

STATISTICAL MODELING OF TOURISM INVESTMENT DECISIONS IN INDONESIA USING SEMIPARAMETRIC APPROACH

**Yossy Maynaldi Pratama^{1*}, Adji Achmad Rinaldo Fernandes²,
Ni Wayan Surya Wardhani³, Nurjannah⁴, Solimun⁵**

^{1,2,3,4,5} *Statistics Department, Faculty of Mathematics and Natural Sciences, Brawijaya University*

Veteran Street, Malang, 65145, Indonesia

Corresponding author's e-mail: *yossmaynaldi@student.ub.ac.id

ABSTRACT

Article History:

Received: 2nd November 2023

Revised: 10th January 2024

Accepted: 28th January 2024

Keywords:

Semiparametric;
Path Analysis;
Investment Decision.

The tourism potential in Indonesia is very large considering that Indonesia consists of tens of thousands of separate islands. Indonesia has many diverse natural landscapes, with all the natural riches and biodiversity that exist within it, which is an attraction for investors who want to invest in Indonesia. The existence of relationships between variables that are linear and nonlinear, where no nonlinear pattern is known, requires a semiparametric approach. This study aims to apply a semiparametric approach to model people's investment decisions in tourism in Indonesia. The data used is in the form of investor respondents who invest in the tourism sector in Indonesia from the National Competitive Basic Research (PDKN) in 2022, totaling 100 respondents. This study uses the semiparametric path analysis method to model tourism investment decisions in Indonesia. The results show that regulatory variables and investment interest variables have a significant and positive effect on investment decision variables. A diversity coefficient of 60.2% indicates that data diversity can be explained by 60.2% with models, while other variables outside the study explain the remaining 38.8%. In other words, the regulatory variable (X) and the investment interest variable (Y_1) can influence the investment decision variable (Y_2) by 60.2%.



This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/).

How to cite this article:

Y. M. Pratama, A. A. R. Fernandes, N. W. S. Wardhani, Nurjannah and Solimun., "STATISTICAL MODELING OF TOURISM INVESTMENT DECISIONS IN INDONESIA USING SEMIPARAMETRIC APPROACH," *BAREKENG: J. Math. & App.*, vol. 18, iss. 1, pp. 0529-0536, March, 2024.

Copyright © 2024 Author(s)

Journal homepage: <https://ojs3.unpatti.ac.id/index.php/barekeng/>

Journal e-mail: barekeng.math@yahoo.com; barekeng_journal@mail.unpatti.ac.id

1. INTRODUCTION

The tourism potential in Indonesia is very large considering that Indonesia consists of tens of thousands of separate islands. Thousands of islands separated and surrounded by two continents cause Indonesia have abundant biodiversity and strong cultural influences, adding to the richness of Indonesian culture [1]. The tourism sector that tourists are most interested in when visiting Indonesia is Indonesia's natural diversity, biodiversity and cultural diversity. The nature potential of Indonesia includes unique flora and fauna such as Komodo dragons, corpse flowers, and other types that are very rare. Indonesia has many diverse landscapes such as Raja Ampat, Bunaken, Badr Island, and others. Indonesia, with all its natural riches and biodiversity, is an attraction for investors who want to invest in Indonesia.

Investment can be defined as a financial activity in a company or institution by providing funds through proof of investment letters with the aim of getting results that are more than the amount of funds provided, such as investment in land, gold, houses, deposits, bonds, stocks, and securities [2]. There are two paradigms of understanding investment. First, investing is seen as a desire, rather than investing money people would prefer saving their money. Second, investment is seen as a necessity where a person will choose to invest his money rather than save it.

Investment interest can be interpreted as a strong desire, tendency, or urge to carry out investment activities with a feeling of pleasure in investing in future profit activities. Several things are thought to affect a person's interest in investing, including the benefits of investing, minimum investment capital, motivation to invest, profits generated from investment, and investment risks.

There are many considerations in determining investment decisions, investors must pay attention to the various risks they will face, both risks for the long term and the short term, because it will affect the results that will be obtained whether it is loss or profit [3]. The understanding that investors have mastered will be a provision for knowing how to avoid risks.

The analysis used to determine the investment decision is influenced by the regulatory system and public investment interest using the semiparametric path analysis method. The semiparametric approach is used when there are two different approaches in the same relationship. A different approach occurs when in one endogenous variable there are two different patterns of relationships between exogenous and endogenous variables. In other words, one of the relationships between exogenous variables to endogenous variables has a linear relationship pattern, so a parametric approach is used. The relationship between exogenous variables and the same endogenous variable has a nonlinear relationship pattern and the form of the relationship is unknown, so a nonparametric approach is needed. The existence of these two different approaches is the basis for using a semiparametric approach.

Based on the description above, this study aims to model the influence of regulatory variables and investment interest variables on investment decisions using semiparametric path analysis and find out which variables have the most influence on investment decision variables.

2. RESEARCH METHODS

2.1 Research Variables

In this study, the following variables were employed: regulation (synchronization) variable as exogenous variable (X), Investment interest as a mediation variable (Y_1), and investment decisions as endogenous variables (Y_2). The data used comes from the 2022 National Competitive Basic Research (PDKN) data. A total of 100 respondents who were investors who invested in the tourism sector in Indonesia were used in this research. In detail, the variables used in this study can be seen in **Table 1**.

Table 1. Research Variables

Research Variable	Indicator
Regulation (X)	1) Simplification

Research Variable	Indicator
	2) Harmonious
	3) Clear
	4) Businesslike
	5) Effective
	6) Efficient
Investment Interest (Y ₁)	1) Benefit
	2) Capital
	3) Motivation
	4) Advantage
	5) Education
Investment Decisions (Y ₂)	1) Profitability
	2) Market Orientation
	3) Tax Protection
	4) Company Growth
	6) Investment Selection

On the basis of a sound legal systematization, synchronization is the activity of observing the conformity or alignment of laws and regulations vertically, specifically between higher and lower laws and regulations [4]. Laws and regulations should not be in conflict with one another or with higher laws and regulations when they are synchronized or aligned [4].

In this study the definition of interest is associated with investment. Investment is one of several development instruments needed by the state in improving the welfare of its people [5]. Investment interest can be defined as the desire to find out about the type of investment starting from advantages, weaknesses, investment performance, and so on [5]. How much time and effort someone puts into learning about, studying, and then practicing a particular type of investment gives insight into their characteristics as investors. A person's understanding of how to invest in the capital market, the amount of capital available to invest, and the desire to demonstrate their existence are all factors that are thought to influence their interest in doing so. [6].

One of the duties of financial management is making investment decisions, which entails distributing money—both internal and external—into different kinds of investments with the goal of increasing profits from the cost of capital in the future. Short-term investments include those made in cash, short-term securities, receivables, and inventories, among other assets. Long-term investment in the form of real estate, construction materials, transportation, manufacturing, and other fixed assets. Maximizing the value of the business is a goal in addition to future profits.

2.2 Parametric Path Analysis

Path analysis is known as causal modeling because it allows testing theoretical proportions about the causal relationships of certain variables. When multiple regression cannot adequately explain complex relationships between variables, path analysis is used [7].

One of the advantages of path analysis is that it can find out the overall influence of the independent variable and break it down into several types of influences. There are various kinds of influence between variables in path analysis, namely direct influence, indirect influence, and total influence [8].

Design of pathway models based on research theories and concepts. Theoretically, relationships between variables can be described with path diagrams or equations to form structural models. In this study, the parametric path model used as follows.

$$Y_{1i} = \beta_{01} + \beta_{XY_1} X_i + \varepsilon_{1i} \quad (1)$$

$$Y_{2i} = \beta_{02} + \beta_{XY_2} X_i + \beta_{Y_1Y_2} Y_{1i} + \varepsilon_{2i} \quad (2)$$

Exogenous variables should be standardized if the units of each variable are different so that each variable has the same unit or units or omits the unit of the variable. Standardization is carried out with standard normal transformations with an average of 0 and a variety of 1, and the following formula can be used.

$$Z_{X_i} = \frac{x_i - \bar{x}}{s_x} \quad (3)$$

$$Z_{Y_i} = \frac{y_i - \bar{y}}{s_y} \quad (4)$$

where:

Z_{X_i} : Standardized exogenous variable values

x_i : Exogenous variable values

\bar{x} : Average of exogenous variables

s_x : Exogenous variable standard deviation

Z_{Y_i} : Standardized endogenous variable values

y_i : Endogenous variable values

\bar{y} : Average of endogenous variables

s_y : Endogenous variable standard deviation

So the standardized path equation can be written as follows:

$$Z_{y_{1i}} = \beta_{XY_1} Z_{X_i} + \varepsilon_{1i} \quad (5)$$

$$Z_{y_{2i}} = \beta_{XY_2} Z_{X_i} + \beta_{Y_1Y_2} Z_{Y_{1i}} + \varepsilon_{2i} \quad (6)$$

2.3 Nonparametric Path Analysis

A spline is a part or piece of a polynomial that has continuous and segmented properties [9]. An advantage of using truncated spline polynomial regression is that it has the ability to identify a unique approach for estimating regression curves. This is possible due to splines have common fusion points that exhibit patterns of data behavior called vertices. The spline's high flexibility is a result of the knot point's ability to adjust local characteristics effectively. Additionally, truncated spline polynomial regression has the benefit of being objective, reducing the amount of subjectivity. The least squares method is used for optimization in order to produce quick, easy, and accurate mathematical calculations that support statistical calculations.

A careful and complicated process is needed to conduct research on independent and characterful splines. The following points are the advantages of nonparametric spline regression: (a) Splines have a unique and excellent statistical interpretation. The Penalized Least Square (PLS) method is optimized to produce the spline model. (b) Spline can handle data and functions efficiently. (c) Splines succeed at handling data that exhibits changes at particular sub-intervals. (d) Spline performs well at generalizing intricate and intricate statistical modeling. [10]. One spline model that is interesting in application is the truncated polynomial spline [11]. The truncated spline nonparametric path analysis model is [12]:

$$Y_{1i} = \hat{f}_{1i} + e_{1i} \quad (7)$$

$$Y_{2i} = \hat{f}_{2i} + e_{2i} \quad (8)$$

The general form of a truncated spline function with one predictor variable is as follows:

$$\hat{f}(x_i) = \hat{\beta}_0 + \sum_{h=1}^p \hat{\beta}_h x_i^h + \sum_{j=p+1}^K \hat{\beta}_j (x_i - k_j)_+^p \quad (9)$$

where:

$\hat{f}(x_i)$: Regression functions whose pattern is not yet known to be suspected.

x_i : Predictor variables value.

i : 1,2, ..., n where n is the number of observations.

h : 1,2, ..., p where p is a polynomial order spline regression.

j : 1,2, ..., K where K is the number of knot points.

k : Knot point.

Here is a truncated spline regression model using linear polynomial degrees with 1 knot point:

$$\hat{f}(x_i) = \hat{\beta}_0 + \hat{\beta}_1 x_i + \hat{\beta}_2 (x_i - k_j)_+ \quad (10)$$

where:

$\hat{f}(x_i)$: Response variables

$$(x_i - k_j)_+^p = \begin{cases} (x_i - k_j)^p; & x_i \geq k_j \\ 0; & x_i < k_j \end{cases}$$

ε_i : Normal Distributed Error with Zero Mean and Variation σ^2

The nonparametric approach is used to estimate data patterns when no information is obtained about the shape of the function or path curve [13]. The purpose of the nonparametric approach is to estimate the path curve that is not yet known. Spline curves can adjust to changes in data behavior [14]. The advantage of spline curves is that they are able to handle changing data patterns at specific subintervals. Truncated spline is one method of approach to nonparametric regression models that is often used. Truncated splines are in the form of polynomial pieces that are continuous and segmented. The benefit of the truncated spline approach to the nonparametric approach is that it tends to find its own estimation of model functions according to the data. This is because in the truncated spline, there are knot points, which are intersection points that indicate changes in data behavior patterns. The spline's high flexibility is a result of the knot point's ability to adjust effectively to local characteristics. The following equations are optimized using the ordinary least squares method (OLS):

$$\min \{ \varepsilon' \varepsilon \} = \min \{ (y - X\beta)' (y - X\beta) \} \quad (11)$$

2.4 Semiparametric Path Analysis

Semiparametric path models can occur when there are some known forms of relationships and some forms of relationships are not yet known. The equation of the semiparametric structural model is as follows [15].

$$y_{1i} = x_{1i}\beta + \varepsilon_{1i} \quad (12)$$

$$y_{2i} = x_{2i}\beta + f(y_{1i}) + \varepsilon_{2i} \quad (13)$$

Based on the above equation where the character of the relationship between $(X \rightarrow Y_1)$ and $(X \rightarrow Y_2)$ using a parametric approach, while the form of relationships $(Y_1 \rightarrow Y_2)$ using a nonparametric approach. So, it can be formed into the equation below which is a semiparametric SEM approximation truncated linear spline with 1 knot:

$$\hat{y}_1 = \hat{\beta}_{01} + \hat{\beta}_{11}x_{1i} \quad (14)$$

$$\hat{y}_2 = \hat{\beta}_{02} + \hat{\beta}_{12}x_{1i} + \hat{\beta}_{22}y_{1i} + \delta_{11}(y_{1i} - K_{11})_+ \quad (15)$$

2.5 Stage of Analysis

The stage of analysis in this study are as follows.

- 1) Input data.

- 2) Assign endogenous variables and exogenous variables.
- 3) Construct the path diagram.
- 4) Calculate the average of each indicator contained in the variable.
- 5) Test linearity assumptions to find out patterns of relationships between variables.
- 6) Perform semiparametric path analysis.
- 7) Test hypotheses on formed semiparametric models.
- 8) Check the validity of the model using R^2 test statistics.
- 9) Interpret the magnitude of the path analysis coefficient and the significance of the influence of parameters in the semiparametric model formed.

3. RESULTS AND DISCUSSION

3.1 Linearity Test

Having established the correlation between the variables, a linearity test was performed using the Ramsey *Reset test*. Latent variable data on the linearity assumption test were obtained using an average score. The following are the linearity test results in **Table 2**.

Table 2. Linearity Test Results

Relationship	F - Value	p-value	Result
X toward Y_1	0.466	0.716	Linear
X toward Y_2	2.832	0.355	Linear
Y_1 toward Y_2	10.049	0.002	Non-linear

Regarding the outcomes of **Table 2**, there is a non-linear relationship that can be explained, specifically the connection between Y_1 towards Y_2 due to the p-value being less than 0.05, the result is rejecting H_0 . In addition, the relationship between X towards Y_1 and X towards Y_2 has a p-value greater than 0.05 thus receiving H_0 which states the two variables have a linear relationship. The results of secondary data have two forms, namely there is a linear relationship and there is a non-linear relationship.

3.2 Direct Effect Testing

Testing of semiparametric path analysis models that explain the form of non-linear relationships, that is, relationships between variables Y_1 towards Y_2 , as well as relationships between variables X towards Y_1 and variable X towards Y_2 which has the form of a linear relationship. The process of semiparametric path analysis is limited to linear shapes (order = 1) and 1 knot point, where the following equation can be formed.

$$\hat{y}_1 = \hat{\beta}_{11}Z_{x_{1i}} \quad (16)$$

$$\hat{y}_2 = \hat{\beta}_{12}Z_{x_{1i}} + f(Z_{y_1}) \quad (17)$$

$$f(Z_{y_1}) = \hat{\beta}_{22}Z_{y_{1i}} + \delta_{11}(Z_{y_{1i}} - K_{11})_+ \quad (18)$$

Based on the above equation, further semiparametric path analysis is carried out based on direct influence. The semiparametric approach in question is carried out on the model \hat{y}_2 Where parametric and non-parametric approaches are carried out together to determine the path coefficients in the model \hat{y}_2 presented on **Table 3**.

Table 3. Direct Effect Test Results

Relationship	Parameter	Coefficient	p-value	Result
X -> Y ₁	$\hat{\beta}_{11}$	0.604	0.002	Significant
X -> Y ₂	$\hat{\beta}_{12}$	0.186	0.012	Significant
Y ₁ -> Y ₂	$\hat{\beta}_{22}$	0.209	0.009	Significant
	δ_{11}	0.281	0.008	Significant

3.3 Result and Discussion

In the context of the analysis's findings utilizing secondary data, it is discovered that regulatory variables (X), Investment interest (Y₁), and investment decision (Y₂) have a substantial impact on one another. Regulation variable (X) in the first measurement model, it directly has a significant and positive effect on investment interest. This can be seen in the path coefficient of 0.604 then p-value <0.000. So that the higher the value of the regulation variable (X) will increase public interest in investing. Every change that occurs in the regulation will have a significant change effect on investment interest. These findings are consistent with studies carried out by [16] which states that the more in harmony or in accordance with laws and regulations with each other in a certain category, it can increase public interest in investing. The investment referred to in the context of this study is tourism investment.

Similar results are shown in the second measurement model where the decision variable invests as the response variable. Directly regulation variable (X) Significant influence on investment decision variables (Y₂). The path coefficient is 0.186 then p-value 0.000 indicates that regulation variables (X) is a significant driving factor and has a positive influence on investment decision variables (Y₂). In other words, the value of the regulation variable increases (X), So also increases the value of investment decision variables (Y₂). This is according to research [17] which states that regulatory factors have a positive and significant influence on investment decisions.

Investment interest variable (Y₁), according to the path analysis testing findings, influence on investment decision variables (Y₂) significantly positive. This can be seen in the value of the path coefficients in both regimes, i.e. in the left regime has a path coefficient of 0.209 and the right is as big as 0.209 + 0.281(y_{1i} - 0.076) with their respective p-values 0.000. The size of the left regime path coefficient is as large as 0.209 has the meaning that from the interest in investing (Y₁) on the range 0% until 53.6% data has a direct influence value of 0.209 then in the range above 53.6% until 100% has a direct influence as large as 0.209+0.281(y_{1i} - 0.076). In other words, the better the score of the investment interest variable score (Y₁) then there will be a spike in the magnitude of the value of the variable score of investment decisions (Y₂). This is consistent with studies carried out by [18] which claims that investment interest significantly influences decisions about investments in a positive way. The coefficient of determination obtained in this model is as large as 60.2% which means that as big as 60.3% Data diversity can be explained in the model and the rest (39.8%) is explained by other variables outside the model.

4. CONCLUSIONS

Our research result can be explained as follow:

1. Estimating functions in semiparametric path analysis is done by combining linear models (parametric) and nonlinear models (nonparametric). As well as obtaining semiparametric models as follows.

$$\hat{y}_1 = 0.604Z_{x_{1i}}$$

$$\hat{y}_2 = 0.186Z_{x_{1i}} + 0.209Z_{y_{1i}} + 0.281(Z_{y_{1i}} - 0.076)_+$$

2. The influential variable in influencing investment decisions is the variable of investment interest with the path coefficient of 0.209.

ACKNOWLEDGMENT

The author gives great appreciation to Dr. Ir. Solimun, MS. and Dr. Adji Achmad Rinaldo Fernandes, S.Si., M.Sc. who have helped in the form of research data derived from the 2022 National Competitive Basic Research (PDKN). As well as to Universitas Brawijaya who helped complete this research.

REFERENCES

- [1] D. Wahyuni, I. K. P. Sari, A. Prasetyo, and M. Berlianaldo, "Environmental Sustainability Analysis in the Tourism Sector: A Perspective on Resilient and Sustainable Tourism Development in Indonesia," 2023. [Online]. Available: <https://www.researchgate.net/publication/372405393>
- [2] M. Artaya, I. Bagus, A. Purbawangsa, L. Gede, and S. Artini, "Pengaruh Faktor Ekonomi Makro, Risiko Investasi dan Kinerja Keuangan Terhadap Return Saham Perusahaan di Bursa Efek Indonesia (BEI)," 2014.
- [3] Y. R. O. Mandagie, M. Febrianti, and L. Fujianti, "Studi Kasus Mahasiswa Akuntansi Universitas Pancasila," 2020. [Online]. Available: <http://journal.univpancasila.ac.id/index.php/RELEVAN>
- [4] M. Rus'an Yasin, M. Akbar, M. Y. Hasmin, and B. I. Hukum, "Kewenangan Ombudsman dalam Penanganan Laporan Pelayanan Publik Ombudsman's Authority in Handling Reports Public Service," *Jurnal Kolaboratif Sains*, vol. 4, no. 5, pp. 241–246, 2021.
- [5] A. Mastura, S. Nuringwahyu, and D. Zunaida, "Pengaruh Motivasi Investasi, Pengetahuan Investasi Dan Teknologi Informasi Terhadap Minat Berinvestasi Di Pasar Modal (Studi Pada Mahasiswa Fia Dan Feb Unisma Yang Sudah Menempuh Mata Kuliah Mengenai Investasi)," *Jiagabi*, vol. 9, no. 1, pp. 64–75, 2020.
- [6] A. Nisa and L. Zulaika, "Pengaruh Pemahaman Investasi, Modal Minimal Investasi, dan Motivasi terhadap Minat Mahasiswa Berinvestasi di Pasar Modal," 2017.
- [7] N. Barbieri, C. Ghisetti, M. Gilli, G. Marin, and F. Nicolli, "A Survey of the Literature on Environmental Innovation Based on Main Path Analysis," *Journal of Economic Surveys*, vol. 30, no. 3. Blackwell Publishing Ltd, pp. 596–623, Jul. 01, 2016. doi: 10.1111/joes.12149.
- [8] A. K. Montoya and A. F. Hayes, "Two-condition within-participant statistical mediation analysis: A path-analytic framework," *Psychol Methods*, vol. 22, no. 1, pp. 6–27, Mar. 2017, doi: 10.1037/met0000086.
- [9] R. Hidayat and M. Sam, "Model Regresi Nonparametrik dengan Pendekatan Spline Truncated," *Prosiding Seminar Nasional*, vol. 03, no. 1, 2018.
- [10] M. J. Lai and L. Wang, "Bivariate penalized splines for regression," *Stat Sin*, vol. 23, no. 3, pp. 1399–1417, Jul. 2013, doi: 10.5705/ss.2010.278.
- [11] Sifriyani, Haryatmi, I. N. Budiantara, and Gunardi, "Geographically weighted regression with spline approach," *Far East Journal of Mathematical Sciences*, vol. 101, no. 6, pp. 1183–1196, Mar. 2017, doi: 10.17654/MS101061183.
- [12] A. A. R. Fernandes, I. Nyoman Budiantara, B. W. Otok, and Suhartono, "Spline estimator for bi-responses nonparametric regression model for longitudinal data," *Applied Mathematical Sciences*, vol. 8, no. 113–116, pp. 5653–5665, 2014, doi: 10.12988/ams.2014.47566.
- [13] Z. Awang, W. M. A. Wan Afthanorhan, and M. A. M. Asri, "Parametric and Non Parametric Approach in Structural Equation Modeling (SEM): The Application of Bootstrapping," *Mod Appl Sci*, vol. 9, no. 9, Jul. 2015, doi: 10.5539/mas.v9n9p58.
- [14] A. Perperoglou, W. Sauerbrei, M. Abrahamowicz, and M. Schmid, "A review of spline function procedures in R," *BMC Medical Research Methodology*, vol. 19, no. 1. BioMed Central Ltd., Mar. 06, 2019. doi: 10.1186/s12874-019-0666-3.
- [15] B. Xu, Y. Luo, R. Xu, and J. Chen, "Exploring the driving forces of distributed energy resources in China: Using a semiparametric regression model," *Energy*, vol. 236, Dec. 2021, doi: 10.1016/j.energy.2021.121452.
- [16] S. Ambec, M. A. Cohen, S. Elgie, and P. Lanoie, "The porter hypothesis at 20: Can environmental regulation enhance innovation and competitiveness?," *Rev Environ Econ Policy*, vol. 7, no. 1, pp. 2–22, 2013, doi: 10.1093/reep/res016.
- [17] P. Bialowolski and D. Weziak-Bialowolska, "External factors affecting investment decisions of companies," *Economics*, vol. 8, no. 1, Mar. 2014, doi: 10.5018/economics-ejournal.ja.2014-11.
- [18] H. Hasanudin, A. Nurwulandari, and R. K. Safitri, "Pengaruh Pengetahuan Investasi, Motivasi Dan Pelatihan Pasar Modal Terhadap Keputusan Investasi Yang Dimediasi Oleh Minat Investas," *Jurnal Ilmiah Managemen, Ekonomi, & Akuntansi (MEA)*, vol. 5, no. 3, pp. 494–512, 2021.