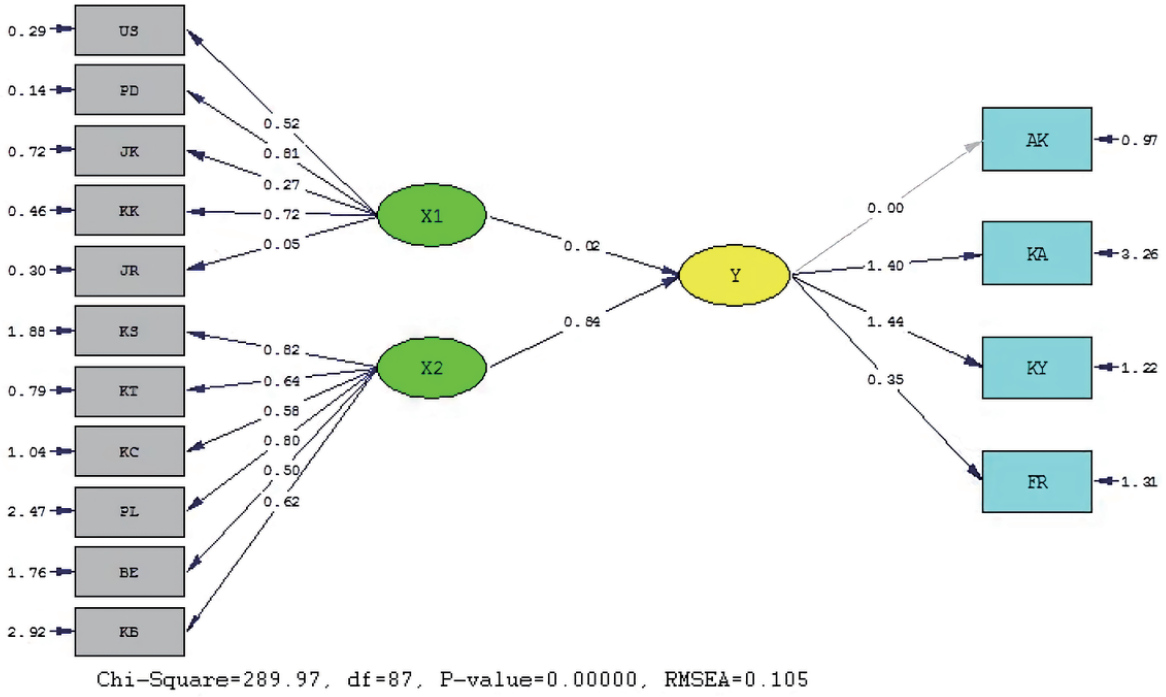

FIGURE 2. Path of model specification (1)

TABLE 1. Validity of manifest variable 1 of mini bus data

No.	Manifest Variable	Latent Variable	λ standard	λ analyze	t standard	t analyze
1.	Age	Passenger Behavior	0.40	0.52	1.96	10.58
2.	Income	Passenger Behavior	0.40	0.81	1.96	14.35
3.	Family Number	Passenger Behavior	0.40	0.27	1.96	4.13
4.	Vehicle Ownership	Passenger Behavior	0.40	0.72	1.96	11.09
5.	Distance	Passenger Behavior	0.40	0.05	1.96	1.16
6.	Safety	Performance Satisfaction	0.40	0.82	1.96	6.71
7.	Accuracy	Performance Satisfaction	0.40	0.64	1.96	7.76
8.	Speed	Performance Satisfaction	0.40	0.58	1.96	6.52
9.	Service	Performance Satisfaction	0.40	0.80	1.96	5.89
10.	Cost	Performance Satisfaction	0.40	0.50	1.96	4.47
11.	Cleanliness	Performance Satisfaction	0.40	0.62	1.96	4.36
12.	Accessibility	Mode Choice	0.40	0.00	1.96	–
13.	Security	Mode Choice	0.40	1.40	1.96	0.05
14	Convenience	Mode Choice	0.40	1.44	1.96	0.05
15	Frequency of Travel	Mode Choice	0.40	0.35	1.96	0.05

TABLE 2. Reliability of latent variable 1 of mini bus data

Latent Variable	**CR**	**Reliability**
Passenger Behavior ($X1$)	0.681	Bad
Performance Satisfaction ($X2$)	0.885	Good
Mode Choice (Y)	0.926	Good

b. Structural Model Fit 1

The structural model obtained from analysis result is $Y = -0.0189X1 + 0.839X2$; $R^2 = 0.704$.

c. Overall Model Fit 1

The overall suitability test of the model can be seen from some measures of Goodness of Fit (GOF), including RMSEA (Root Mean Square Error of Approximation), GFI (Goodness of Fit Index), AGFI (Adjusted Goodness of Fit Index) and CFI (Comparative Fit Index).

TABLE 3. Result of model feasibility 1 of mini bus data

Goodness of Fit	Cut off Value	Analysis Result	Model Evaluation
RMSEA	≤ 0.08	0.105	Bad
GFI	≥ 0.90	0.843	Bad
AGFI	≥ 0.90	0.784	Bad
CFI	≥ 0.95	0.704	Bad

Based on Table 3 above, it can be seen that the value of the feasibility test of the models is at a criterion that is not good or below the standard value.

5) Respecifications

Based on Table 1 there are still variables that have less than standard loading factor values. The results of testing the feasibility of the model also indicate that the model still has poor criteria. Therefore. it is respecified that it is removing the variables and then modifying the model so that a model that has a good match with the data can be found.

Model modification is carried out with the same stage as in the previous SEM analysis stage.

1) Model Specification 2

The structural model after modification is obtained as shown in Figure 3.

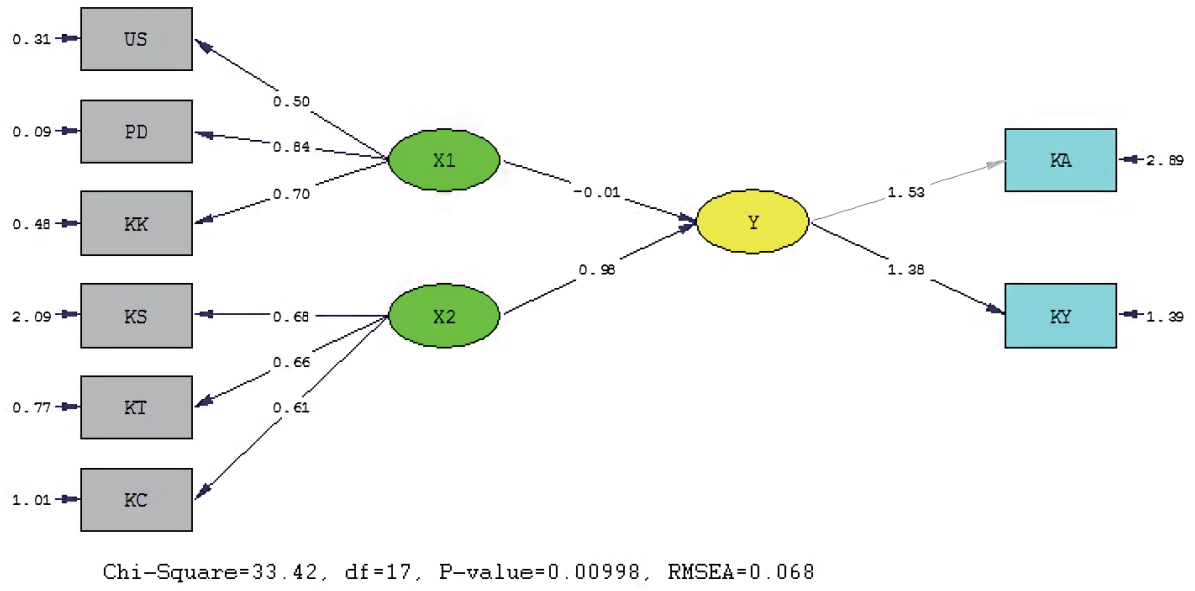

FIGURE 3. Path of model specification (2)

2) Identification 2

The degree of freedom value after the model specification is 17 (the amount of known data is greater than the estimated amount of data). This shows that the model can be analyzed because it includes the over-identified model.

3) Estimation 2

The estimation method produced with the help of the LISREL 9.3 software stopped at the 12th iteration.

4) Goodness of Fit 2

a. Measurement Model Fit 2

• Validity 2

The results of the modified model validity can be seen in Table 4 below.

Table 4. Validity of manifest variable 2 of mini bus data

No.	Manifest Variable	Latent Variable	λ standard	λ analyze	t standard	t analyze
1.	Age	Passenger Behavior	0.40	0.50	1.96	9.91
2.	Income	Passenger Behavior	0.40	0.84	1.96	14.49
3.	Vehicle Ownership	Passenger Behavior	0.40	0.70	1.96	10.59
4.	Safety	Performance Satisfaction	0.40	0.65	1.96	5.50
5.	Accuracy	Performance Satisfaction	0.40	0.66	1.96	7.78
6.	Speed	Performance Satisfaction	0.40	0.61	1.96	6.74
7.	Security	Mode Choice	0.40	1.53	1.96	–
8.	Convenience	Mode Choice	0.40	1.38	1.96	7.43

• Reliability 2

The results of the model reliability after modification show good CR values on all three latent variables. This can be seen in Table 5 below.

Table 5. Reliability of latent variable 2 of mini bus data

Latent Variable	**CR**	**Reliability**
Passenger Behavior ($X1$)	0.813	Good
Performance Satisfaction ($X2$)	0.773	Good
Mode Choice (Y)	1.120	Good

b. Structural Model Fit 2

The structural models obtained from the analysis results is $Y = -0.00655X1 + 0.984X2$ with $R^2 = 0.968$.

c. Overall Model Fit 2

The overall suitability test on the modified model can be seen in Table 6 below.

Table 6. Result of model feasibility 2 of mini bus data

Goodness of Fit	**Cut off Value**	**Analysis Result**	**Model Evaluation**
RMSEA	≤ 0.08	0.068	Good
GFI	≥ 0.90	0.964	Good
AGFI	≥ 0.90	0.924	Good
CFI	≥ 0.95	0.962	Good

Based on Table 6 above it can be seen that the value of the feasibility test of the RMSEA, GFI, AGFI and CFI models is in the criteria of good or above the standard value.

4. **Conclusions.** In this paper, we concluded that indicators of passenger behavior and performance satisfaction influenced the mini-bus mode choice. That indicators are Age, Income, Vehicle Ownership, Safety, Accuracy, Speed, Security, and Convenience. While the model produced is $Y = -0.00655X1 + 0.984X2$ with $R^2 = 0.968$. The statement means that the structural model above can explain the actual situation in the field by

96.8%. We hope this discussion can expand the view about mode choice and SEM method. The future research can use larger amount of data and more variables.

REFERENCES

[1] P. A. Steenbrink, *Optimization of Transport Network*, Wiley, London, 1974.

[2] J. B. Ullman, Structural equation modelling, in *Using Multivariate Statistics (4th Edition)*, B. G. Tabachnick and L. S. Fidell (eds.), Boston, Allyn & Bacon, 2001.

[3] Y. Shiftan, M. L. Outwater and Y. Zhou, Transit market research using structural equation modeling and attitudinal market segmentation, *Transport Policy*, vol.15, no.3, pp.186-195, 2008.

[4] K. Deutsch, S. Y. Yoon and K. Goulias, Modeling travel behavior and sense of place using a structural equation model, *Journal of Transport Geography*, vol.28, pp.155-163, 2013.

[5] J. Chen and S. Li, Mode choice model for public transport with categorized latent variables, *Mathematical Problems in Engineering*, vol.2017, pp.1-11, 2017.

[6] R. A. Faroh, S. Abusini and Marjono, Mode choice factors of the bus and train route of Malang-Surabaya, *American Journal of Engineering Research,* vol.8, pp.350-358, 2019.

[7] E. K. Morlok, *Introduction to Transportation Engineering and Planning*, Mc. Graw-Hill Kogakuha, 1978.

[8] C. Bowersox, *Introduction to Transportation*, Macmillan Publishing Co. Inc, New York, 1981.

[9] J. H. Hair, R. I. Tatham and R. E. Andenson, *Multivariate Data Analysis*, 5th Edition, Prentice Hall International, New York, 1998.

[10] S. Kumar, Structural equation modeling basic assumptions and concepts: A novices guide, *Asian Journal of Management Sciences*, vol.3, no.7, pp.25-28, 2015.

[11] N. Timm, *Applied Multivariate Analysis*, Springer, New York, 2002.

[12] K. A. Bollen and J. S. Long, *Testing Structural Equation Models*, Sage Publications, USA, 1993.

[13] A. Nayernia, A review of structural equation modeling (SEM) and its application in language education research, *MJAL*, 2013.

[14] B. Byrne, *Structural Equation Modeling with AMOS: Basic Concepts, Applications, and Programming (2nd Edition)*, Routledge Taylor & Francis Group, London, 2010.

[15] S. Sharma, *Applied Multivariate Techniques*, John Willey & Sons Englewood Chiffs New Jersey, New York, 1996.