Segmentation of Business Interest using FIMIX-PLS and REBUS-PLS

Hendra H. Dukalang[1], Bambang Widjanarko Otok[2,*], Cindy Cahyaning Astuti[3], and Mohammad Isa Irawan[4]

[1] Department of Sharia Banking, Faculty of Islamic economics and business, IAIN Sultan Amai Gorontalo, Gorontalo 96112, Indonesia
[2] Department of Statistics, Faculty of Science and Data Analytics, Institut Teknologi Sepuluh Nopember, Surabaya 60111, Indonesia
[3] Information Technology Education, Faculty of Psychology and Education, Universitas Muhammadiyah Sidoarjo, Sidoarjo 61215, Indonesia
[4] Department of Mathematics, Faculty of Science and Data Analytics, Institut Teknologi Sepuluh Nopember, Surabaya 60111, Indonesia
*Corresponding author. dr.otok.bw@gmail.com

Abstract. This research aims to determine the heterogeneity of business interest models using the Structural equation model partial least squares (SEM PLS). SEM PLS analysis does not consider the heterogeneity in segments of the observation unit. One analysis that can identify unobserved heterogeneity is finite mix partial least square (FIMIX-PLS) and Response Unit Segmentation Partial Least Square (REBUS PLS). Other research aims to compare the two methods for unobserved heterogeneity in the observation unit. The analysis results show that the FIMIX PLS approach is better than REBUS PLS if seen based on the increase in the R^2 value. If we look at the consistency of the R^2 value, the REBUS-PLS model is the best because each segment has almost the same R^2 value and a value greater than Global R^2.

Keywords: business interest models, SEM-PLS, FIMIX-PLS, REBUS-PLS

1 Introduction

Partial Least Squares Structural Equation Modeling (PLS-SEM) is a statistical method used to analyze the relationship between variables in a complex structural model. The structural model used in the PLS-SEM analysis is a conceptual framework that describes the relationship between variables in a study or research.[1]. The structural model consists of exogenous variables, namely variables not influenced by other variables in the model, and endogenous variables, which are influenced by other variables. PLS-SEM models the relationship between variables, including latent variables, which cannot be measured directly[2].

Unobserved heterogeneity in the context of PLS-SEM analysis refers to variations in the population that cannot be measured or observed directly in the data sample used. This means that differences between individuals or elements in the people are not reflected in the variables observed or measured in the data sample. This heterogeneity is often an

H. S. Panigoro et al. (eds.), *Proceedings of the 2nd International Conference on Sciences, Mathematics, and Education 2023 (ICOSMED 2023)*, Advances in Social Science, Education and Humanities Research 927,
https://doi.org/10.2991/978-2-38476-410-5_7

important issue in statistical analysis because it can affect the analysis results and make the model incompatible with the actual variation in the population. In the PLS-SEM analysis, there is no consideration of heterogeneity that may exist in the observation units on latent variables.
An analysis that can overcome the problem of unobserved heterogeneity in PLS-SEM is the Finite Mixture Partial Least Squares (FIMIX PLS) approach.[3] and A Response-Based Procedure for Detecting Unit Segments in PLS Path Modeling (REBUS PLS). [4]. FIMIX PLS is one of the methods. This method assumes that the population consists of several groups or subpopulations, each with a different structure or pattern of relationships between variables in the structural model. FIMIX PLS is often used when researchers have reason to assume that the population under study is heterogeneous and that the same model does not apply to all individuals or elements in the population.[5]. REBUS-PLS uses a segmentation approach called response-based segmentation, where the segmentation process is carried out based on the unit's response to the model proposed in the PLS-SEM analysis. REBUS-PLS uses a hierarchical approach in the segmentation process, starting with a global model that includes all units. Then, the segmentation process is carried out into several more homogeneous segments. REBUS-PLS evaluates the quality of modeling with and without segmentation. The goal is to determine whether modeling that considers segments significantly improves modeling quality compared to the global model.

FIMIX PLS was first introduced by [3], then some research on the development of theory and application related to FIMIX PLS was carried out by [6] on quality, customer satisfaction, and loyalty data for stores. Furthermore,[7] conducted in marketing data shows that FIMIX PLS produced good segmentation in modeling empirical data in formative and reflective measurement models. In 2016 [8], [9] conducted a further study on segmentation using the FIMIX-PLS approach; the case study demonstrated the ability of FIMIX-PLS to identify whether unobserved heterogeneity significantly affects structural model relationships. In addition, it shows that FIMIX-PLS is very useful for determining the number of segments to be extracted from the data. While research with REBUS PLS analysis was first introduced by [4], several theoretical development and application studies related to REBUS PLS were conducted by [10], which shows that the REBUS PLS approach is more objective and reliable for capturing heterogeneity in consumer behavior. Other research by [11] showed that REBUS PLS analysis is suitable for evaluating the effects of indicators related to labor, tax systems, non-profit organizations, and migrants in Italy.

In this study, a comparison of FIMIX PLS and REBUS PLS analysis will be carried out in overcoming unobserved heterogeneity applied to a case study of the effect of financial literacy on interest in entrepreneurship, where the impact of financial literacy is represented by three variables, namely financial knowledge, financial behavior, financial attitude. The variables that represent the influence of financial literacy are chosen based on the main dimensions of financial literacy. Financial Knowledge relates to understanding basic financial concepts such as interest, inflation, risk diversification, etc. This knowledge is essential for entrepreneurs. This knowledge is essential for entrepreneurs to make the right financial decisions. Financial behavior reflects a person's actions in managing their finances, such as spending, saving, investing, and taking credit. For entrepreneurs, good financial behavior can help run the business smoothly and reduce the risk of financial failure. Financial attitude, i.e., attitude towards money and finance, influences decisions. Someone with a risky attitude may be more open to investing in new business ideas, while a conservative may be more cautious [12–14].

2 Research Methods

The data used in this study is primary data from data collection using questionnaires from Imogen product business actors in Gorontalo Province, which became the observation unit in 2022.

In this paper, business interest consists of 6 indicators, namely self-confidence (Y1), originality (Y2), leadership spirit (Y3), task-oriented (Y4), future-oriented (Y5), and daring to take risks (Y6). The business interest variable is considered an endogenous latent variable that is influenced by the following exogenous latent variables: financial attitude, which consists of five indicators, namely financial basics (X1.1), financial management (X1.2), credit and debt (X1.3), savings and investment (X1.4) and risk and insurance (X1.5). Financial behavior consists of 5 indicators, namely budgeting (X2.1), saving money (X2.2), controlling money (X2.3), investing (X2.4), and paying obligations (X2.5). and financial knowledge consists of interest rates (X3.1), installments (X3.2), financial management (X3.3), investment (X3.4) and financial statements (X3.5)

The research methods used in this study are FIMIX PLS and REBUS PLS, while the research stages of FIMIX PLS and REBUS PLS are as follows:

1. Collect the data used in research consisting of endogenous latent variables symbolized by η (eta) and exogenous latent variables represented by ξ (ksi).
2. We are designing the inner model (relationship between latent variables) and the outer model (relationship between indicators on latent variables).
3. Construct path diagrams following theoretical concepts to explain the relationship between variables.
4. Transforming the path diagram into a system of equations
5. Estimate each model parameter consisting of weight estimates, weight coefficients, path coefficients, and average estimates.
6. Hypothesis testing using bootstrap standard error resampling
7. Evaluate the results of bootstrap resampling on measurement and structural models by comparing the t statistics value with the t table.
8. After the model meets the overall statistical criteria, heterogeneity detection can be done using FIMIX PLS and REBUS PLS.
    a. The FIMIX-PLS algorithm performs try and error to find the number of relevant segments. The recommended number of segments is 2-6 segments. Then, compare the output value based on each segment: the smallest Akaike Information Criterion (AIC) value and the most significant Normed Entropy (EN) value. If the EN value is 0 to 1, the EN value closer to 1 means that the segment class separation is getting better, while if the EN value is less than 0.5, it means that the segment class separation is not good.

b. The REBUS PLS algorithm performs first-stage clustering based on structural and communal residuals from the global model and second-stage clustering on observation units using hierarchy clustering. The clustering process determines the number of clusters according to the optimal number of clusters by the segmentation process using FIMIX PLS, namely k = 3. Furthermore, the clustering results are modeled again using SEM-PLS

9. I was comparing the results of FIMIX-PLS and REBUS PLS by looking at the R value2 in the best segment.

3 Result and Discussion

3.1 Model Specifications

The first step in the SEM model using the PLS algorithm is to make specifications of the measurement model (outer model) and structural model (inner model). The specification model can be seen in Figure 1 below. The PLS algorithm is used to estimate latent variable scores and model parameters.

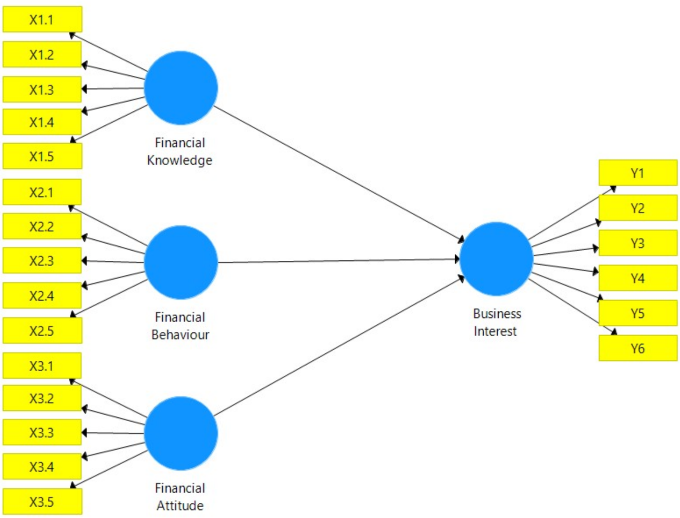

Fig. 1. Model Specification of Business Interest.

3.2 Measurement Model Parameter Estimation

The feasibility of a measurement model can be seen from the value of the loading factor of more than 0.5. And the —t-stat value is greater than the t-table. The test results can be seen in Table 1.

Table 1. Factor Loading Value

Indicator	Factor Loading	t-Statistics	P Values
Financial Knowledge (X1)			
X1.1	0.644	7.968	0.000*
X1.2	0.615	6.701	0.000*
X1.3	0.818	15.170	0.000*
X1.4	0.695	10.758	0.000*
X1.5	0.849	40.084	0.000*
Financial Behavior (X2)			
X2.1	0.812	16.817	0.000*
X2.2	0.682	8.929	0.000*
X2.3	0.666	9.215	0.000*
X2.4	0.712	11.493	0.000*
X2.5	0.831	28.425	0.000*
Financial Attitude (X3)			
X3.1	0.716	8.702	0.000*
X3.2	0.843	22.277	0.000*
X3.3	0.858	26.556	0.000*
X3.4	0.728	12.806	0.000*
X3.5	0.718	12.538	0.000*
Business Interests			
Y1	0.887	34.138	0.000*
Y2	0.863	24.531	0.000*
Y3	0.880	29.845	0.000*
Y4	0.887	35.429	0.000*
Y5	0.806	18.599	0.000*
Y6	0.751	14.036	0.000*

* Significant at $\alpha = 5\%$

Based on table 1. It is obtained that the value of the loading factor is greater than 0.5, and the t-stat value is greater than the t table (1.98). So that all latent variables are considered valid to describe each construct variable. For the reliability test using Cronbach's Alpha or composite reliability, interpreted the same as Cronbach's Alpha, the Cronbach Alpha and Composite reliability values for each variable are more than 0.7. This indicates that the indicators set for each variable have been able to measure the latent variable. The higher correlation between indicators that make up the construct (latent) indicates a better convergent validity value. This can be seen by the AVE value of more than 0.5 for each variable.

3.3 Structural Model Parameter Estimation

The structural model is used to predict the causal relationship between variables. The test results for the structural equation model of Business interest can be seen in Figure 2. The test results with the bootstrapping process can be seen in Table 2.

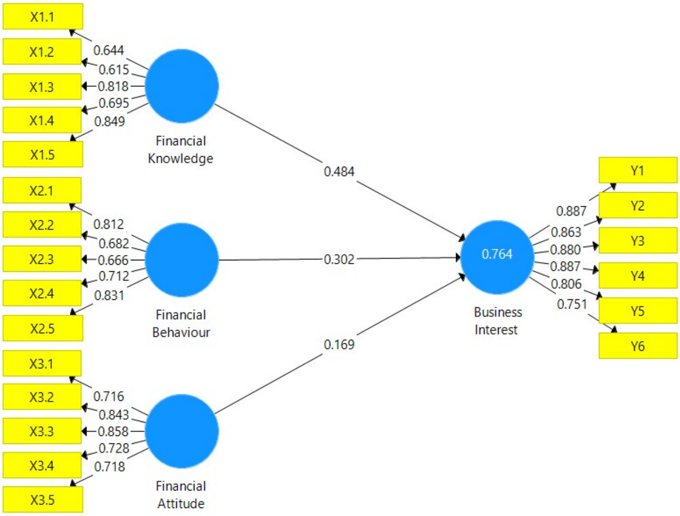

Fig. 2. Structural Equation Model of Business Interest.

Table 2. Path Coefficient Results

Variable	Path Coef.	Standard Error	t-stat	P-values
$X1 \rightarrow Y$	0.169	0.084	2.006	0.045*
$X2 \rightarrow Y$	0.302	0.076	3.954	0.000*
$X3 \rightarrow Y$	0.484	0.073	6.655	0.000*

* Significant at $\alpha = 5\%$

Table 3 shows that each exogenous variable significantly affects endogenous variables with a significant level of 5%. The feasibility of the structural model using R2. In this study, the R2 value was 0.764 . This figure explains that the variability of endogenous variables that the variability of exogenous variables can explain is 76.40%. Thus, the structural equation model of this study is as follows:

$$Y = 0,484X_1 + 0,302X_2 + 0,169X_3 \quad (1)$$

The low value of the coefficient of determination of the global SEM model indicates that the assumption that all observations follow only one model is incorrect. Therefore, finding the number of object segments involved in this study is essential. To find the right segment, we used FIMIX-PLS and REBUS-PLS. In these methods, the segments found are characterized by different structural model coefficients.

3.4 Segmentation with FIMIX PLS

The results obtained based on the analysis of FIMIXPLS are the formation of the number of groups based on predetermined criteria, namely the AIC value and EN (Normed Entropy).

Table 3. Criteria for Determining the Best Segment

	K = 2	**K = 3**	**K = 4**	**K = 5**
AIC	143,488	135,219	145,219	155,219
N.E	0.324	0.527	0.3555	0.329

Based on Table 3, the smallest AIC value is at k = 3 It has the most prominent NE of 0.527 . This indicates that the best segment that is formed is 3 . Furthermore, the results of the percentage of data grouping in segment three can be seen in Table 4.

Table 4. Segment Size

Group	**Segment 1**	**Segment 2**	**Segment 3**
Percentage Respondent	57.6	27.8	14.8

In the number of 3 segments, the percentage of respondents in each segment 1 is 57.6% of the total number of respondents. Segment 2 has a segment size of 27.8% of the respondents. Meanwhile, segment 3 is 0.148 or 14.8% of the total respondents. Next is to determine the value of the Path coefficient on segments 1, 2, and 3, presented in Table 5.

FIMIX PLS on Segment 1

$$Y = 0,166X_1 + 0,260X_2 + 0,424X_3 \tag{2}$$

FIMIX PLS on Segment 2

$$Y = -0,020X_1 + 0,408X_2 + 0,663X_3 \tag{3}$$

Table 5. Path Coefficient for 3 FIMIX-PLS segments

Variable	Global PLS	Segment 1	Segment 2	Segment 3
Financial Attitude (X1)	0.169	0.861	0.509	0.343
Financial Behavior (X2)	0.302	0.846	0.139	0.477
Financial Knowledge (X3)	0.484	0.509	0.350	0.149

FIMIX PLS on Segment 3

$$Y = 0,991X_1 - 0,018X_2 + 0,027X_3 \tag{4}$$

Models 2, 3, and 4. indicate that in each segment in the FIMIX-PLS model, the coefficient of each predictor variable is different. This is influenced by the number of respondents in each segment, which is different, so the coefficient is also different.

3.5 Segmentation with REBUS PLS

Based on results from the FIMIX PLS model that m Based on the results of the FIMIX PLS model, the best model using three segments, the REBUS PLS model also uses three segments. The results of the study can be seen in Table 6.

Table 6. Comparison of Financial Variables (X1, X2, X3) across Segments

Variable	Global PLS	Segment 1	Segment 2	Segment 3
Financial Attitude (X1)	0.169	0.861	0.509	0.343
Financial Behavior (X2)	0.302	0.846	0.139	0.477
Financial Knowledge (X3)	0.484	0.509	0.350	0.149

REBUS PLS in Segment 1

$$Y = 0,861X_1 + 0,846X_2 + 0,509X_3 \tag{5}$$

REBUS PLS in Segment 2

$$Y = 0,509X_1 + 0,139X_2 + 0,350X_3 \tag{6}$$

REBUS PLS in Segment 3

$$Y = 0,343X_1 + 0,477X_2 + 0,149X_3 \tag{7}$$

Models 5, 6, and 7. indicate that in each segment in the REBUS-PLS model, the coefficient of each predictor variable is also different. This is influenced by the number of respondents in each segment, so the coefficient is also different as in the FIMIX-PLS model. Furthermore, a significant comparison of each segment's FIMI-PLS and REBUS-PLS models will be seen.

3.6 Evaluation of FIMIX-PLS and REBUS-PLS Models

Based on the analy12sis using the FIMIX PLS and REBUS PLS models, the results of R Square on the Global Model, Segment 1, Segment 2, and Segment 3 are presented in Table 7 as follows:

Table 7. Comparison of Financial Variables between Global PLS, FIMIX PLS, and REBUS PLS models

Variable	Global PLS	FIMIX PLS			REBUS PLS		
		Segm 1	Segm 2	Segm 3	Segm 1	Segm 2	Segm 3
Financial Attitude (X1)	2,006	0.573	0.050	8,820	5,629	2,247	2,212
Financial Behavior (X2)	3,954	0.792	8,904	0.083	0.882	4,311	4,060
Financial Knowledge (X3)	6,655	3,072	15,126	0.286	2,225	1,369	1,263
R^2	0.764	0.590	0.969	0.999	0.861	0.802	0.802

Table 7 is the t-statistics value of each variable in Global PLS, FIMIX PLS, and REBUS PLS in the t-statistics value in Table 7 compared to the $t_{(-0.95,95)} = 1.98$ so that it can be seen that the exogenous variables that have a significant effect on business interest in each segment. Modeling using Global PLS variables X1, X2, and X 3 has a significant impact on Y . Modeling using FIMIX PLS in segment one variable X3 has a significant effect on Y, in segment two variables X2 and X3 has a significant effect on Y, and in segment three variable X 1 has a significant impact on Y. Modeling using REBUS PLS in segment one variable X1 and X2 have a significant effect on Y, in segment 2 and 3 variables X1, X2 and X3 have a significant impact on Y. Modeling using REBUS PLS in segment one variables X1 and X2 have a significant effect on Y.

Segmentation results using FIMIX-PLS and REBUS-PLS can generally increase the value of R^2 compared to the Global PLS model. However, in segment 1, the FIMIX PLS model experienced a decrease in the value of R^2. Segmentation results using REBUS-PLS are more consistent than segmentation using FIMIX-PLS. So, it can be concluded that segmentation using REBUS- PLS is better than FIMIX- PLS.

4 Conclusion

In this study, the assumption that only one structural model exists for all research data is irrelevant. Attempting all respondents to follow only the structural model will result in an invalid model. This study is proof that it is very likely that in a population, there is heterogeneity in the SEM model. The results show that there are three viable segments to classify respondents who have business interests. The structural model using FIMIX-PLS has a coefficient of determination of more than 90% in each segment. Meanwhile, the structural model using REBUS-PLS has a coefficient of determination of more than 80% in each segment. This shows that the structural model in each segment has a good fit.

When viewed from the magnitude of the R2 value for the global model and the FIMIX-PLS and REBUS-PLS models, the segmentation process on data containing unobserved heterogeneity can increase the R2 value compared to the Global PLS model. Segmentation results using REBUS-PLS are more consistent than segmentation using FIMIX-PLS. So, it can be concluded that segmentation using REBUS-PLS is better than FIMIX-PLS.

References

1. B. Otok, D. Sustrami, P. Hastuti, S. Purnami, Sutikno, A. Suharsono, International Journal of Civil Engineering and Technology **9**(12), 926–938 (2018)
2. A. Rumengan, J. Rumengan, C. Wibisono, B. Otok, International Journal of Mechanical Engineering and Technology (IJMET **9**(10), 632–644 (2018)
3. K. Jedidi, H. Jagpal, W. Desarbo, Marketing Science **16**(1), 39–59 (1997)
4. V. Vinzi, L. Trinchera, S. Squillacciotti, M. Tenenhaus, in *Applied Stochastic Models in Business and Industry* (2008), p. 439–458
5. M. Mukid, B. Otok, A. Suharsono. Segmentation of public health status using finite
6. C. Hahn, M. Johnson, A. Herrmann, F. Huber, Schmalenbach Business Review **54**, 243–259 (2002)
7. M. Sarstedt, C. Ringle, J Appl Stat **37**(8), 1299–1318, (2010)
8. J. Hair, M. Sarstedt, L. Matthews, C. Ringle, European Business Review **28**(1), 63–76 (2016)
9. L. Matthews, M. Sarstedt, J. Hair, C. Ringle, European Business Review **28**(2), 208–224 (2016)
10. M. Mehmetoglu, Journal of Targeting, Measurement and Analysis for Marketing **19**(3–4), 165–172, (2011)
11. C. Quintano, P. Mazzocchi, Journal of Economic Studies **47**(2), 405–430, (2020)
12. T. Sugiyanto, W. Radianto, T. Efrata, L. Dewi, Advances in Economics, Business and Management Research **353–358** (2019)
13. K. Rai, S. Dua, M. Yadav, FIIB Business Review **8**(1), 51–60, (2019)
14. R. Yogasnumurti, I. Sadalia, N. Irawati, in *Proceedings of the 2nd Economics and Business International Conference, Scitepress* (2021), p. 649–657