Variance
Journal of Statistics and Its Applications
P-ISSN: 2685-8738 E-ISSN: 2685-872X
Variance: Journal of Statistics and Its Applications
Volume 6 Issue 2 | October 2024
10.30598/variancevol6iss2page229-238

SEGMENTATION OF FRESH GRADUATES' JOB INTEREST AND MOTIVATION BASED ON FINITE MIXTURE PARTIAL LEAST SQUARES (FIMIX-PLS)

Mohamad David Hermawan[1], Ardi Kurniawan[2*], M. Fariz Fadillah Mardianto[3], Sediono[4]
[1,2,3,4] Department of Mathematics, Faculty of Science and Technology, Universitas Airlangga
Merr C Campus, Jl. Dr. Ir. H. Soekarno, Mulyorejo, Mulyorejo District, Surabaya, 60115, East Java , Indonesia

Corresponding Author's E-mail: *ardi-k@fst.unair.ac.id*

Abstract: *Rapid technological advancements have significantly transformed the global labor market, impacting various industries and workers in Indonesia who may need to be adequately prepared to adapt. The current landscape demands individuals who can acquire knowledge quickly and adapt to modern technologies. The unemployment rate in Indonesia, especially among fresh graduates, is still a concern. Lack of motivation and interest in finding a job and high expectations of working conditions contribute to this problem. This research aims to address the gap in research by using the Finite Mixture Partial Least Squares (FIMIX-PLS) approach to examine the segmentation of fresh graduate characteristics about their interest and motivation in finding a job. Segmentation based on latent variable relationships in the structural model can be overcome with Finite Mixture Partial Least Square (FIMIX-PLS) to identify more homogeneous characteristics. This research analyzes explicitly the impact of compensation, work environment, and company reputation on the interest and motivation of fresh graduates in finding a job. This research resulted in the best segmentation of two segments: the 1st segment at 77.8% (265 samples) and the 2nd segment at 22.2% (75 samples).*

Keywords: *FIMIX-PLS, Fresh graduate, Interest and Motivation, Job Search, Segmentation.*

1. INTRODUCTION

Technological advances in artificial intelligence, mechanization, and automation have transformed the global labor market, affecting industries and workers in Indonesia who are unprepared to adapt. Companies now need individuals who are quick to learn and adapt to new technologies [1]. According to DBS, 48% of companies in Indonesia will undergo digital transformation by 2022, and this number is expected to continue to rise [2]. This transformation is changing the way employees are recruited and retained as well as how individuals search for jobs that match their qualifications. The Open Unemployment Rate (OER) in Indonesia in August 2023 was 5.32%, with the majority of the unemployed aged 20-24 years old [3], [4]. Lack of motivation and interest in finding a job, coupled with high expectations of working conditions, are contributing factors to unemployment among New Graduates [5].

In multivariate analysis, SEM can simultaneously test and estimate the relationship between one or more dependent variables and multiple factors [6]. Many researchers have utilized SEM because it can estimate complex models with many structural paths, constructs, and indicators without paying attention to distribution assumptions on the data. SEM has two types of models, namely structural models and measurement models. The structural model describes the relationship between latent variables. In estimating structural equation models, it is often assumed that the data collected is homogeneous. However, this assumption is sometimes rational if the data is obtained from several units or groups. Finite Mixture Partial Least Square (FIMIX-PLS) is an alternative approach solution that can be utilized to identify diversity in the relationship pattern of linear path models [7]. FIMIX-PLS addresses inhomogeneity or heterogeneity by estimating segment probabilities for each observation and simultaneously estimating path coefficients for all segments [8].

Submitted: 2024-08-07 *Accepted: 2024-10-30*

In previous research, it was stated that compensation has a positive influence on interest and motivation to find a job [9]. In contrast, a company's reputation does not affect interest and motivation to find a job. On the other hand, in Pragita's research (2023), using the same method, namely multiple linear regression, the results explain that compensation and company reputation significantly affect the intention of fresh graduates to find a job [10]. Alifuddin's research (2023) with the Structural Equation Modeling (SEM) method shows that compensation and company reputation significantly have a positive effect on interest in finding a job [11]. However, there has yet to be research using the Finite Mixture Partial Least Square (FIMIX-PLS) approach to identify the segmentation of new graduate characteristics related to interest and motivation to find a job.

This research aims to fill this void by analyzing the effect of compensation, work environment, and company reputation on interest and motivation to seek employment using the FIMIX-PLS approach. The results of this research can help formulate strategic policies for companies and universities and provide insight to the public regarding the factors that influence the interest and motivation of new graduates to find a job.

2. METHODOLOGY

2.1. Data Source

This research uses primary data or data from researchers through survey questionnaires. This research took place in public and private universities in Surabaya. The survey was conducted by distributing questionnaires online through social media platforms. The questionnaire was distributed from June 2024 to July 2024.

2.2. Research Variables

This research examines latent variables, which include endogenous variables such as Motivation for Finding a Job (η_1) and Interest in Finding a Job (η_2), and the mediating variable of company reputation (η_3). Meanwhile, the exogenous variables include compensation (ξ_1) and work environment (ξ_2). This research uses latent variables and indicators, which are presented in Table 1.

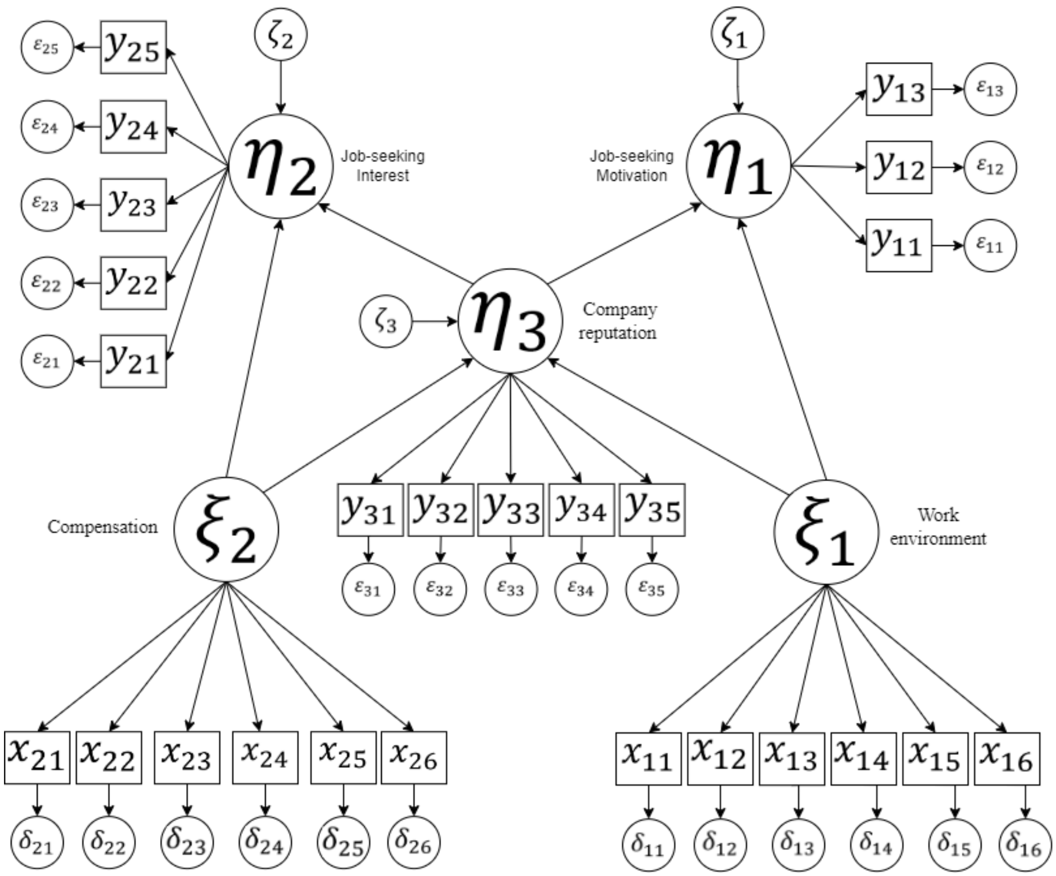
Table 1. Latent Variables and Indicators

Variable	Indicators
Motivation for Finding a Job (η_1)	The existence of hopes and dreams (y_{11})
	The existence of needs (y_{12})
	self-reliance (y_{13})
Interest in Finding a Job (η_2)	Pay more attention to job openings that are in demand (y_{21})
	A desire to get to know the company of interest better (y_{22})
	Curiosity about jobs in companies of interest (y_{23})
	Feeling happy about a job at a company that is in demand (y_{24})
	There is an emotional attraction to the company of interest (y_{25})
Company reputation (η_3)	Good company image in the eyes of job applicants (y_{31})
	Strong product and service brands (y_{32})
	Company vision and leadership (y_{33})
	Corporate social and environmental responsibility (y_{34})
	Good financial performance of the company (y_{35})
Compensation (ξ_1)	Appropriate salary (x_{11})
	Satisfying incentives (x_{12})
	Career advancement (x_{13})
	Health Insurance (x_{14})
	Allowances (x_{15})

Variable	Indicators
	Old age guarantee (x_{16})
Work environment (ξ_2)	Facilities (x_{21})
	Easy access (x_{22})
	Relationships between colleagues (x_{23})
	Superior-subordinate relationship (x_{24})
	Professional working rules (x_{25})
	Workplace comfort (x_{26})

2.3. Path Diagram

This research has a path diagram provided in Figure 1 as follows.



Picture 1. Research Path Diagram

2.4. Sampling Techniques

This research collected data using a questionnaire with a purposive sampling technique. The ideal minimum sample number should be ten times the number of indicators [9]. In the context of this research, 25 indicators are used. Based on PLS guidelines, the minimum number of samples required is 10 times the number of indicators, which is a minimum of 250 samples [10].

The sample collection approach is carried out by filling out a questionnaire. The researcher has provided answer options for respondents to choose directly from the available alternatives. The participants were given five answer options: very appropriate, appropriate, moderate, not appropriate, and very inappropriate. This process involves formulating a research design and, afterwards, conducting tests to assess the validity and reliability of the instrument.

2.5. Data Analysis Methods

This research uses data analysis techniques Finite Mixture Partial Least Square (FIMIX-PLS) [11]. FIMIX-PLS is an alternative method that can be used to detect variations in the relationship patterns of linear route models [12]. FIMIX-PLS addresses the problem of heterogeneity by calculating the likelihood that each observation belongs to a particular segment and simultaneously estimating the path coefficients for all segments [13]. Before using FIMIX-PLS, SEM-PLS assesses the predicted relationship between constructs by checking for the presence of influences or relationships between constructs (latent or conceptual entities that can be quantified and viewed) [14].

The analysis process of this research involves the following steps. Create a conceptual model consisting of a measurement and structural models [11].

a) Create a path diagram to represent the concept.
b) Formulate based on the path diagram outlined in step 1.

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

c) Perform parameter estimation for measurement models and structural models.
d) Evaluate the model in the following way.
    i. Measurement Model Evaluation
        1) Loading factor

        Value loading factor used to assess the validity of convergence. Value loading factor must exceed 0.7 [12]. There are several criteria for grades outer loading [14]. Outer loading less than 0.4 should be removed because its contribution is low to the model, while if it exceeds 0.7, the indicator is considered acceptable or worthy of use. However, indicators with a value outer loading between 0.4 and 0.7 require further analysis of the impact of elimination on AVE values and Composite Reliability.

        2) Average Variance Extracted (AVE)

        The AVE value should be greater than 0.5, which indicates that 50% or more of the variance of the indicator can be explained [12]. The formula for calculating AVE is as follows:

$$AVE_i = \frac{\sum_{k=1}^{n} \lambda_{ik}^2}{\sum_{k=1}^{n} \lambda_{ik}^2 + \sum_i var(\varepsilon_i)} \tag{2}$$

        3) Composite Reliability

        Latent variable indicators are assessed and measured using internal consistency metrics. This procedure is used to ensure the instrument's accuracy, uniformity, and dependability in determining latent variables. A reliability value is anticipated to be greater than 0.7, although a value of 0.6 is still considered acceptable [12]. The reliability value of the composite can be calculated using the following equation (3) formula.

$$\rho_{ci} = \frac{(\sum_{k=1}^{n} \lambda_{ik})^2}{(\sum_{k=1}^{n} \lambda_{ik}) + \sum_i var(\varepsilon_i)} \tag{3}$$

        4) Cronbach's Alpha

Cronbach's Alpha is used to measure the lower bound of the reliability of a latent variable, which reflects the overall reliability of the indicator in the model. For confirmatory research, Cronbach's Alpha value should be more than 0.7, while for explanatory research, this value should be more than 0.6.

ii. Structural Model Evaluation

Structural models can be assessed by examining the value of R-Squared. This value is used to evaluate the extent to which a particular independent latent variable significantly impacts the dependent latent variable. A value of 0.67 for R-Square indicates a moderately large or good model, while 0.33 indicates a moderate model, and 0.19 indicates a weak model [12]. Formula R-Square are as follows.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \widehat{y_i})^2}{\sum_{i=1}^{n}(y_i - \overline{y_i})^2} \tag{4}$$

e) Hypothesis testing

f) Doing FIMIX-PLS estimation

FIMIX-PLS estimation is performed on the assumption that heterogeneity arises in the structural model because the samples come from different populations, which requires segmentation [15]. FIMIX-PLS uses the results of estimating latent variables and modifying the relationship in the inner model.

$$\boldsymbol{B\eta_i} + \boldsymbol{\Gamma\xi_i} = \boldsymbol{\zeta_i} \tag{5}$$

Assuming that it is distributed as $\boldsymbol{\eta_i}$a finite mixture, then this can be stated as follows.

$$\boldsymbol{\eta}_i \sim \sum_{k=1}^{K} \boldsymbol{\rho}_k \left[\frac{|\boldsymbol{B}_k|}{\sqrt[M]{2\pi}\sqrt{|\boldsymbol{\Psi}_k|}} e^{-\frac{1}{2}(B_k\boldsymbol{\eta}_i + \boldsymbol{\Gamma}_k\xi_i)'\boldsymbol{\Psi}_k^{-1}(B_k\boldsymbol{\eta}_i + \boldsymbol{\Gamma}_k\xi_i)} \right] \tag{6}$$

g) Choosing segment criteria

Selecting criteria based on AIC_3 (modified by a factor of 3) and BIC, as they can show the correct number in 84% of cases where both criteria show the same number of segments AIC_3, BIC [16]. In addition, the value EN (Entropy Criterion) has a specification value above 0.5 and the lowest. Based on these criteria and can be formulated as follows.

$$AIC_3 = -2 \cdot lnL + 3 \cdot N_k \tag{7}$$

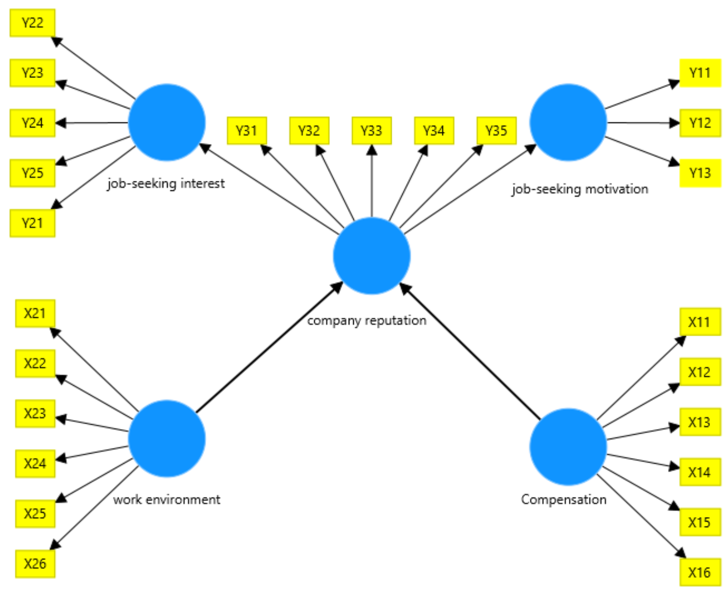
$$BIC = -2 \cdot lnL + ln(N) \cdot N_k \tag{8}$$

$$EN = 1 - \frac{\left[\sum_{i=1}^{I}\sum_{s=1}^{S} -P_{is} \cdot ln\, P_{ik}\right]}{I \cdot ln(K)} \tag{9}$$

3. RESULTS AND DISCUSSION

3.1. SEM-PLS Analysis

The structural model formed based on indicators and significant path relationships is as follows.



Picture 2. Final Model

3.2. Evaluation of Measurement Models

1. Loading factor

Measurement models for formative indicators involve assessment outer loading of each indicator using the bootstrapping to determine its relevance. Significance test for outer loading can be seen in Table 2.

Table 1. Value Outer loading

Variable	Indicators	Outer loading	T Statistics	P Values	Information
Compensation	X11	0.818	23.023	0.000	Significant
	X12	0.784	20.579	0.000	Significant
	X13	0.820	29.642	0.000	Significant
	X14	0.821	22.483	0.000	Significant
	X15	0.792	22.906	0.000	Significant
	X16	0.800	21.860	0.000	Significant
Work Environment	X21	0.839	26.767	0.000	Significant
	X22	0.799	20.226	0.000	Significant
	X23	0.798	22.579	0.000	Significant
	X24	0.828	28.756	0.000	Significant
	X25	0.858	31.112	0.000	Significant
	X26	0.873	37.770	0.000	Significant
Motivation for Finding a Job	Y11	0.817	23.466	0.000	Significant
	Y12	0.763	20.059	0.000	Significant
	Y13	0.814	32.924	0.000	Significant
Interest in Finding a Job	Y21	0.792	21.904	0.000	Significant
	Y22	0.727	16.834	0.000	Significant
	Y23	0.782	23.950	0.000	Significant
	Y24	0.791	24.978	0.000	Significant
	Y25	0.924	63.952	0.000	Significant
Company Reputation	Y31	0.771	18.925	0.000	Significant
	Y32	0.764	17.604	0.000	Significant
	Y33	0.761	19.598	0.000	Significant
	Y34	0.765	18.806	0.000	Significant
	Y35	0.824	22.976	0.000	Significant

From Table 2, it is explained that for the loading factor evaluation, it is proven that all are significant (more than 0.7). This means that these indicators have contributed to the latent variable and measured consistently.

2. Average Variance Extracted (AVE)

The significance test for AVE is as follows.

Table 2. AVE Scores

Variable	Average Variance Extracted (AVE)	T Statistics	P Values	Information
Compensation	0.649	12.874	0.000	Significant
Work Environment	0.694	14.995	0.000	Significant
Interest in Finding a Job	0.649	15.768	0.000	Significant
Motivation for Finding a Job	0.637	15.125	0.000	Significant
Company Reputation	0.604	11.315	0.000	Significant

Table 3 illustrates that the value of each AVE is significant (more than 0.5). This value means that the construct can explain at least 50% of the variance of its indicators, indicating that the construct has sufficient convergent validity.

3. Composite Reliability

The significance test for Composite Reliability is as follows.

Table 3. Value Composite Reliability

Variable	Composite Reliability	T Statistics	P Values	Information
Compensation	0.892	34.97	0.000	Significant
Work Environment	0.913	45.554	0.000	Significant
Interest in Finding a Job	0.864	33.589	0.000	Significant
Motivation for Finding a Job	0.722	13.717	0.000	Significant
Company Reputation	0.837	21.366	0.000	Significant

Table 4 illustrates that the value of each Composite Reliability is significant (more than 0.7). This value indicates that the indicators are quite consistent in measuring the same construct.

4. Cronbach's Alpha

Table 4. Value Cronbach's Alpha

Variable	Cronbach's Alpha	T Statistics	P Values	Information
Compensation	0.892	34.634	0.000	Significant
Work Environment	0.911	44.239	0.000	Significant
Interest in Finding a Job	0.862	32.673	0.000	Significant
Motivation for Finding a Job	0.716	13.46	0.000	Significant
Company Reputation	0.836	21.125	0.000	Significant

Table 5 illustrates that each Cronbach's Alpha value is significant (more than 0.7). A value of 0.7 or higher is generally considered adequate, indicating good internal reliability. Values of 0.8 or more indicate excellent reliability.

3.3. Structural Model Evaluation

Evaluation of the structural model in SEM-PLS involves assessing the goodness of the model by using the R^2 determination coefficient for each endogenous latent variable, which represents the predictive power of the model. The R^2 values are obtained as follows:

Table 5. Value R-Square

Variable	R-Square
Interest in Finding a Job	0.411
Motivation for Finding a Job	0.568
Company Reputation	0.782

From the resulting R^2 values, it is indicated that the model is classified as a moderate model (more than 0.333 less than 0.67). Evaluation of the structural model can also be done by looking at the path parameter coefficients (path coefficients) between latent variables.

Table 6. Path Coefficient Value

Relationship	Path coefficients
Compensation -> Company Reputation	0.319
Work Environment -> Company Reputation	0.589
Company Reputation -> Interest in Finding a Job	0.641
Company Reputation -> Motivation for Finding a Job	0.754

From Table 7, the exogenous variable (Compensation) to the mediating variable (Company Reputation) has a value of 0.319, and the exogenous variable (Work Environment) to the mediating variable (Company Reputation) has a value of 0.589. Therefore, the direction of the relationship between the two is positive. On the other hand, the mediating variable (Company Reputation) on the endogenous variable (Interest in Finding a Job) has a value of 0.641. It has a value of 0.754 on the endogenous variable (Motivation for Finding a Job).

3.4. Segmentation Analysis with FIMIX-PLS

Results from the FIMIX-PLS analysis using SmartPLS 4 is the formation of the number of segments based on several criteria that have been determined.

Table 7. AIC3, BIC, and EN Criteria

Criterion	1 segment (without segmentation)	2 segments
AIC3	1932.023	1361.176
BIC	1951.826	1403.61
EN	0	0.968

Based on Table 8, a comparison of 2-segment models has better AIC3, BIC, and EN values (smaller, EN > 0.5) compared to unsegmented models. Segmentation is obtained from the probability value of each segment membership. Here are the segment sizes (segment size) from the analysis carried out.

Table 8. Segment Size (segment size) in percent

	Segment1	Segment2
%	0.778	0.222

Table 9 shows that the size of the 1st segment is 77.8% (265 samples), and the 2nd segment is 22.2% (75 samples). Segment 1 (77.8%): This group is likely to have more consistent relationships between the variables in the model, according to the overall relationship structure in the path coefficient table. As this segment accounts for most of the sample, the general pattern or relationship between the variables in the model is more characteristic of this segment. Segment 2 (22.2%): This smaller segment may have different characteristics when assessing the relationships between variables, e.g., the preferences or factors that influence their interest and motivation to find a job may differ from Segment 1. Heterogeneity in this segment may indicate the presence of unique groups in the data that respond differently to the variables.

4. CONCLUSION

Based on the results of the discussion, there are 25 significant indicators. In this research model, the R2 level is 41.1% for the interest in finding a job variable and 56.8% for the motivation variable to find a job where this figure has meaning for the moderate model. In the structural equation model path, there are 4 significant paths with 2 paths being mediation, the path is the influence of compensation on the company's reputation, the work environment on the company's reputation, the company's reputation on the interest in finding a job, and the company's reputation on the motivation to find a job. The segmentation formed through FIMIX-PLS analysis based on the AIC3, BIC, and EN criteria formed 2 segmentations with the 1st segment 77.8% (265 samples) and the 2nd segment 22.2% (75 samples).

REFERENCES

[1] L. Saragih, “Identifikasi Dampak Perkembangan Teknologi Terhadap Tenaga Kerja Toko Ritel Indonesia: Studi Kasus Toko X,” *Jurnal Kependudukan Indonesia*, vol. 14, no. 1, pp. 13–28, Jun. 2019.

[2] DBS, “Key Change,” Jakarta, 2022.

[3] Badan Pusat Statistik, “Keadaan Angkatan Kerja Di Indonesia Agustus 2023,” Jakarta, 2023.

[4] Badan Pusat Statistik, “Keadaan Ketenagakerjaan Indonesia Agustus 2023,” Jakarta, 2023.

[5] P. Supriyoso and W. Muamadah, “Analisis Disiplin Kerja dan Motivasi Kinerja Karyawan PT. Piala Sutra Wida Muamadah,” *Prosiding FRIMA (Festival Riset Ilmiah Manajemen & Akuntansi)*, no. 4, pp. 417–423, 2021.

[6] D. I. Febrianti and C. Hendratmoko, “Pengaruh Kompensasi, Karakteristik Pekerjaan, Tanggung Jawab Sosial Perusahaan dan Reputasi Perusahaan terhadap Minat Melamar Pekerjaan,” *Kinerja: Jurnal Ekonomi dan Manajemen*, vol. 19, no. 2, pp. 257–267, 2022.

[7] M. P. Pragita, “Pengaruh Reputasi Perusahaan Dan Kompensasi Terhadap Intensi Melamar Kerja Pada Generasi Z (Studi Kasus Pada Fresh Graduate Ilmu Administrasi Bisnis Universitas Lampung),” Undergraduate Thesis, Universitas Lampung, Fakultas Ilmu Sosia dan Ilmu Politik Univesitas Lampung, 2023.

[8] P. R. M. Alifuddin, “Pengaruh Reputasi, E-Recruitment, Dan Lingkungan Kerja Terhadap Minat Melamar Pekerjaan Oleh Mahasiswa,” Dissertation, STIE Indonesia Banking School, Jakarta Selatan, 2023.

[9] J. F. Hair, G. T. M. Hult, C. M. Ringle, M. Sarstedt, N. P. Danks, and S. Ray, *Partial Least Squares Structural Equation Modeling (PLS-SEM) Using R: A Workbook*. in Classroom Companion: Business. Cham: Springer International Publishing, 2021.

[10] S. Haryono and P. Wardoyo, *Structural Equation Modelling Untuk Penelitian Manajemen Menggunakan AMOS 18.00*. Bekasi: PT Intermedia Personalia Utama, 2012.

[11] J. Hair Jr, J. F. Hair Jr, M. Sarstedt, C. M. Ringle, and S. P. Gudergan, *Advanced issues in partial least squares structural equation modeling*. Sage Publications, 2023.

[12] I. Ghozali, *Structural Equation Modeling Metode Alternatif dengan Partial Least Square (PLS)*, 4th ed. Badan Penerbit Universitas Dipenegoro, 2014.

[13] C. Hahn, M. D. Johnson, A. Herrmann, and F. Huber, “Capturing Customer Heterogeneity using a Finite Mixture PLS Approach,” *Schmalenbach Bus Rev*, vol. 54, no. 3, pp. 243–269, Jul. 2002.

[14] J. F. Hair, G. T. M. Hult, C. M. Ringle, and M. Sarstedt, *A primer on partial least squares structural equation modeling (PLS-SEM)*, 3rd ed. Los Angeles London New Delhi Singapore Washington DC Melbourne: SAGE, 2022.

[15] C. M. Ringle, S. Wende, and A. Will, “The Finite Mixture Partial Least Squares Approach - Methodology and Application,” 2006.

[16] M. Sarstedt, J.-M. Becker, C. M. Ringle, and M. Schwaiger, “Uncovering and Treating Unobserved Heterogeneity with FIMIX-PLS: Which Model Selection Criterion Provides an Appropriate Number of Segments?,” *Schmalenbach Bus Rev*, vol. 63, no. 1, pp. 34–62, Jan. 2011.