STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **14**, August 2025, pp 541–555.
Published online in International Academic Press (www.IAPress.org)

Modelling Labor Market Access Among Vocational Education Graduates in Indonesia: A Heckman Selection Approach

Fachrizal [1], Mia Rahma Romadona [1], Suryadi Suryadi [1], Andi Budiansyah [1,*], Syahrizal Maulana [1], Maulana Akbar [1], Ratna Sri Harjanti [2], Arif Hidayat [2]

[1] *National Research and Innovation Agency, Indonesia*
[2] *Politeknik LPP Yogyakarta, Indonesia*

Abstract Understanding the labour market outcomes of vocational education graduates is essential for evaluating the effectiveness of mid-level tertiary education. This study uses the Heckman two-step selection model to estimate the returns on vocational education for graduates of Diploma (D1-D4) programs using nationally representative data from Indonesia's 2023 National Labor Force Survey (SAKERNAS). In the first stage, a probit model estimates the probability of labour market participation as a function of major industry sectoral, educational attainment, demographic characteristics, household factors, and urban residence. The second stage estimates the determinants of wages, conditional on employment, with the Inverse Mills Ratio (IMR) from the first stage included to correct for selection bias. Results indicate that higher vocational qualifications significantly increase both the likelihood of employment and monthly wages. However, women with children and urban residents are less likely to be employed, whereas female workers and vocational graduates residing in rural areas tend to have lower earnings. This study contributes methodologically by demonstrating how selection-corrected statistical modelling can improve the accuracy of returns to education estimates, particularly for vocational education. Substantively, the findings offer policy-relevant insights into the evolving role of vocational education in promoting economic inclusion in Indonesia's post-secondary education and labour market systems.

Keywords Vocational Education, Labour Market Access, Heckman Selection Model, Indonesia, SAKERNAS

DOI: 10.19139/soic-2310-5070-2510

1. Introduction

In a labour force survey like the Indonesian National Labor Force Survey (SAKERNAS), wage information is only available for those who are currently employed, leaving out individuals who may be actively seeking jobs or who have exited the labour market entirely [1]. Advanced econometric techniques, such as the Heckman selection model, provide a framework to correct this bias by simultaneously accounting for employment probability and wage determination [2]. As a result, wage estimates obtained from standard models may not accurately reflect the true returns to vocational education, as they fail to account for the differences between employed and those who are not.

The effectiveness of education in improving labour market outcomes is influenced by the broader economic and institutional context, including the structure of the labour market and the availability of supportive policies [3, 4]. Universities and vocational training institutions should strengthen their career services, counselling, and industry partnerships to align educational outcomes with the needs of labour markets [5]. Many graduates experience job mismatches or are overeducated for their positions, which can diminish the returns on their educational investments

[*]Correspondence to: Andi Budiansyah (Email: andi.budiansyah@brin.go.id). Research Centre for the Economics of Industry, Services, and Trade National Research and Innovation Agency (BRIN).

ISSN 2310-5070 (online) ISSN 2311-004X (print)
Copyright © 2025 International Academic Press

[6]. By analyzing how graduates transition into the workforce, policymakers and educators can better understand whether vocational education successfully enhances job accessibility and economic mobility. A comprehensive evaluation of labour market outcomes can provide valuable insights into whether vocational graduates secure stable, well-paying jobs or face barriers to entry into the formal economy.

Despite its significance, evaluating the impact of vocational education on labour market outcomes poses considerable methodological challenges. Traditional wage regression models, which assess earnings differentials across various education levels [7], often disregard the inherent sample selection bias in employment data, or the process of selecting the sample may not be independent of the dependent variable [2]. Tackling sample selection bias is crucial for generating more accurate and policy-relevant findings regarding the effects of vocational education on the labour market, such as income [8]. By including these adjustments, researchers can more reliably estimate the real impact of vocational education on earnings and employment prospects. Such detailed analysis is vital for identifying potential labour market barriers that vocational graduates encounter and developing policies that enhance their employment opportunities.

A better understanding of vocational graduates' labour market outcomes can inform educational reforms and workforce policies in developing economies [9]. Ensuring that vocational education aligns with labour market needs requires ongoing evaluation and adaptation of curricula, industry partnerships, and government support programs [6, 10]. Researchers and policymakers can gain more precise insights into how vocational education contributes to economic inclusion by tackling methodological challenges and refining analytical approaches. These findings can guide evidence-based policy interventions to strengthen vocational training systems and maximize their role in labour market integration.

Motivation and Main Contributions

This study enhances the estimation of vocational education's impact on employment and wages by addressing sample selection bias, which arises because wage data in cross-sectional labour force surveys are typically only observed for employed individuals. By applying the Heckman selection model, the study provides more accurate estimates of vocational education's returns, ensuring that non-employed individuals are also considered in the analysis. Unlike broader studies on returns to education, this research provides a detailed, single-year cross-sectional analysis of Diploma 1 (D1) to Diploma 4 (D4) graduates' employment outcomes in Indonesia's current labour market. The study highlights how vocational education translates into job opportunities and wage outcomes under prevailing economic conditions.

2. Data and Methods

The theory of human capital in education focuses on how investments in education enhance individuals' knowledge and skills, thereby increasing their productivity and earning capacity [11]. This study uses data from the SAKERNAS in 2023, a nationally representative cross-sectional survey conducted by the Indonesian Statistics (BPS). SAKERNAS collects detailed information on labour force participation, employment characteristics, wages, and educational attainment. The dependent and independent variables in this study are based on the most used from the previous studies (Table 1) and the availability from SAKERNAS.

Job Opportunity (working status in SAKERNAS) and Monthly Wages are the dependent variables, while Age, Education Level (D1-D3 and D4 Level), Employment Status (Formal or Informal), Marital Status, Number of Children, Job Experience, Gender, and Business Sectors as the independent variables. Considering Indonesia as an archipelagic country, this study uses the Urban-Rural dummy as the independent variable since city populations are most at risk of open unemployment, especially for those who are less educated [12] and women in urban areas receive lower incomes than men even with equal levels of education [13].

The sample is not randomly selected in this study, leading to biased and inconsistent estimates. This bias is common in various fields, where the outcome of interest is only observed for a non-random subset of the population [14, 15]. As wage data in SAKERNAS is only available for employed individuals, estimating the returns to

Table 1. Previous studies of job opportunity and wage
Source: Authors' Work

Dependent Variables	Independent Variables	Previous Studies
Job Opportunity	Age	[19], [20], [21], [22]
	Age at First Marriage	[21]
	Earnings	[23]
	Education	[19], [20], [21], [22], [24]
	Employment Status	[21], [23]
	Ethnicity	[23]
	Gender	[20]
	Health	[20]
	Household Income	[19], [21]
	Marital Status	[19], [20], [22], [23]
	Non-wage Income	[20]
	Number of Children	[19], [20]
	Race	[23]
	Registration with Public Employment Service	[22]
	Sex	[22]
Wages	Age	[25], [26]
	Self-Employment	[22]
	City Size	[25]
	Civil Status	[27]
	Demographics	[28]
	Education	[25], [29]
	Education Level	[28], [30], [31]
	Ethnicity	[25]
	Experience	[25], [26], [27], [29], [30], [32]
	Field of Study	[28]
	Firm Size	[25]
	Gender	[25], [26], [33]
	Health	[29]
	Household Demographics	[29]
	Immigration Status	[30]
	Informal Employment	[27]
	Job Tenure	[26]
	Location	[27]
	Marital Status	[30]
	Migrant Status	[25]
	Mobility	[30]
	Pre-enrollment earnings	[28]
	Region	[25]
	Sectors of Occupation	[6], [26], [27], [29], [30], [34]
	Vocational vs General Education	[32]
	Years of Schooling	[26], [27], [32]

vocational education without correcting for selection into employment would result in biased estimates. The study applies the Heckman two-step selection model [16].

The first stage models the probability of labour market participation using probit regression to identify factors that affect a person's chances of being included in the observed sample (only those who work have wage data). In this first stage, an Inverse Mills Ratio (IMR) value is obtained to correct the potential bias that occurs [17, 18]. The second stage estimates the determinants of wages, conditional on employment, using Ordinary Least Squares (OLS) with the IMR from the first stage included to correct for selection bias. If the IMR is significant, it suggests that there is indeed a selection bias that needs to be corrected. Conversely, if the IMR is not significant, it may indicate that selection bias is not a concern, and alternative models might be more appropriate.

In this study, the sample is restricted to individuals who completed vocational education programs Diploma 1 to Diploma 4. Respondents with lower educational attainment (high school or below) or university degrees (S1 or higher) are excluded because the main analysis of this study is limited to vocational education only. The final sample consists of approximately 17,844 respondents aged 18 and above. Missing values in the start working month variable were imputed as January (month = 1), while the start working year variable with the value 9999 was excluded from the analysis. Outliers such as unusual years of work experience were removed. This study assumes that individuals could have started working after age 15. The Statistical Program software (STATA 16) was used to process the data.

2.1. Model Specification

The research steps undertaken using the Heckman two-step selection model are illustrated in Figure 1.

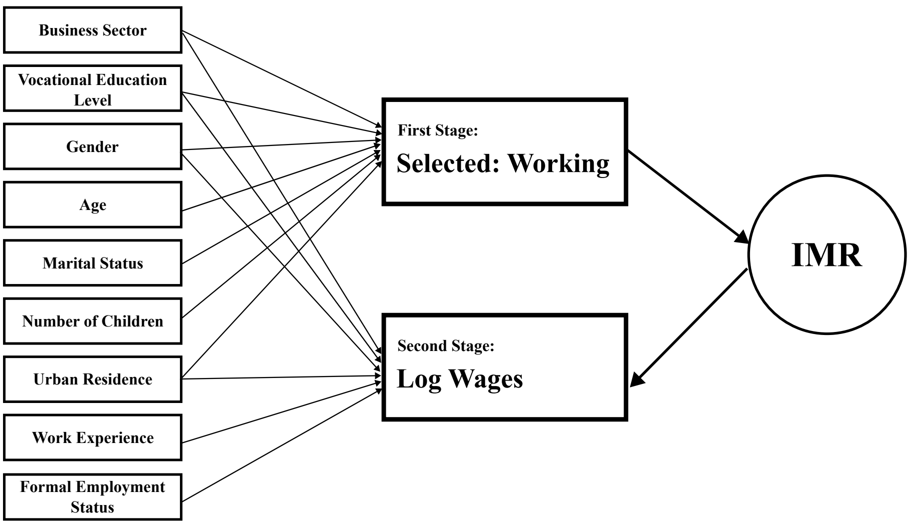

Figure 1. Research step using Heckman two-step selection model
Source: Authors' Work

First Stage: Labor Market Participation Model
The probability of labour market participation is specified as follows:

$$\Pr(Y_i) = \Phi(Z_i\gamma) \tag{1}$$

where Y_i=1 if an individual is employed and 0 otherwise. The vector Z_i includes business sector, vocational education level, gender, age, marital status, number of children, and urban residence.

Second Stage: Wage Determination Model
For individuals who are employed, the log of monthly wages is modelled as follows:

$$\ln(W_i) = \beta X_i + \lambda \cdot \text{IMR}_i + \varepsilon_i \tag{2}$$

where W_i defined the natural log of monthly wages. The variable X_i includes business sector, vocational education level, work experience, formal employment status, gender, and urban residence.

2.2. Marginal Effects Estimation

Marginal effects are computed from the first-stage probit model to estimate the incremental impact of explanatory variables on the probability of labour market participation. This provides interpretable estimates of how higher levels of vocational education influence employment probabilities.

2.3. Diagnostic and Robustness Checks

To ensure the robustness and validity of the model, several diagnostic procedures were performed. First, alternative models were estimated to validate the findings and confirm that the conclusions remain consistent across specifications. Multicollinearity among the explanatory variables in the wage equation was assessed using Variance Inflation Factors (VIF). Residual diagnostics were also conducted by examining Normal Quantile-Quantile (Q-Q) plots to evaluate the normality of the residuals. If the plotted points align closely with the diagonal line, the residuals are considered normally distributed; however, notable deviations may indicate issues such as heteroskedasticity, omitted variable bias, or non-linearity. Lastly, the presence of selection bias was examined through the statistical significance of the Inverse Mills Ratio in the wage equation, where a significant coefficient is interpreted as evidence of sample selection bias.

2.4. Descriptive Data

The demographic profile of vocational graduate respondents in the 2023 SAKERNAS shows variation in wage levels (Table 2), with an average log-transformed wage of 14.77 and a standard deviation of 0.88. The education level variable describes that only 10.18% of individuals hold a D4-level vocational degree or equivalent. The age distribution shows an average of 39.78 years, with a minimum of 18 years and a maximum of 89 years, reflecting a wide range of working-age individuals. Gender distribution reveals that approximately 61.61% of respondents are female. Residence distribution shows that most respondents live in the urban area (69.58%) and work as formal workers (74.18%).

Table 2. Summary statistics of Log Wages, Education Attainment, Age, Gender, and Urban Residence of the Vocational Graduate Respondents in the 2023 SAKERNAS
Source: Authors' Work

Variable	Obs	Mean	Std. Dev.	Min	Max
log_wage	12,271	14.77294	0.8753639	11.15625	18.9803
diploma4	17,844	0.101883	0.3025028	0	1
age	17,844	39.78884	12.91224	18	89
gender	17,844	0.6161175	0.4863435	0	1
urban	17,844	0.6958081	0.4600772	0	1
formal	13,594	0.7417979	0.4376618	0	1

As for gender working and not working (Figure 2), 81.26% male vocational graduates on this survey are working. Meanwhile, only 66.09% female vocational graduates are working. The business sectors based on Indonesian Standard Industrial Classification (KBLI) 2020 were simplified into three: agriculture (1), manufacturing (2-6), and services (7-17). Female vocational graduates dominate in all business sectors and mostly work in services (Figure 3).

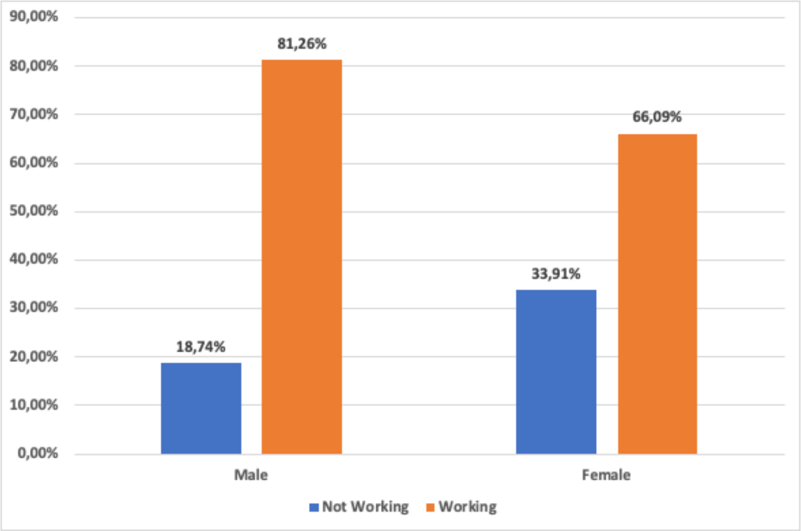

Figure 2. Gender working status of vocational graduates
Source: Authors' Work

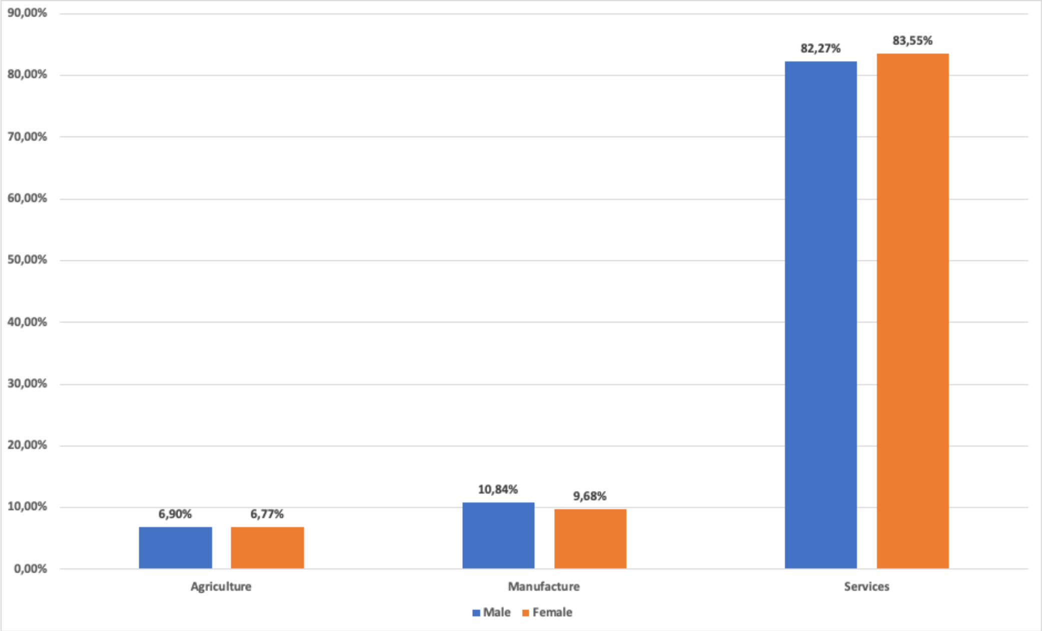

Figure 3. Gender distribution of vocational graduates based on business sector
Source: Authors' Work

The shape of the wage distribution (Figure 4) appears approximately normal, with a slight right skew, indicating that while most individuals have wages concentrated around the central values, some earn significantly higher wages. The distribution peak suggests that most workers fall within the middle-wage range, around 15, which translates into an approximate real wage level depending on the base used for the logarithmic transformation. The density curve overlayed on the histogram confirms the bell-shaped pattern, though some deviations and minor peaks suggest slight variations within wage groups.

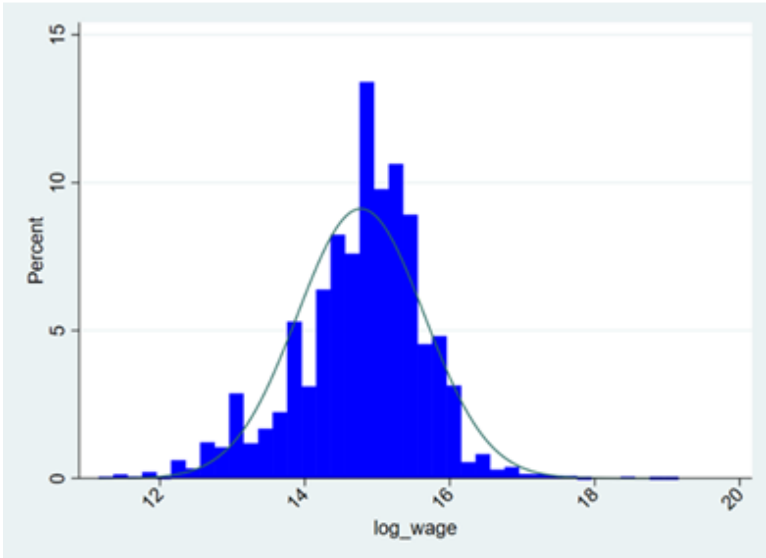

Figure 4. The histogram distribution of log-transformed wages (log wage)
Source: Authors' Work

The low density is on the distribution's lower and upper ends, implying that only a few individuals earn extremely low or extremely high wages. The presence of outliers in higher wage brackets might suggest a disparity in earnings among vocational education graduates. Additionally, the steep decline on the right tail indicates that fewer individuals earn exceptionally high wages compared to the average. Overall, the histogram suggests a relatively symmetric but slightly right-skewed wage distribution, which is common in labour market studies, especially when analyzing log-transformed wages. Figure 5 shows that the income disparity in Indonesia's provinces is quite significant, even among provinces that are still on the same island.

Figure 5. The income disparity in Indonesia's provinces
Source: Authors' Work

3. Results and Discussion

3.1. Selection Equation: Employment Probability (First-Stage Probit Model)

As shown in Table 3, the first-stage Probit model estimates the likelihood of employment (working), and the results indicate several key determinants. The employment opportunities of vocational graduates in the manufacturing sector are 26.38% higher than those of the agricultural sector (β = 0.2638). In contrast, the employment opportunities in the service sector are no different from those in agriculture (insignificant). Higher vocational education significantly increases employment probability (β = 0.1225). Age variable impacts employment probability (β = -0.0231), suggesting that older individuals may face challenges in the job market due to employer preferences for younger workers.

Women are significantly less likely to be employed than men (β = -0.6154), highlighting persistent gender-based employment barriers. Individuals with more children have lower employment probabilities (β = -0.1088), which may be linked to family responsibilities that limit job-seeking efforts. Marriage is positively associated with employment (β = 0.2218), likely due to economic necessity in dual-income households. Urban residents are less likely to be employed than rural workers (β = -0.1339), indicating that higher labour competition in cities may offset urban job advantages. These findings suggest that while vocational education is crucial in improving employability, gender disparities and urban labour competition remain major challenges.

Table 3. Heckman Selection Model Result
Source: Authors' Work

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
log_wage						
work_sec3						
Manufacturing	.0874333	.0427077	2.05	0.041	.0037278	.1711388
Services	.0974989	.0341027	2.86	0.004	.0306588	.164339
diploma4	.0668847	.0274145	2.44	0.015	.0131532	.1206162
work_exp	.0228659	.0011275	20.28	0.000	.020656	.0250758
formal	.3941391	.0215745	18.27	0.000	.351854	.4364243
urban	.3874272	.0193793	19.99	0.000	.3494445	.4254099
gender	-.3060957	.0292266	-10.47	0.000	-.3633787	-.2488126
_cons	14.20995	.0478815	296.77	0.000	14.1161	14.30379
working						
work_sec3						
Manufacturing	.2637896	.060105	4.39	0.000	.1459858	.3815933
Services	.0299144	.0469439	0.64	0.524	-.0620939	.1219227
diploma4	.1225429	.0401086	3.06	0.002	.0439314	.2011544
age	-.0230955	.0010055	-22.97	0.000	-.0250663	-.0211248
kids	-.1087955	.0222379	-4.89	0.000	-.1523809	-.06521
married	.221759	.0301727	7.35	0.000	.1626215	.2808964
urban	-.1338954	.0264033	-5.07	0.000	-.1856448	-.0821459
gender	-.6154447	.0258008	-23.85	0.000	-.6660134	-.564876
_cons	1.777866	.0675465	26.32	0.000	1.645477	1.910254
/mills						
lambda	-.3447133	.089883	-3.84	0.000	-.5208807	-.1685459
rho	-0.41535					
sigma	.82993937					

3.2. Marginal Effects Estimation

Marginal effects from the first-stage probit model estimate the incremental impact of explanatory variables on the probability of labour market participation (Figure 6). All explanatory variables are significant at the 95% confidence level except for the services sector's marginal effect on the agricultural sector. The estimation of the marginal effect of the manufacturing sector is 8.24% higher than that of the agricultural sector. This means that

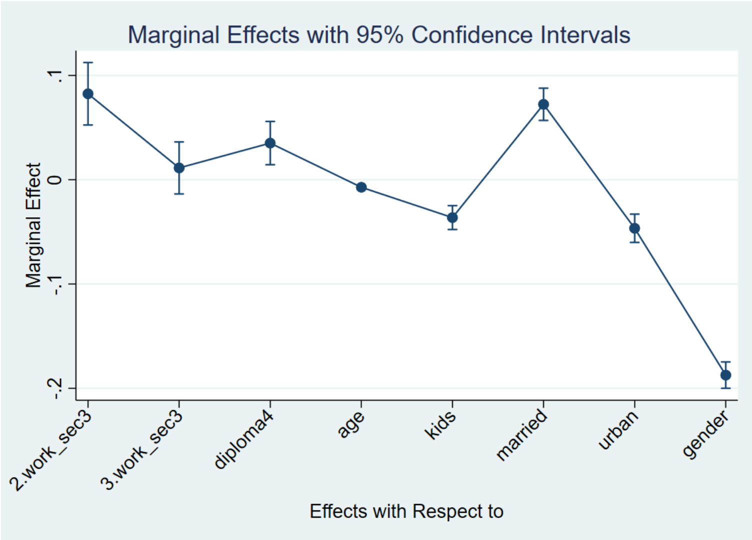

Figure 6. The average marginal effect of explanatory variables on the probability of labour market participation
Source: Authors' Work

vocational graduates have more chances to be employed in the manufacturing sector than in the agriculture sector. Diploma explanatory variables (education level D1-D3 to D4) marginal effect estimation is +3.51%, indicating that holding a higher vocational diploma level (D4) increases employment probability by 3.51 percentage points. Higher vocational education (from D1-D3 to D4) significantly enhances job access. The age explanatory variables' marginal effect estimation is -0.72%, which means that each additional year of age decreases employment probability by 0.72 percentage points. Older individuals face increasing difficulty in securing employment.

The Number of Children (kids) explanatory variables marginal effect estimation is -3.36%. Having an additional child reduces employment probability by 3.36 percentage points, suggesting that family responsibilities may limit labour force participation. Marital Status (married) explanatory variables marginal effect estimation is +7.24%, which means that married individuals are 7.24 percentage points more likely to be employed. This could be due to economic necessity and household financial obligations.

Urban Residence (urban) explanatory variables marginal effect estimation is -4.65%. This means that urban residents are 4.65 percentage points less likely to be employed than rural residents. This suggests that higher job competition in cities might offset urban job market advantages. Gender (Female = 1, Male = 0) explanatory variables marginal effect estimation is -18.73%, meaning that women are 18.73 percentage points less likely to be employed than men. This indicates strong gender disparities in vocational labour market access.

3.3. Wage Determination: Heckman-Corrected OLS (Second-Stage Regression)

The second-stage wage equation estimates the impact of vocational education and other factors on earnings (log wage). The results confirm several important labour market dynamics. Vocational graduates working in the manufacturing sector have an income 8.74% higher than those who work in the agricultural sector ($\beta = 0.0874$). While in the services sector have an income 9.74% higher than those who work in the agricultural sector ($\beta = 0.0974$). Vocational education is positively associated with wages ($\beta = 0.0669$), meaning that improving vocational education from D1-D3 to D4 increases earnings by approximately 6.69%. This underscores the importance of vocational education in wage growth. Work experience significantly improves wages ($\beta = 0.0229$), with each additional year of experience raising wages by 2.29%. This aligns with previous studies, where experience enhances productivity [35, 36, 37] and earnings potential [38, 39].

Formal employment leads to substantially higher wages ($\beta = 0.3798$), as workers in the formal sector earn 37.98% more than those in informal jobs. This reflects the greater job security and benefits of formal employment.

Urban workers earn significantly more than rural workers ($\beta = 0.3941$), supporting the idea that urban labor markets provide higher-paying opportunities despite the higher competition observed in the selection equation. A large gender wage gap exists ($\beta = -0.3061$), indicating that women earn 30.61% less than men, even after accounting for other factors. This may reflect structural barriers such as occupational segregation, discriminatory pay practices, or differences in job types between men and women. Advocate for employer partnerships to align curricula with regional labor needs (e.g., agriculture in rural areas, tech in urban areas). These findings reinforce the positive economic impact of vocational education but also expose structural inequalities, particularly regarding gender and job sector (formal vs. informal employment).

3.4. Diagnostic and Robustness Checks

Validate with Alternative Models

Empirical studies have shown that alternative methods can provide more accurate and reliable estimates in various contexts [40, 41]. Propensity Score Matching (PSM) can be an alternative to the Heckman two-step selection model as it is a statistical technique used to reduce bias in estimating treatment effects from observational data by matching treated and untreated subjects with similar propensity scores, as the probability of receiving the treatment given observed covariates [42, 43]. The PSM result (Table 4) is quite similar to the Heckman First-Stage Probit Model, suggesting that selection bias is a major concern in this dataset.

Table 4. Propensity Score Matching (PSM) Result
Source: Authors' Work

```
Probit regression                               Number of obs     =     13,594
                                                LR chi2(8)        =    1072.90
                                                Prob > chi2       =     0.0000
Log likelihood = -7598.2582                     Pseudo R2         =     0.0659
```

working	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
work_sec3						
Manufacturing	.2721914	.059455	4.58	0.000	.1556618	.388721
Services	.0350523	.0464783	0.75	0.451	-.0560434	.1261481
diploma4	.1109898	.0398131	2.79	0.005	.0329576	.189022
age	-.0227312	.0009934	-22.88	0.000	-.0246783	-.0207841
kids	-.1145428	.0219974	-5.21	0.000	-.1576568	-.0714288
married	.2284278	.0298814	7.64	0.000	.1698614	.2869941
urban	-.1468494	.0260659	-5.63	0.000	-.1979377	-.0957611
gender	-.5913811	.0255659	-23.13	0.000	-.6414893	-.5412729
_cons	1.777921	.0667143	26.65	0.000	1.647163	1.908679

Multicollinearity Check

Variance Inflation Factors (VIF) are calculated for explanatory variables in the wage equation. Table 5 shows that the value of the explanatory variables' VIF is less than 10, indicating that there is no serious multicollinearity among the independent variables in the model.

Residual Diagnostics

Normal Quantile-Quantile (Q-Q) plots are used to assess the normality of residuals from the model. Figure 7 shows that the residuals in the model slightly follow a normal distribution where the points are close to the 45-degree line.

Table 5. Multicollinearity test result
Source: Authors' Work

Variable	VIF	1/VIF
work_sec3		
2	2.37	0.421407
3	2.37	0.421362
diploma4	1.04	0.964629
work_exp	1.38	0.725442
formal	1.15	0.873226
urban	1.14	0.878530
gender	2.69	0.371397
imr	3.12	0.320640
Mean VIF	1.91	

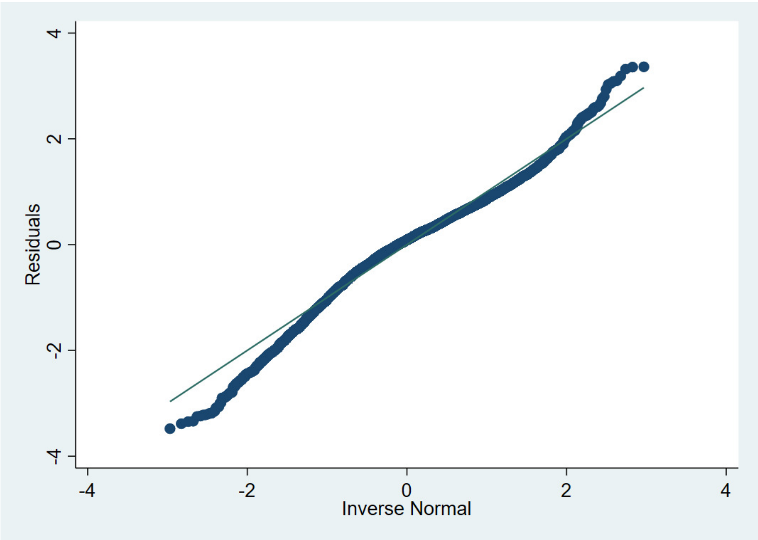

Figure 7. The Residual Diagnostics result
Source: Authors' Work

While Q-Q plots are a visual method of checking normality, statistical tests can also be used to test the normality of residuals (Table 6). The residuals from the regression model are approximately normally distributed (skewness-kurtosis test, Adj chi2(2) > 0.05), supporting the validity of the normality assumption. This means the model has no heteroskedasticity (non-constant variance), omitted variable bias, or non-linearity.

Table 6. Skewness-kurtosis test result
Source: Authors' Work

Variable	Obs	Pr(Skewness)	Pr(Kurtosis)	Adj chi2(2)	Prob>chi2
resid	9,299	0.0000	0.0000	437.00	0.0000

Selection Bias and the Inverse Mills Ratio (IMR)

The IMR (λ = -0.3447, p $>$ 0.05) is statistically significant, suggesting that selection bias is a major concern in this dataset. This implies that wage estimation using OLS without selection correction can lead to a biased estimation. Using Heckman's model provides important insights into employment probability determinants and reinforces the robustness of wage estimates.

Some findings of this study confirm the previous study. Vocational education significantly increases employment probability, which proves that higher education enhances job accessibility, and an additional level of vocational training also increases earnings [44]. The Indonesian government is also implementing the Independent Learning Independent Campus (MBKM) program in student learning. According to the results, 97% of student respondents stated that the program is valuable in developing their competencies and skills as preparation for entering the workforce, which could ultimately help reduce unemployment rates among vocational graduates [45].

The age variable in this study hurts employment probability, suggesting that older individuals may face challenges in the job market due to employer preferences for younger workers. However, work experience significantly improves wages, whereas experience enhances productivity and earnings potential. This finding is in line with [46], [47], and [48]. Older individuals are encouraged to engage in economic activities to fulfil their needs, especially since not all receive a pension or social security. Labour wages, non-labour income, and household status influence their working hours. The government recommends focusing on appropriate industrial sectors and family care programs to support elderly workers [49].

As for gender, in this study, women are significantly less likely to be employed than men, which is also in line with [50] and [51]. The individuals with more children in this study also have lower employment probabilities, which may be linked to family responsibilities that limit job-seeking efforts. Women with more children are less likely to be employed due to increased childcare responsibilities [23, 52, 53] and tend to work as informal workers [23]. In this study, formal workers earned 38% more than informal workers. This gender income gap is consistent with findings from previous studies that used the Blinder-Oaxaca decomposition method. For instance, gender discrimination and social expectations were found to contribute to lower wages for women in Cali, Colombia [54], while in the Lebanese banking sector, a significant portion of the gender wage gap was attributed to discrimination [55].

Structural reforms and globalization in Indonesia often struggle against systemic biases amplified during economic shocks [56]. The pre-employment card (Kartu Pra-Kerja), published in 2020 to coincide with the COVID-19 Pandemic, is quite significant as it has a positive impact on women's employability [57] and expand vocational training to vulnerable groups and in rural schools [57, 58]. Some studies critized the implementation of the pre-employment card because of the structural barriers related to anti-discrimination and childcare for female workers [59, 60].

This study also found that marriage is positively associated with employment. Among all gender and marital status groups, unmarried men exhibit the highest unemployment rate [23]. Urban residents are less likely to be employed than rural workers, indicating that higher labour competition in cities may offset urban job advantages. This makes urban workers earn significantly more than rural workers. Regarding this issue, Indonesia has prioritized upper-secondary vocational education to improve employment outcomes. However, the effectiveness of this policy has been mixed, with vocational graduates in urban areas faring better in terms of workplace-provided training and pensions compared to their rural counterparts [61].

The additional factors that might influence labor market outcomes, such as regional economic conditions, industry-specific trends, and access to professional networks have not been discussed in this study. Specific interventions such as subsidized childcare programs, targeted technical training for women, tax incentives for rural employers hiring vocational graduates, or mobile job centers is not available in SAKERNAS. Further research is expected to include these factors.

4. Conclusion

This research employs the Heckman two-step selection model to assess the impact of vocational education on employment opportunities and wage outcomes. The first-stage Probit model indicates that having higher vocational education, being married, and residing in rural areas increases the likelihood of employment. Conversely, older individuals, women with children, and urban residents face greater difficulties in securing jobs, as these factors are negatively associated with employment probability. In the second-stage wage regression, the findings reveal that female workers and vocational graduates living in rural areas experience lower earnings than their counterparts. This suggests that while vocational education enhances employability, challenges such as gender inequality in the labour market and competition for jobs in urban areas persist. To address these disparities and maximize the benefits of vocational education, the Indonesian government should adopt more targeted interventions. For instance, providing childcare support or flexible work arrangements could help increase labour force participation among women, particularly those with children. In rural areas, establishing partnerships between vocational schools and local industries—such as agriculture, agri-tech, or small-scale manufacturing—can expand employment opportunities and reduce dependency on urban job markets. Moreover, location-based wage incentives or relocation support programs could help balance regional disparities in income outcomes for vocational graduates.

Acknowledgement

The authors express their sincere gratitude to the Research Centre for the Economics of Industry, Services, and Trade, as well as the Research Organization for Governance, Economy, and Community Welfare (OR TKPEKM), National Research and Innovation Agency (BRIN) for their financial support from CRC 2024 OR TKPEKM and for providing Indonesia's 2023 National Labor Force Survey (SAKERNAS), which significantly facilitated the analysis in this study. The authors declare that there are no potential conflicts of interest related to this work's research, writing, or publication. The authors have made equal contributions to this paper.

REFERENCES

1. R. Hasan and K. R. L. Jandoc, *Workers' Earnings in the Philippines: Comparing Self-Employment with Wage Employment*, Asian Dev Rev, vol. 27, no. 01, pp. 43–79, 2010.
2. K. Parker, *Correcting for Sampling Bias in Education Production Functions*, South African Journal of Economics, vol. 75, no. 1, pp. 118–124, 2007.
3. T. Matkovic and I. Kogan, *All or Nothing? The Consequences of Tertiary Education Non-Completion in Croatia and Serbia*, Eur Sociol Rev, vol. 28, no. 6, pp. 755–770, 2012.
4. M. Triventi, *The role of higher education stratification in the reproduction of social inequality in the labor market*, Res Soc Stratif Mobil, vol. 32, pp. 45–63, 2013.
5. A. S. Alanazi and H. Benlaria, *Bridging Higher Education Outcomes and Labour Market Needs: A Study of Jouf University Graduates in the Context of Vision 2030*, Soc Sci, vol. 12, no. 6, p. 360, 2023.
6. G. Cherednichenko, *Employment and Labor Market Outcomes of College and Vocational School Graduates*, Voprosy Obrazovaniya / Educational Studies Moscow, no. 1, pp. 256–282, 2020.
7. G. Johnes, *'Don't Know Much About History...': Revisiting the Impact of Curriculum on Subsequent Labour Market Outcomes*, Bull Econ Res, vol. 57, no. 3, pp. 249–271, 2008.
8. J. Meer, *Evidence on the returns to secondary vocational education*, Econ Educ Rev, vol. 26, no. 5, pp. 559–573, 2007.
9. K. F. Zimmermann, *Youth Unemployment and Vocational Training*, Foundations and Trends® in Microeconomics, vol. 9, no. 1–2, pp. 1–157, 2013.
10. J. D. Hawley and J. Paek, *Developing human resources for the technical workforce: a comparative study of Korea and Thailand*, Int J Train Dev, vol. 9, no. 1, pp. 79–94, 2005.
11. D. Gillies, *Human Capital Theory in Education*, in Encyclopedia of Educational Philosophy and Theory, Singapore: Springer Singapore, 2015, pp. 1–5.
12. R. Qibthiyyah and A. J. Utomo, *Family Matters: Demographic Change and Social Spending in Indonesia*, Bull Indones Econ Stud, vol. 52, no. 2, pp. 133–159, 2016.
13. S. Susilo, E. Yuniashri, and R. P. Handrito, *Women's tendency in work involvement: A study of female labour force in Indonesia*, Economics & Sociology, vol. 17, no. 1, pp. 151–165, 2024.
14. J. Galimard, S. Chevret, C. Protopopescu, and M. Resche-Rigon, *A multiple imputation approach for MNAR mechanisms compatible with Heckman's model*, Stat Med, vol. 35, no. 17, pp. 2907–2920, 2016.

15. J. Zhao, H.-J. Kim, and H.-M. Kim, *New EM-type algorithms for the Heckman selection model*, Comput Stat Data Anal, vol. 146, p. 106930, 2020.
16. J. J. Heckman, *Sample Selection Bias as a Specification Error*, Econometrica, vol. 47, no. 1, p. 153, 1979.
17. H. Tauchmann, *Consistency of Heckman-type two-step estimators for the multivariate sample-selection model*, Appl Econ, vol. 42, no. 30, pp. 3895–3902, 2010.
18. W. Du, X. Wu, and H. Tong, *Fair Regression under Sample Selection Bias*, in 2022 IEEE International Conference on Big Data (Big Data), IEEE, 2022, pp. 1435–1444.
19. N. Lari, A. Awadalla, M. Al-Ansari, and E. Elmaghraby, *Determinants of female labor force participation: implications for policy in Qatar*, Cogent Soc Sci, vol. 8, no. 1, 2022.
20. T. K. McNamara and J. B. Williamson, *Race, Gender, and the Retirement Decisions of People Ages 60 to 80: Prospects for Age Integration in Employment*, The International Journal of Aging and Human Development, vol. 59, no. 3, pp. 255–286, 2004.
21. Yulmardi, Erfit, and D. N. Putra, *Employment status and women's fertility: Do working women have fewer children?*, Jurnal Perspektif Pembiayaan dan Pembangunan Daerah, vol. 12, no. 5, pp. 483–498, 2024.
22. I. Unay Gailhard and K. Kataria, *Economic crisis and labour force transition to inactivity: a comparative study in German rural and urban areas*, Studies in Agricultural Economics, vol. 116, no. 1, pp. 25–32, 2014.
23. B. McKinney, *The Effect of Family Organization and Race and Ethnicity on Women's Economic Status in New York State*, American Behavioral Scientist, vol. 53, no. 2, pp. 223–238, 2009.
24. L. T. Giang, A. Kikkawa, D. Park, and T. D. Nguyen, *Employment among older men and women in Vietnam: How their health played a role?*, Working with Older People, vol. 28, no. 3, pp. 233–246, 2024.
25. H. Zuo, *Formal and Informal Employment in China: Probability of Employment and Determinants of Monthly Wages*, Australian Economic Review, vol. 46, no. 4, pp. 405–423, 2013.
26. M. H. Kim, E. Y. V. López, and Y. Han, *Does education matter for the earnings of young adults?*, The New Educational Review, vol. 56, no. 2, pp. 172–183, 2019.
27. F. Pastore, *Returns to education of young people in Mongolia*, Postcommunist Econ, vol. 22, no. 2, pp. 247–265, 2010.
28. P. Böckerman, M. Haapanen, and C. Jepsen, *Back to school: Labor-market returns to higher vocational schooling*, Labour Econ, vol. 61, p. 101758, 2019.
29. H. Hansen, J. Rand, and N. W. Win, *The gender wage gap in Myanmar: Adding insult to injury?*, J Asian Econ, vol. 81, p. 101511, 2022.
30. H. E. Stokke, *The gender wage gap and the early-career effect: the role of actual experience and education level*, LABOUR, vol. 35, no. 2, pp. 135–162, 2021.
31. J. B. Rukundo, *Informal Employment and Monthly Wage Earnings: Evidence from Food Processing Firms*, African Journal of Business and Economic Research, vol. 14, no. 4, pp. 5–25, 2019.
32. A. Zhangaliyeva and M. Nakabayashi, *Legacy of the Czar: Complementarity between education and work in Russia*, J Appl Econ, vol. 26, no. 1, 2023.
33. A. Mamun and Md. Arfanuzzaman, *The Effects of Human Capital and Social Factors on The Household Income of Bangladesh: An Econometric Analysis*, J Econ Dev, vol. 45, no. 3, pp. 29–49, 2020.
34. M. Somantri and R. Pramudita, *Enhancing Industry's Role in Vocational Education: An Analysis of Challenges and Opportunities Based on a Literature Review*, in 2024 9th International STEM Education Conference (iSTEM-Ed), IEEE, 2024, pp. 1–7.
35. S. Chowdhury, E. Schulz, M. Milner, and D. Van De Voort, *Core employee based human capital and revenue productivity in small firms: An empirical investigation*, J Bus Res, vol. 67, no. 11, pp. 2473–2479, 2014.
36. W. J. Moore, R. J. Newman, and G. K. Turnbull, *The Experience-Earnings Profile: Productivity-Augmenting or Purely Contractual? Evidence from the UK*, J Labor Res, vol. 28, no. 3, pp. 417–435, 2007.
37. V. Mohanty and M. B. Kulkarni, *Employee Experience*, 2025, pp. 249–282.
38. M. Vilalta-Bufi, *On the industry experience premium and labor mobility*, Labour Econ, vol. 17, no. 3, pp. 547–555, 2010.
39. G. Brunello and S. Comi, *Education and earnings growth: evidence from 11 European countries*, Econ Educ Rev, vol. 23, no. 1, pp. 75–83, 2004.
40. A. E. Sales, M. E. Plomondon, D. J. Magid, J. A. Spertus, and J. S. Rumsfeld, *Assessing response bias from missing quality of life data: The Heckman method*, Health Qual Life Outcomes, vol. 2, no. 1, p. 49, 2004.
41. D. Naouri, C. Bussiere, and N. Pelletier-Fleury, *What Are the Determinants of Dental Care Expenditures in Institutions for Adults With Disabilities? Findings From a National Survey*, Arch Phys Med Rehabil, vol. 99, no. 8, pp. 1471–1478, 2018.
42. Y. Wang et al., *Optimal Caliper Width for Propensity Score Matching of Three Treatment Groups: A Monte Carlo Study*, PLoS One, vol. 8, no. 12, p. e81045, 2013.
43. P. C. Austin, *Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples*, Stat Med, vol. 28, no. 25, pp. 3083–3107, 2009.
44. T. Vold, H. Haave, and A. Kaloudis, *On Work Relevance of Adult Education: A Case Study Narrative*, Electronic Journal of Knowledge Management, vol. 18, no. 2, 2020.
45. C. M. Zellatifanny, Q. D. K. W, and Y. A. A. S, *Optimization of Digital Media to Escalate the Acceptance Level of Workforce Skills Information*, in 2021 2nd International Conference on ICT for Rural Development (IC-ICTRuDev), IEEE, 2021, pp. 1–6.
46. K. Charni, *Do employment opportunities decrease for older workers?*, Appl Econ, vol. 54, no. 8, pp. 937–958, 2022.
47. G. Öylü, A. Motel-Klingebiel, and S. Kelfve, *Age Differences in Unemployment Risk and Reemployment Outcomes in Late Working Life in Sweden*, J Aging Soc Policy, pp. 1–26, 2024.
48. V. Flek, M. Hála, and M. Mysíková, *Assessing the Job-Finding Probability of Older and Prime-Age Unemployed Workers*, Prague Economic Papers, vol. 29, no. 4, pp. 424–444, 2020.
49. S. Subanti, A. L. Riani, E. P. Lestari, A. R. Hakim, and S. Susiyanti, *The determinant of hours of work (the empirical case for Indonesian old–age)*, J Phys Conf Ser, vol. 1943, no. 1, p. 012140, 2021.
50. T. J. Cooke, *Gender role beliefs and family migration*, Popul Space Place, vol. 14, no. 3, pp. 163–175, 2008.

51. L. Gobillon, D. Meurs, and S. Roux, *Estimating Gender Differences in Access to Jobs*, J Labor Econ, vol. 33, no. 2, pp. 317–363, 2015.
52. A. Cinque, C. Poggi, and J. Miluka, *How Does Fertility Affect Female Employment? Evidence from Albania*, J Dev Stud, vol. 60, no. 8, pp. 1227–1245, 2024.
53. S. Chirkova and S. Kalantaryan, *The Gender-Specific Determinants of Labour Supply in Rural and Urban Contexts: Evidence from Armenia's Transition Economy*, in Gendering Post-Soviet Space, Singapore: Springer Singapore, 2021, pp. 137–159.
54. D. F. Vargas Calderón, J. F. Rios Obando, C. I. Rengifo Ortiz, and K. G. Serna Muñoz, *Brecha salarial de género: un análisis del mercado laboral de Cali-Colombia y su Área Metropolitana*, Revista Venezolana de Gerencia, vol. 29, no. Especial 11, pp. 667–683, 2024.
55. A. Dah and A. Fakih, *Decomposing Gender Wage Differentials Using Quantile Regression: Evidence from the Lebanese Banking Sector*, International Advances in Economic Research, vol. 22, no. 2, pp. 171–185, 2016.
56. N. T. Nguyen, V. A. Phung, K. Digdowiseiso, and N. Afni, *Bridging the gender wage gap: the role of education in developing economies*, Journal of Economic Studies, 2025.
57. M. S. Al Ayyubi, D. S. Pratomo, and F. Prasetyia, *Does pre-employment card program improve Indonesian youth labor market performance in pandemic era?*, Cogent Economics & Finance, vol. 11, no. 2, 2023.
58. M. Caliendo and S. Kopeinig, *Some Practical Guidance for The Implementation of Propensity Score Matching*, J Econ Surv, vol. 22, no. 1, pp. 31–72, 2008.
59. D. McKenzie, *How Effective Are Active Labor Market Policies in Developing Countries? A Critical Review of Recent Evidence*, World Bank Res Obs, vol. 32, no. 2, pp. 127–154, 2017.
60. P. Nguyen, F. Putra, M. Considine, and A. Sanusi, *Activation through welfare conditionality and marketisation in active labour market policies: Evidence from Indonesia*, Australian Journal of Public Administration, vol. 82, no. 4, pp. 488–506, 2023.
61. S. Choi, H. Li, K. Ogawa, and Y. Tanaka, *Secondary vocational education and decent work in Indonesia: differences between urban and rural areas*, International Journal of Training Research, vol. 22, no. 2, pp. 125–146, 2024.