An Analysis of Vertical Mismatch of College Graduates in Central Java in 2025 from a Gender Perspective Using a Multinomial Logistic Regression

Camila Nurahma[(1)], Arief Agoestanto[(2)]

Program Studi Statistika Terapan dan Komputasi, Fakultas MIPA, Universitas Negeri Semarang

Jl. Raya Banaran, Sekaran, Kec. Gunung Pati, Kota Semarang, Jawa Tengah 50229
e-mail: camilanurahma@gmail.com[(1)], agoestanto.aa@gmail.com[(2)]

ABSTRAK

Ketidaksesuaian vertikal merupakan kondisi ketika tingkat pendidikan individu tidak sesuai dengan kualifikasi pekerjaan yang dijalani. Meskipun berbagai penelitian telah mengkaji faktor – faktor yang memengaruhi ketidaksesuaian vertikal, kajian mengenai perbedaan gender pada lulusan perguruan tinggi di Provinsi Jawa Tengah masih terbatas. Penelitian ini bertujuan untuk menganalisis pengaruh gender terhadap ketidaksesuaian vertikal lulusan perguruan tinggi di Provinsi Jawa Tengah tahun 2025. Data yang digunakan merupakan data mikro Survei Angkatan Kerja Nasional (Sakernas) Tahun 2025 dengan sampel sebanyak 4.100 lulusan perguruan tinggi yang bekerja di Provinsi Jawa Tengah. Analisis dilakukan menggunakan regresi logistik multinomial dengan kategori match sebagai kategori referensi dan Average Marginal Effects (AME) untuk menginterpretasikan perubahan probabilitas pada setiap kategori. Hasil penelitian menunjukkan bahwa overeducation merupakan bentuk ketidaksesuaian vertikal yang paling dominan. Gender berpengaruh signifikan terhadap status ketidaksesuaian vertikal setelah dikontrol oleh umur, status perkawinan, pendidikan tertinggi, klasifikasi tempat tinggal, upah, dan jam kerja. Hasil AME menunjukkan bahwa perempuan memiliki probabilitas mengalami overeducation yang lebih rendah sebesar 16,11%, probabilitas berada pada kondisi match yang lebih tinggi sebesar 14,43%, dan probabilitas mengalami undereducation yang lebih tinggi sebesar 1,68% dibandingkan laki-laki. Temuan penelitian ini dapat menjadi masukan bagi perguruan tinggi dan pembuat kebijakan dalam memperkuat keterkaitan antara pendidikan tinggi dan kebutuhan pasar kerja dengan mempertimbangkan perbedaan gender.

Kata kunci: Ketidaksesuaian Vertikal; Gender; Regresi Logistik Multinomial; Average Marginal Effects

ABSTRACT

Vertical mismatch is a condition in which an individual's level of education does not match the qualifications of the job being undertaken. Although various studies have examined the factors affecting vertical mismatch, empirical evidence regarding gender differences in college graduates at the provincial level is still limited, especially in Central Java Province. This research offers novelty by examining gender differences in match status, overeducation, and undereducation at the provincial level using multinomial logistic regression and Average Marginal Effects (AME). This research aims to analyze the effect of gender on the vertical mismatch of college graduates in Central Java Province in 2025. The data used were Sakernas microdata in 2025. The sample was 4,100 college graduates who worked in Central Java Province. The data analysis was carried out using multinomial logistic regression and AME. The research results showed that overeducation was the most dominant form of vertical maladjustment. Gender had a significant effect on the probability of an individual being in each category of vertical nonconformity. The AME results showed that women had a lower probability of experiencing overeducation of 16.11%, a higher probability of being in a match condition of 14.43%, and a higher probability of experiencing undereducation of 1.68% compared to men. These findings can provide input for universities and policymakers in strengthening the link between higher education and labor market needs by considering gender differences.

Keywords: *Vertical Mismatch; Gender; Multinomial Logistic Regression; Average Marginal Effect*

INTRODUCTION

Vertical mismatch between educational attainment and employment remains a key issue in the Indonesian labor market. The data from the Central Statistics Agency (2025) showed that 36.36% of the workforce experienced vertical mismatch, consisting of 22.36% overeducation and 13% undereducation. This indicates that a substantial share of workers are employed in positions that do not correspond to their educational qualifications. From the perspective of Human Capital Theory, education is viewed as an investment in human capital that enhances individual's productivity, skill, and competencies, thereby increasing their opportunities to obtain employment that matches their educational qualifications [1]. However, improvements in educational attainment do not automatically lead to suitable job placement because labor demand, occupational structure, and job allocation also shape the outcomes achieved by the workforce. Labor Market Segmentation Theory explains that the labor markets are divided into primary and secondary segment, each characterized by different employment conditions, wages, and career opportunities [2]. Consequently, individuals with similar education levels may face different opportunities to obtain jobs that match their qualifications, including differences between men and women. These differences in opportunities can occur across various groups of workers, including men and women. In labor economics literature, this phenomenon is generally divided into overeducation and undereducation. Overeducation occurs when an individual's education level exceeds the qualifications required for the job they hold, while undereducation occurs when an individual's education level falls below the job requirements [3].

Central Java Province provides an important regional context for analyzing vertical mismatch among college graduates. In 2025, Central Java had 339 universities with a total of 277,776 graduates, with a higher proportion of women than men [4]. The Open Unemployment Rate for university graduates remains at 4.66% [5] while the informal sector still dominated the Central Java labor market at 59.64% [6]. This combination of high graduate output, graduate unemployment, and an informal-sector-dominated labor market indicates that the growing number of university graduates entering the labor market has not been fully matched by the availability of jobs that correspond to their educational qualifications. Furthermore, there are indications that men and women face different labor market conditions. This difference is seen in the Open Unemployment Rate for university graduates by gender and educational level, as presented in Figure 1.

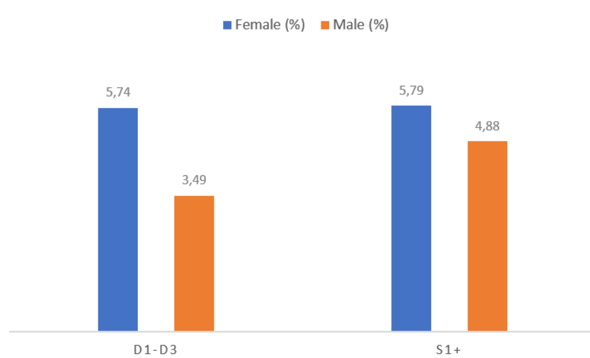

Figure 1. Unemployment Rate Among Higher Education Graduates by Gender and Educational Level, Central Java, 2025

According to Figure 1, the Open Unemployment Rate for women at both the diploma and undergraduate levels is higher than that of men. At the diploma level, the Open Unemployment Rate (TPT) for women reached 5.74%, while for men it was 3.49%. Meanwhile, at the undergraduate level and above, the TPT for women was 5.79%, while for men it was 4.88%. This difference indicates that men and women face unequal labor market conditions. Despite having the same level of education, men and women may have different opportunities to obtain jobs that match their qualifications [7]. This reinforces the notion that opportunities for suitable employment are not evenly distributed between genders in Central Java Province. Previous research also indicates that differences in labor market conditions between men and women can influence the likelihood of vertical mismatch [8][9].

However, empirical evidence about gender differences in vertical mismatch among university graduates at the provincial level is still limited, especially in Central Java Province. Research in Indonesia generally focuses on identifying factors influencing wages and their impact on wages [10][11], while differences in opportunities for vertical mismatch between men and women have rarely been discussed in depth. However, there are still few studies that specifically examine gender differences in the three categories of vertical mismatch (match, overeducation, and undereducation) at the provincial level and present the results in probability form. Therefore, this study extends the existing literature by placing gender as the main explanatory variable, not just one of the control variables. In addition, this research simultaneously analyzes the probability of individuals being in the match, overeducation, and undereducation categories using a multinomial logistic regression model, as well as using Average Marginal Effects (AME) to provide a more intuitive interpretation of probability differences based on gender.

Therefore, to fill this research gap, this research used the data from the August 2025 National Labor Force Survey for Central Java Province to analyze the distribution of vertical mismatches based on gender, examine the effect of gender on the probability of experiencing overeducation and undereducation after controlling for individual and job characteristics, and estimate the changes in probability in each category of mismatch. This research contributes to the literature on gender differences in vertical mismatch among college graduates at the provincial level by showing how Average Marginal Effects (AME) make it easier to interpret probabilities from multinomial logistic regression. The findings of this research also provide practical implications for universities and local governments in designing matching strategies between education and employment that are more responsive to gender differences in the labor market.

METHOD

Data and Data Sources

This research used secondary data from the August 2025 National Labor Force Survey held by the Central Statistics Agency. Sakernas is a household survey conducted by the Central Statistics Agency to provide data and information regarding employment conditions in Indonesia [12]. The unit of analysis in this research was the working population of college graduates in Central Java Province. The sample was restricted to employed individuals with educational attainment ranging from diploma to doctoral level. The research sample was limited to individuals who had complete information for all variables used in the analysis. Applying these criteria resulted in a final sample of 4,100 employed college graduates in Central Java Province. The workers in the KBJI 0 position group, which included the members of the Indonesian National Army and the Indonesian National Police, were not included in the analysis. It was because the recruitment and job placement process in this position group did not fully reflect the labor market mechanisms applying to the majority of the workforce [13][14]. Then, the data processing and analysis were carried out using RStudio software.

The final sample for this study consisted of 4,100 employed college graduates in Central Java Province with complete information on all variables included in the analysis, namely vertical mismatch status, gender, age, marital status, highest educational attainment, residential classification, wages, and working hours. Therefore, this study did not require statistical imputation of the data.

The descriptive statistics used Sakernas sampling weights to produce representative estimates of the population of college graduates employed in Central Java Province. Meanwhile, the multinomial logistic regression model was estimated without using sample weights because the purpose of the regression analysis was to examine the relationship between gender and vertical mismatch, not to estimate population characteristics. As explained by Lumley and Scott [15], the use of weights in regression analysis depends on the purpose of the inference and the characteristics of the survey design. Therefore, the regression coefficients and Average Marginal Effects (AME) are interpreted as conditional relationships within the analysis sample, while the descriptive statistics represent the target population.

Research Variables

The variables used in this study consisted of a dependent variable, a primary independent variable, and the independent control variables. The dependent variable in this study was vertical mismatch status, while gender served as the primary independent variable. In addition, this study also used several control variables: age, marital status, education level, residential location, wages, and working hours. Operational definitions of each variable are presented in Table 1.

Table 1. Research Variables

Dependent Variable	Y	Status Mismatch
Primary Independent Variable	X_1	Gender
Control Independent Variables	X_2	Age
	X_3	Marital Status
	X_4	Highest Level of Education
	X_5	Residential Classification
	X_6	Wages
	X_7	Working Hours

Vertical Mismatch Classification

Vertical mismatch classification was conducted using a job analysis approach, referring to the job skill level classification according to the International Standard Classification of Occupations (ISCO-08), which was linked to the education level according to the International Standard Classification of Education (ISCED). Because the Sakernas's data used the 2020 Indonesian Standard Classification of Occupations (KBJI), each occupational group was mapped to an ISCO level to obtain the education level generally required for each occupation. This mapping is presented in Table 2 [16]. The workers in KBJI job group 0, which included the members of the Indonesian National Armed Forces and the Indonesian National Police, were not included in the analysis.

Table 2. Mapping of KBJI Occupational Group and ISCO Skill Levels to Required Education Levels

KBJI	ISCO	Required Level of Education
1	4	S1/S2/S3
2	4	S1/S2/S3
3	3	D1/D2/D3
4	2	SMA/SMK
5	2	SMA/SMK
6	2	SMA/SMK
7	2	SMA/SMK
8	2	SMA/SMK
9	1	SD

In addition, vertical mismatch status was determined by comparing an individual's highest educational attainment with the required educational level for the job. Individuals were categorized as matched if their educational level matched the job qualifications, overeducated if their educational level was higher, and undereducated if their educational level was lower than the required educational level. The classification matrix used in this study is presented in Table 3.

Table 3. Classification of Vertical Mismatch by Education Level and ISCO

Level of Education	ISCO 1	ISCO 2	ISCO 3	ISCO 4
D1-D3	Overeducation	Overeducation	Match	Undereducation
S1	Overeducation	Overeducation	Overeducation	Match
S2	Overeducation	Overeducation	Overeducation	Match
S3	Overeducation	Overeducation	Overeducation	Match

Although this classification approach is widely used in studies of vertical mismatch, it has several limitations. This study has limitations in measuring vertical mismatches because it uses the 2014 KBJI mapping to ISCO-08 and ISCED to determine the educational level required for a job. This approach uses standard educational levels based on occupational group classifications, so variations in educational requirements among jobs within the same occupational group may not be fully captured [17].

Multinomial Logistic Regression

Moreover, multinomial logistic regression was used in this study because the dependent variable, vertical mismatch status, consisted of three nominal and mutually exclusive categories: matched, overeducated, and undereducated. Because the dependent variable in this study consists of nominal categories, multinomial logistic regression is an appropriate analytical method for modelling the relationship between the response variable and a set of categorial or continuous predictor variables [18]. This modelling approach has also been succesfully in statistical studies in Indonesia, demonstrating its suitability for analysing nominal response variables with multiple outcome categories [19]. In this study, the matched category was used as the reference category. Therefore, the model produces two logit functions comparing the overeducation and undereducation categories against the match category, namely [20]:

$$g_1(x) = ln\left(\frac{P(Y=1)}{P(Y=0)}\right) = \beta_{01} + \beta_{11}X_1 + \cdots + \beta_{n1}X_n \qquad (1)$$

$$g_2(x) = ln\left(\frac{P(Y=2)}{P(Y=0)}\right) = \beta_{02} + \beta_{12}X_1 + \cdots + \beta_{n2}X_n \qquad (2)$$

where Y = 0 represents the matched, Y = 1 represents the overeducation, and Y = 2 represents the undereducation category.

Before interpreting the parameters, multicollinearity among the independent variables was checked using the Variance Inflation Factor (VIF). Next, the overall model significance was evaluated using the Likelihood Ratio Test (LRT), and the significance of each parameter was tested using the Wald Test. The dependent variable in this study represents educational mismatch status (*match*, *overeducation*, and *undereducation*), which is defined by comparing an individual's educational attainment with the educational requirements of the occupation. Accordingly, the three outcome categories are conceptually mutually exclusive [21]. Therefore, the Independence of Irrelevant Alternatives (IIA) assumption was considered based on the conceptual characteristics of the outcome variable. Model fit was subsequently evaluated using McFadden's, Cox and Snell's, and Nagelkerke's pseudo R^2 statistics.

Average Marginal Effects (AME)

Furtheremore, the coefficients obtained from multinomial logistic regression are expressed in log-odds units, making direct interpretation in terms of probability less intuitive. Therefore, this study employed Average Marginal Effects (AME) to estimate the average change in predicted probabilities associated with changes in the independent variabels [22]. In the multinomial logit model, AME was obtained by calculating the marginal effects for each observation and then averaging them. Mathematically, AME can be expressed as follows:

$$AME = \frac{1}{n}\sum_{i=1}^{n} p_{ij}(\beta_{kj} - \overline{\beta_k}) \qquad (3)$$

where

n : number of observations

$p_{ij}(\beta_{kj} - \overline{\beta_k})$: marginal effect for each observation

The AME value indicated the changes in the average probability of a category due to changes in the independent variable, assuming other variables were constant. In this study, AME was used to interpret the effects of gender and other variables on the probability of an individual being overeducated, undereducated, or matched [23].

RESULT AND DISCUSSION

Distribution of Vertical Mismatch

The distribution of vertical mismatch was analyzed to provide an initial overview of the match between education level and occupation for college graduates in Central Java Province in 2025. The overall distribution of vertical mismatch status is presented in Figure 2.

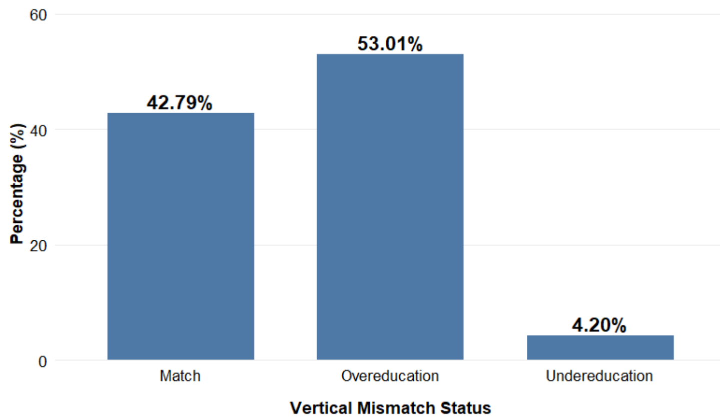

Figure 2. Distribution of Vertical Mismatch in Central Java

Based on Figure 2, the overeducation category was the most dominant condition among college graduates in Central Java Province, accounting for 53.01% of the total sample. This proportion was higher than the matched category, which accounted for 42.79%, while the undereducation category had the lowest proportion at only 4.20%. These findings indicated that more than half of college graduates in Central Java Province worked in jobs that required a lower level of education than their own, making overeducation the most dominant form of vertical mismatch. This is consistent with various studies in Indonesia that report that the increase in the number of college graduates has not been accompanied by a comparable increase in job opportunities for jobs requiring higher education qualifications. This pattern suggests that the high proportion of overeducation in this study not only reflects the results of individual job searches but also indicates a structural imbalance between the increasing number of college graduates and the structure of jobs available in the Central Java Province labor market.

In addition, the distribution of vertical mismatch was analyzed by gender to examine differences in patterns between male and female graduates. The number 1 in blue represents the male group, while the number 2 in orange represents the female group. The results are presented in Figure 3.

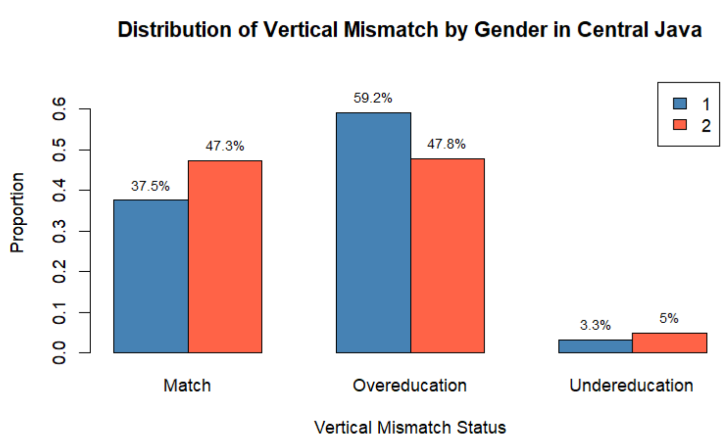

Figure 3. Distribution of Vertical Mismatch by Gender in Central Java

Based on Figure 3, the proportion of men experiencing overeducation reached 59.2%, higher than the proportion of women experiencing 47.8%. Conversely, the proportion of women experiencing a match was higher (47.3%) than men (37.5%). The proportion of undereducation was relatively small in both gender groups, at 3.3% for men and 5% for women. This descriptive pattern differs from the national findings of Khoiruddin et al. [10], who reported that women are more likely to experience overeducation in Indonesia. This difference is likely influenced by differences in sample characteristics, study area coverage, survey period, occupational composition, and the mismatch classification approach used. Nevertheless, various studies in Indonesia consistently show that gender is a significant factor related to educational mismatch [24][25]. Overall, these findings suggest that the gender differences observed in Central Java Province require further multivariate analysis and cannot be directly interpreted as evidence of a direct causal relationship.

Multicollinearity Test

The result of the multicollinearity test using the Variance Inflation Factor (VIF) are presented in Table 4.

Table 4. VIF

Variable	VIF
Gender	1.0569
Age	1.4184
Marital Status	1.3703
Highest Level of Education	1.0217
Residential Classification	1.0034
Wages	1.0494
Working Hours	1.0391

Based on Table 4, all independent variables have VIF values well below the commonly accepted threshold, indicating that multicollinearity is unlikely to affect the stability of the estimated regression coefficients. Following the recommendations in previous methodological studies, all independent variables were retained in the multinomial logistic regression model [26].

Multinomial Logistic Regression Model Parameter Estimation

The results of multinomial logistic regression parameter estimation are presented in Table 5. The model used the match category as a reference to produce two logit functions: the logit function for the overeducation category against match and the logit function for the undereducation category against match.

Table 5. Estimation of the Multinomial Logistic Regression Model

Variable	Category	Overeducation	Undereducation
Gender	Female	- 0.7417	-0.0799
Age		- 0.0074	-0.0082
Marital Status	Married	- 0.5199	- 0.3988
	Divorced	- 0.1434	0.4047
	Widowed	- 0.3383	- 0.2946
Highest Level of Education	D4	- 0.6698	- 19.8949

	S1	- 1.4593	- 18.7967
	S2	- 2.2593	- 18.6893
	Applied Master's Degree	- 2.0979	-19.7526
	S3	- 3.5636	- 19.0648
Residential Classification	Rural	- 0.3964	0.3572
Wages		- 1.86×10^{-7}	- 2.82×10^{-8}
Working Hours		0.0123	0.0065

Based on the estimation results, the logit function for the overeducation category against the match category was obtained as follows:

$$g_1(x) = 2.7156 - 0.7417X_1 - 0.0074X_2 - 0.5199X_{3a} - 0.1434X_{3b} - 0.3383X_{3c} - 0.6698X_{4a} - 1.4593X_{4b} - 2.2593X_{4c} - 2.0979X_{4d} - 3.5636X_{4e} - 0.3964X_5 - 1.86 \times 10^{-7}X_6 + 0.0123X_7$$

Meanwhile, the logit function for the undereducation category against the match category was obtained as follows:

$$g_1(x) = 0.9528 - 0.0799X_1 - 0.0082X_2 - 0.3988X_{3a} + 0.4047X_{3b} - 0.2946X_{3c} - 19.8949X_{4a} - 18.7967X_{4b} - 18.6893X_{4c} - 19.7526X_{4d} - 19.0648X_{4e} - 0.3572X_5 - 2.82 \times 10^{-8}X_6 + 0.0065X_7$$

with X_1: gender, X_2: age, X_3: marital status, X_4: highest level of education, X_5: residential classification, X_6: wages, dan X_7: working hours.

The very large negative coefficients for the undereducation category at the D4, S1, S2, Applied Master's Degree and S3 levels were related to the absence of undereducated respondents in these educational groups in the study sample. Based on the vertical mismatch classification matrix, D4, S1, S2, Applied Master's Degree and S3 graduates cannot be classified as undereducated, resulting in structural zero cells in these education groups. This condition creates complete or quasi-complete separation, which causes the coefficient estimates to be very large in absolute terms [27]. Therefore, the coefficients in the undereducation category are more appropriately understood in conjunction with the Average Marginal Effects (AME) results, as AME provides interpretation in terms of more easily understood probability changes. Furthermore, these findings suggest that the observed patterns are influenced not only by the data distribution but also by the operational definitions used to classify vertical mismatch [28][29].

Simultaneous Significance Test

The results of the simulaneous significance test using the Likelihood Ratio Test (LRT) are presented in Table 6.

Table 6. Likelihood Ratio Test

Model	Resid. df	Resid. Dev	Test	Df	Lr Statistic	P-value
1	8198	6874.377				
2	8172	5522.762	1 vs 2	26	1351.615	< 0.001

Where :

Model 1 : model without independent variables

Model 2 : model with all independent variables

Based on Table 6, the residual deviance value decreased from 6874.377 in the model without independent variables to 5522.762 in the model containing all independent variables. The decrease in the residual deviance value showed that the addition of independent variables significantly improved the fit compared to a model that only contains an intercept. The Likelihood Ratio Test results produced an LR statistical value of 1351.615 with a p-value < 0.001. Because the p-value was smaller than 0.05, H_0 was rejected [30]. These results indicated that the multinomial logistic regression model was significant simultaneously, so that the independent variables used together contributed significantly in explaining variations in the vertical mismatch status of college graduates in Central Java Province.

Partial Significance Test

Partial significance tests were conducted using the Wald test to identify independent variables significantly associated with vertical mismatch status [31]. The results of the partial significance test for parameters are presented in Table 7.

Table 7. Partial Test

Mismatch	Variabel	Kategori	P-Value	Keputusan
Overeducation	Gender	Female	<0.001	Signifikan
	Age		<0.001	Signifikan
	Marital Status	Married	<0.001	Signifikan
		Divorced	<0.001	Signifikan
		Widowed	<0.001	Signifikan
	Highest Level of Education	D4	<0.001	Signifikan
		S1	<0.001	Signifikan
		S2	<0.001	Signifikan
		Applied Master's Degree	<0.001	Signifikan
		S3	<0.001	Signifikan
	Residential Classification	Rural	<0.001	Signifikan
	Wages		<0.001	Signifikan
	Working Hours	Working Hours	<0.001	Signifikan
Undereducation	Gender	Female	<0.001	Signifikan
	Age		<0.001	Signifikan
	Marital Status	Married	<0.001	Signifikan
		Divorced	<0.001	Signifikan
		Widowed	<0.001	Signifikan
	Highest Level of Education	D4	<0.001	Signifikan
		S1	<0.001	Signifikan
		S2	<0.001	Signifikan
		Applied Master's Degree	<0.001	Signifikan
		S3	<0.001	Signifikan
	Residential Classification	Rural	<0.001	Signifikan
	Wages	Wages	0.1037	Signifikan
	Working Hours	Working Hours	<0.001	Signifikan

Based on Table 7, the partial significance test results showed that all independent variables in the overeducation category had p-values less than 0.05. Therefore, gender, age, marital status (married, divorced, and widowed), highest level of education (D4, S1, S2, Applied Master's Degree and S3), residential classification, wages, and working hours significantly of college graduates being classified as overeducated relative to matched employment. These findings indicate that both individual characteritics and labour market condiitions play an important role in explaining the occurrence of overeducation among college graduates in Central Java Province.

In the undereducation category, gender, age, marital status (married, divorced, and widowed), highest level of education (D4, S1, S2, Applied Master's Degree, and S3), residential classification, and working hours also showed significant effects with p-values less than 0.05. However, the wage variable was not statistically significant (p = 0.1037), indicating that after controlling for the other variables in the model, wages were not associated with the likelihood of college graduates being classified as undereducated relative to match employment.

Furthermore, the significance of the gender variable in both the overeducation and undereducation categories indicated that gender differences played an important role in explaining variations in vertical mismatch status among college graduates. This finding is consistent with previous studies in Indonesia, which also identified gender as a significant determinant of educational mismatch [11][32]. One factor that explains these findings is the disparity in employment opportunities and job allocation experienced by men and women in the labor market, resulting in different chances of securing a job that matches their educational qualifications [33][34]. Accordingly, differences in vertical mismatch by gender may be attributed not only to differences in educational attainment but also to labour market mechanisms that shape employment opportunities for college graduates.

Goodness of Fit

In addition to the Likelihood Ratio Test, model fit was evaluated using three pseudo R^2 measures, namely McFadden's, Cox and Snell's, and Nagelkerke's pseudo R^2. The results are presented in Table 8.

Table 8. Goodness of Fit

Pseudo R^2	**Value**
McFadden's	0.1966
Cox and Snell's	0.2808
Nagelkerke's	0.3454

The goodness of fit value indicated that the multinomial logistic regression model had an adequate fit to the data. The McFadden's Pseudo R^2 value of 0.1966, Cox and Snell's Pseudo R^2 of 0.2808, and Nagelkerke's Pseudo R^2 of 0.3454 indicate that the model with explanatory variables provided a better fit than the model that only contained the intercept. The Pseudo R^2 value was used as a Likelihood based measure of goodness of fit and cannot be interpreted as the proportion of variation explained by the model as in linear regression. In general, a larger pseudo R^2 value indicates better model fit, although there are no standard criteria regarding the value that indicates optimal model fit [35]. Thus, the goodness of fit results show that the model built is adequate for analysing the relationship between gender and vertical mismatch among college graduates in Central Java Province.

Average Marginal Effects (AME)

Moreover, the Average Marginal Effects (AME) were used to measure the average differences inthe probabilities of belonging to each vertical mismatch category between male and female graduates, while holding other variables constant. Unlike logistic regression coefficients, which were interpreted in log odds units, AME provided a direct interpretation in the form of probability changes, making it easier to understand [36]. The results of the AME calculation are presented in Table 9.

Table 9. Average Marginal Effect

Kategori	AME	P-value	95% CI
Match	0.1443	<0.001	[0.1434 ; 0.1452]
Overeducation	-0.1611	<0.001	[-0.1620 ; -0.1603]
Undereducation	0.0168	<0.001	[0.0162 ; 0.0174]

Based on Table 8, female graduates have a 14.43 percentage point higher probability of being in the match category, a 16.11 percentage point lower probability of being overeducated, and a 1.68 percentage point higher probability of being undereducated compared to male graduates. All AME values were significant at the 5% level. This result indicated that gender significantly influenced an individual's probability of being in each vertical mismatch category. These results showed that after controlling for other variables used in the model, female graduates had a greater chance of obtaining a job that matched their educational qualifications and a lower chance of experiencing overeducation than male graduates.

Furthermore, the AME values indicated that the most dominant influence of gender occurred in the overeducation category. Women had a lower probability of being overeducated than men. From the perspective of Assignment Theory, differences in education-job matching occur because workers are allocated to jobs that do not necessarily maximize the correspondence between their educational qualifications and job requirements. As a result, education-job mismatch may arise even among highly educated workers [37]. Labor Market Segmentation Theory explains that men and women may face different job opportunities in various segments of the labor market. These differences in job opportunities can affect an individual's chances of obtaining a job that matches their educational level, resulting in different probabilities of being overeducated between men and women [2].

In addition to having a lower probability of overeducation, women also had a higher probability of being in the match category than men. This finding indicated that women had a greater chance of obtaining a job that matches their educational level. The results of this study were in line with Lu and Li [7], who showed that men and women had different probabilities of experiencing a mismatch between education and employment due to differences in opportunities in the labor market. Similarly, Tran et al. [6] showed that vertical mismatch in college graduates was influenced not only by educational characteristics, but also by demographic factors and labor market structure. Thus, the findings of this study confirmed that vertical mismatch was the result of the interaction between individual characteristics and labor market mechanisms, resulting in different opportunities for men and women to obtain jobs that match their educational level.

CONCLUSION

Based on the research results, the following conclusions were reached.

1. The most dominant vertical mismatch among college graduates in Central Java Province in 2025 was overeducation. Furthermore, there were differences in the pattern of vertical mismatch between men and women. The proportion of overeducation was higher among men, and women were more likely to be in a matched condition.
2. Gender significantly affected the likelihood of college graduates being overeducated and undereducated after controlling for age, marital status, highest level of education, residential classification, wages, and working hours.
3. Changing gender from male to female decreased the probability of being overeducated by 16.11%, increased the probability of being in a matched condition by 14.43%, and increased the probability of being undereducated by 1.68%.

Beyond its practical implications, this study also contributes to labor market research by providing recent empirical evidence on the relationship between gender and vertical mismatch among college graduates in Central Java Province. The findings indicate that gender is an important determinant of vertical mismatch after controlling for individual and job-related characteristics, thereby enriching the empirical evidence on education job mismatch in the Indonesian labor market.

The findings of this study indicate that universities need to strengthen the link between higher education and labor market needs by aligning curricula, enhancing career development services, and increasing collaboration with the business and industrial sectors. Furthermore, these findings can serve as a foundation for employment agencies and local governments in formulating policies that support improved job matching and reduce vertical mismatches while taking gender differences into account. Further research is recommended to examine horizontal mismatches, field of study mismatches, and labor market outcomes using longitudinal data, thereby providing a more comprehensive understanding of the dynamics of mismatches between education and employment.

REFERENCE

[1] G. S. Becker, *Human Capital: A Theoretical and Empirical Analysis*, 1st ed. Chicago, IL: University of Chicago Press, 1964.

[2] M. Reich, D. M. Gordon, and R. C. Edwards, “A Theory of Labor Market Segmentation,” in *Economics and Social Justice*, S. Bowles and T. E. Weisskopf, Eds., Edward Elgar Publishing, 1998, pp. 65–72. doi: 10.4337/9781035351428.00010.

[3] M. Hanri and N. K. Sholihah, “KAJIAN PERLINDUNGAN SOSIAL DAN TENAGA KERJA,” vol. 5, 2024.

[4] Kementerian Pendidikan Tinggi, Sains, dan Teknologi Republik Indonesia, *Statistik Pendidikan Tinggi 2025*. Jakarta: Sekretariat Jenderal, Kementerian Pendidikan Tinggi, Sains, dan Teknologi, 2026. Accessed: Jul. 17, 2026. [Online]. Available: https://data.kemdiktisaintek.go.id/publikasi/buku-statistik-pendidikan-tinggi-2025

[5] Badan Pusat Statistik, “Tingkat Pengangguran Terbuka Berdasarkan Tingkat Pendidikan, 2025.” Accessed: Jan. 30, 2026. [Online]. Available: https://www.bps.go.id/id/statistics-table/2/MTE3OSMy/tingkat-pengangguran-terbuka-berdasarkan-tingkat-pendidikan.html

[6] BPS Provinsi Jawa Tengah, “Keadaan Angkatan Kerja Provinsi Jawa Tengah Agustus 2024,” Badan Pusat Statistik Provinsi Jawa Tengah, 2025.

[7] D. Yansyah, E. R. K. Wati, M. Nurrizalia, M. I. Kurniawan, A. Batrisya, and R. Wulandari, “Pengaruh Pendidikan Bagi Perempuan untuk Mendapat Kesempatan Kerja Guna Meningkatkan Perekonomian Keluarga,” *J. Pendidik. Non Form.*, vol. 1, no. 3, p. 13, Apr. 2024, doi: 10.47134/jpn.v1i3.352.

[8] T. Quang Tran, H. T. T. Nguyen, N. B. T. Vu, and D. V. Le, “Factors Associated With Vertical and Horizontal Job-Education Mismatches: Insights From University Graduates in Vietnam,” *Eval. Rev.*, p. 0193841X251409287, Dec. 2025, doi: 10.1177/0193841X251409287.

[9] Y. Lu and X. Li, “Lost in Translation: Gender Gap in Education–Occupation Mismatch among Highly Educated Workers,” *Work Occup.*, p. 07308884251380099, Sep. 2025, doi: 10.1177/07308884251380099.

[10] M. A. Khoiruddin, A. M. Setyanti, A. Suman, F. Prasetyia, and S. Susilo, “Exploring Determinants of Education-Job Mismatch Among Educated Workers in Indonesia,” *J. Ekon. Pembang. Kaji. Masal. Ekon. Dan Pembang.*, pp. 263–281, Dec. 2024, doi: 10.23917/jep.v25i2.23994.

[11] P. Wicaksono, I. Theresia, and B. Al Aufa, “Education–occupation mismatch and its wage penalties: Evidence from Indonesia,” *Cogent Bus. Manag.*, vol. 10, no. 3, p. 2251206, Dec. 2023, doi: 10.1080/23311975.2023.2251206.

[12] Badan Pusat Statistik Provinsi Jambi, “Yuk, Kenalan dengan Sakernas - Berita dan Siaran Pers.” Accessed: Jun. 10, 2026. [Online]. Available: https://jambi.bps.go.id/id/news/2022/08/23/269/yuk--kenalan-dengan-sakernas.html

[13] Badan Pusat Statistik Indonesia, *Tabel Kesesuaian KBJI 2014-2002*. Jakarta: Badan Pusat Statistik, 2015. Accessed: Jun. 10, 2026. [Online]. Available: https://www.bps.go.id/id/publication/2015/03/13/2948ff3fdc52c2298acacbb3/tabel-kesesuaian-kbji-2014-2002.html

[14] E. Hasibuan and D. Handayani, “Pengaruh Qualification Mismatch terhadap Upah Tenaga Kerja Di Indonesia,” *J. Ekon. Dan Pembang.*, vol. 29, no. 1, pp. 1–16, Aug. 2021, doi: 10.14203/JEP.29.1.2021.1-16.

[15] T. Lumley and A. Scott, “Fitting Regression Models to Survey Data,” *Stat. Sci.*, vol. 32, no. 2, May 2017, doi: 10.1214/16-STS605.

[16] International Labour Office, *ISCO-08*, vol. 1. Geneva: International Labour Office, 2012.

[17] S. McGuinness, K. Pouliakas, and P. Redmond, “How Useful is the Concept of Skills Mismatch?,” *SSRN Electron. J.*, 2017, doi: 10.2139/ssrn.2979934.

[18] R. Crisdianto, M. F. A. Rakafalih, A. Saputri, L. Sopiansyah, A. Fairuzindah, and D. S. Rini, “Analysis of the Timeliness of Graduation of FMIPA College KIP Students at Bengkulu University Using Binary Logistic Regression,” *J Stat.*, vol. 16, no. 2, pp. 574–584, 2023.

[19] L. P. S. Pratiwi and I. M. P. P. Wijaya, “Binary Logistic Regression Analysis on the Spread of Dengue Fever in Bali Province,” *J Stat.*, vol. 17, no. 1, pp. 614–623, 2024.

[20] B. L. Fauzan, T. Agustin, and A. M. H. Mahmudah, "Prediksi Klasifikasi Kecelakaan Lalu Lintas di Kota Surakarta dengan Menggunakan Metode Regresi Logistik Multinomial," *Sustain. Civ. Build. Manag. Eng. J.*, vol. 1, no. 4, p. 9, Aug. 2024, doi: 10.47134/scbmej.v1i4.3159.

[21] M. Van Smeden, J. A. De Groot, S. Nikolakopoulos, L. C. Bertens, K. G. Moons, and J. B. Reitsma, "A generic nomogram for multinomial prediction models: theory and guidance for construction," *Diagn. Progn. Res.*, vol. 1, no. 1, p. 8, Dec. 2017, doi: 10.1186/s41512-017-0010-5.

[22] R. Williams, "Using the Margins Command to Estimate and Interpret Adjusted Predictions and Marginal Effects," *Stata J. Promot. Commun. Stat. Stata*, vol. 12, no. 2, pp. 308–331, Jun. 2012, doi: 10.1177/1536867X1201200209.

[23] J. N. Wulff, "Interpreting Results From the Multinomial Logit Model: Demonstrated by Foreign Market Entry," *Organ. Res. Methods*, vol. 18, no. 2, pp. 300–325, Apr. 2015, doi: 10.1177/1094428114560024.

[24] A. Setiyana and S. I. Oktora, "Analisis horizontal mismatch pada tenaga kerja lulusan Sekolah Menengah Kejuruan di Indonesia," *J. Kependud. Indones.*, vol. 19, no. 2, pp. 109–124, 2024.

[25] S. Maryanti, N. Bachtiar, S. Maryati, and D. Novianto, "Impact of Educational Mismatch on Graduates Employment Prospects in Indonesia: Education and Employment Context," *TEC Empres.*, vol. 20, no. 2, pp. 1135–1164, 2025.

[26] R. M. O'Brien, "A Caution Regarding Rules of Thumb for Variance Inflation Factors," *Qual. Quant.*, vol. 41, no. 5, pp. 673–690, Oct. 2007, doi: 10.1007/s11135-006-9018-6.

[27] R. K. Putri, M. Athoillah, and A. Haqiqiyah, "Analisis Faktor Yang Mempengaruhi Ketepatan Kelulusan Mahasiswa Dengan Algoritma Regresi Linear," Jurnal Lebesgue: Jurnal Ilmiah Pendidikan Matematika, Matematika dan Statistika, vol. 5, no. 2, pp. 671–680, 2024.

[28] S. McGuinness, K. Pouliakas, and P. Redmond, "How useful is the concept of skills mismatch?," IZA Discussion Papers, 2017.

[29] M. Salas-Velasco, "Mapping the (mis) match of university degrees in the graduate labor market," *J. Labour Mark. Res.*, vol. 55, no. 1, p. 14, 2021.

[30] F. Arofah and A. Sofro, "Penerapan Regresi Logistik Multinomial untuk Analisis Model Tingkat Depresi pada Lansia," *MATHunesa J. Ilm. Mat.*, vol. 10, no. 1, pp. 84–93, Apr. 2022, doi: 10.26740/mathunesa.v10n1.p84-93.

[31] P. Sembiring, "Analisis Regresi Logistik Untuk Menentukan Faktor-Faktor Yang Mempengaruhi Kesejahteraan Masyarakat Kabupaten/Kota Di Pulau Nias," *FARABI J. Mat. Dan Pendidik. Mat.*, vol. 6, no. 1, pp. 25–31, 2023.

[32] S. Maryanti, N. Bachtiar, S. Maryati, and D. Novianto, "Impact of Educational Mismatch on Graduates Employment Prospects in Indonesia: Education and Employment Context," *TEC Empres.*, vol. 20, no. 2, pp. 1135–1164, 2025.

[33] C. Desmond, K. Watt, S. Naicker, J. Behrman, and L. Richter, "Girls' schooling is important but insufficient to promote equality for boys and girls in childhood and across the life course," *Dev. Policy Rev.*, vol. 42, no. 1, p. e12738, Jan. 2024, doi: 10.1111/dpr.12738.

[34] L. U. Pradana, P. Lestari, and S. Khu, *Ketimpangan Gender dalam Dunia Kerja dan Rumah Tangga di Indonesia: Analisis Teori Feminis terhadap Isu-Isu dan Solusi*. ResearchGate, 2025.

[35] M. Van Smeden, J. A. De Groot, S. Nikolakopoulos, L. C. Bertens, K. G. Moons, and J. B. Reitsma, "A generic nomogram for multinomial prediction models: theory and guidance for construction," *Diagn. Progn. Res.*, vol. 1, no. 1, p. 8, Dec. 2017, doi: 10.1186/s41512-017-0010-5.

[36] M. C. Perraillon, "Interpreting Model Estimates: Marginal Effects," Jul. 2019, [Online]. Available: https://clas.ucdenver.edu/marcelo-perraillon/sites/default/files/attached-files/perraillon_marginal_effects_lecture_lisbon_0.pdf

[37] M. Sattinger, "Assignment models of the distribution of earnings," *J. Econ. Lit.*, vol. 31, no. 2, pp. 831–880, 1993.