

RESEARCH ARTICLE

Vertical education–job mismatch and worker welfare: Evidence from secondary school graduates in Indonesia

Kukuh Arisetyawan[†], Devanto Shasta Pratomo^{†*}, David Kaluge, and Wildan Syafitri

Department of Economics, Faculty of Economics and Business, Brawijaya University, Malang, East Java, Indonesia

Abstract

Education–job mismatch is an important labor market concern in developing countries, particularly among upper-secondary graduates navigating heterogeneous urban and rural employment opportunities. This study examines the factors associated with vertical education–job mismatch and the associations between mismatch and worker welfare among graduates of general and vocational upper-secondary schools in Indonesia, with particular attention to urban–rural differences. Using the August 2023 National Labor Force Survey (SAKERNAS), we estimate multinomial logit models for mismatch status and weighted linear regression models for worker welfare outcomes. Results indicate that only 49.3% of workers are well matched; 40.4% are educationally overqualified, and 10.3% are educationally underqualified. Age, gender, training participation, and urban residence are systematically associated with mismatch status. Educational underqualification is positively associated with the Job Quality Index in the baseline model, whereas educational overqualification is not statistically significant. The associations between mismatch and earnings vary across spatial contexts. Educational overqualification is positively associated with earnings in the rural reference group, although this association is offset in urban areas, whereas educational underqualification is negatively associated with rural earnings but positively associated with earnings in urban areas. Multinomial probit estimates are broadly consistent with the multinomial logit results, whereas selection-corrected analyses indicate that the welfare associations are sensitive to non-random sorting across mismatch categories. Favorable outcomes among some mismatched workers may therefore reflect unobserved skills, selective retention, and labor market adaptation rather than efficient initial allocation. The findings suggest that policies should combine locally differentiated worker–job matching systems with measures to improve job quality and expand productive employment opportunities.

Keywords: Education–job mismatch; Educational overqualification; Secondary education; Job quality; Earnings; Indonesia

[†]These authors contributed equally to the work.

***Corresponding author:**
Devanto Shasta Pratomo
(devanto@ub.ac.id)

Citation: Arisetyawan, K., Pratomo, D. S., Kaluge, D., & Syafitri, W. (2026). Vertical education–job mismatch and worker welfare: Evidence from secondary school graduates in Indonesia. *International Journal of Population Studies*. 12(5):026140073. <https://doi.org/10.36922/IJPS026140073>

Received: March 30, 2026

Revised: July 9, 2026

Accepted: August 12, 2026

Published online: September 1, 2026

Copyright: © 2026 Author(s). This is an Open-Access article distributed under the terms of the Creative Commons Attribution License, permitting distribution, and reproduction in any medium, provided the original work is properly cited.

Publisher's Note: AccScience Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

Upper secondary education plays a strategic role in facilitating the transition from school to work, particularly as many young individuals enter both formal and informal labor

markets at this stage. Within the human capital and school-to-work transition framework, secondary education functions not only as a terminal level of education for a substantial share of the workforce but also as an important foundation for developing general and vocational skills that support the transition into productive and adequately remunerated employment (McGowan & Andrews, 2017; Rocha *et al.*, 2025). This is especially evident in developing countries, where employment structures are often characterized by sectoral segmentation, limited creation of quality jobs, and substantial frictions in matching educational qualifications with labor market demands (Lu *et al.*, 2026).

In this context, education–job mismatch has become an increasingly important labor market issue. Mismatch not only reflects inefficiencies in worker–job allocation but is also associated with differences in job quality, productivity, labor mobility, and overall economic well-being (Allen & van der Velden, 2001; Wardani *et al.*, 2024). In Indonesia, the issue is particularly relevant because upper-secondary school graduates constitute a large share of the labor force and are among the groups most exposed to school-to-work transition challenges (Ariansyah *et al.*, 2024; Botezat *et al.*, 2024). Evidence from the 2023 National Labor Force Survey (SAKERNAS) indicates that vertical mismatch among these graduates remains substantial. Nationally, only about 49.3% of workers are employed in jobs that match their educational attainment, while 40.4% are educationally overqualified (higher-than-required education) and 10.3% are educationally underqualified (lower-than-required education). This pattern suggests that vertical mismatch at the secondary education level is not a marginal phenomenon but rather a structural feature of Indonesia's labor market.

Interestingly, the relationship between mismatch and worker welfare is not uniform. Descriptive evidence from SAKERNAS 2023 shows that well-matched workers earn an average monthly income of approximately 2.06 million Indonesian rupiah (US\$135.25), whereas both educationally overqualified and educationally underqualified workers earn approximately 2.26 million Indonesian rupiah (about US\$148). Their corresponding mean log monthly incomes are 14.6322 and 14.6288, respectively, compared with 14.5386 for well-matched workers. United States (U.S.) dollar equivalents are calculated using the 2023 annual average exchange rate of 15,236.9 Indonesian rupiah per U.S. dollar. The pattern also differs spatially: educationally underqualified workers have the lowest average income in rural areas but the highest in urban areas. These descriptive patterns suggest that the relationship between mismatch and economic well-being is neither linear nor

homogeneous; instead, the observed earnings premium or penalty appears to vary with local job structures, sectoral composition, and worker sorting (Hakim & Rosini, 2026; Khoiruddin *et al.*, 2024; Xiao *et al.*, 2023).

Although the literature on education–job mismatch is extensive, most existing studies focus on higher education graduates and primarily examine educational overqualification as a source of wage penalties or reduced productivity (McGowan & Andrews, 2017). Empirical evidence on upper-secondary graduates remains relatively limited, despite their dominant role during the early stages of the school-to-work transition in many developing countries (Blanck *et al.*, 2025).

Furthermore, much of the existing literature assumes that the relationship between mismatch and worker welfare is homogeneous across labor markets, overlooking substantial differences in local employment structures. Job opportunities, sectoral segmentation, informality, and labor market screening mechanisms often differ considerably between urban and rural areas (Xiao *et al.*, 2023). In Indonesia, relatively few studies simultaneously examine education–job mismatch among upper-secondary graduates, the factors associated with mismatch, and the associations between mismatch and worker welfare while explicitly incorporating urban–rural heterogeneity using nationally representative data. To our knowledge, no previous study has jointly examined these dimensions using nationally representative labor force data from Indonesia. This gap highlights the need for a more integrated and context-sensitive analysis.

Against this background, this study aims to (i) examine the factors associated with vertical education–job mismatch using variables related to individual characteristics, human capital, school-to-work transition, and job characteristics; and (ii) examine the associations between vertical education–job mismatch and worker welfare, particularly earnings and job quality, while accounting for heterogeneity between urban and rural labor markets.

The analysis focuses on vertical mismatch because it directly captures discrepancies between educational attainment and job requirements that are more closely associated with earnings and job quality than horizontal mismatch (McGuinness, 2006). This approach is particularly appropriate for developing-country settings, where data limitations and relatively less specialized education systems make horizontal mismatch substantially more difficult to measure accurately.

This study makes three important contributions to the literature. First, it treats education-based vertical mismatch as an imperfect but policy-relevant indicator of

worker–job alignment and explicitly distinguishes labor market adaptation from allocative efficiency (Choi *et al.*, 2024; Hakim & Rosini, 2026; Indrawati & Kuncoro, 2021). Unlike previous studies that often interpret favorable outcomes among mismatched workers as evidence against mismatch penalties, this study argues that such outcomes may instead reflect compensating skills, selective retention, or post-entry adaptation within an initially imperfect match rather than efficient labor market allocation. Second, it jointly analyzes the factors associated with vertical mismatch and its associations with job quality and earnings among upper-secondary graduates, a population that remains underexamined in the mismatch literature. Third, it evaluates whether these relationships differ across spatially segmented urban and rural labor markets, thereby providing new evidence on how local labor market conditions shape the observed associations between educational mismatch and worker welfare.

1.1. Conceptual and theoretical foundations

Education–job mismatch occurs when workers' educational attainment differs from the education required for their jobs. Vertical mismatch comprises educational overqualification (higher-than-required education) and educational underqualification (lower-than-required education), whereas horizontal mismatch concerns discrepancies between a worker's field of study and job tasks (Allen & van der Velden, 2001; Di Stasio, 2017). This study focuses on vertical mismatch because the occupational and educational classifications available in SAKERNAS permit a nationally consistent comparison between workers' educational attainment and the educational requirements of their occupations.

An education-based mismatch measure is not equivalent to a direct measure of skill mismatch. Workers holding the same educational credential may differ substantially in cognitive, technical, digital, and socioemotional skills, while occupations classified under the same International Standard Classification of Occupations (ISCO) code may involve different task requirements. Consequently, educationally underqualified workers may possess competencies acquired through work experience or training, whereas well-matched workers may still experience skill underutilization. Accordingly, the ISCO–International Standard Classification of Education (ISCED) classification should be interpreted as a proxy for formal qualification alignment rather than as a complete measure of multidimensional worker–job fit (Allen & De Weert, 2007; Green & McIntosh, 2007).

The conceptual framework of this study is primarily informed by three complementary theoretical perspectives. Human capital theory argues that education enhances

workers' productivity and improves labor-market rewards through the accumulation of productive knowledge and skills (Becker, 1964; Hartog, 2000). From this perspective, educational overqualification may indicate that acquired human capital is not fully utilized in the assigned job, whereas educational underqualification may reflect the presence of skills acquired outside formal education (Capsada-Munsech, 2017).

Assignment theory complements the human capital perspective by emphasizing that productivity and earnings depend not only on workers' qualifications but also on how efficiently workers are matched with jobs requiring their skills (Kim & Choi, 2018; Sattinger, 2012). Consequently, workers with similar educational attainment may experience different labor-market outcomes depending on the quality of worker–job matching.

Search-and-matching theory further explains why mismatch may persist even in competitive labor markets. Because workers and employers face imperfect information, search costs, mobility constraints, and limited vacancies, worker–job matches are often imperfect and gradually improve through labor-market adjustment (Mimani, 2025; Mortensen & Pissarides, 1999).

These theoretical perspectives suggest that education–job mismatch reflects not only differences in workers' human capital but also the interaction between worker characteristics, employer behavior, and labor-market institutions. They therefore provide the conceptual foundation for examining both the factors associated with mismatch and the associations between mismatch and worker welfare.

1.2. Factors associated with vertical mismatch

Human capital and search-and-matching perspectives suggest that education–job mismatch is likely to emerge during the school-to-work transition, particularly when workers have limited information about available jobs and employers have incomplete information about applicants' productivity (Farber & Gibbons, 1996; Mimani, 2025; Mortensen & Pissarides, 1999). Early-career workers may therefore accept imperfect job matches while searching for more suitable employment. As workers accumulate experience, develop occupation-specific skills, and obtain additional training, worker–job alignment may improve through occupational mobility and workplace learning. Employer learning may also contribute to this process as firms gradually obtain better information about workers' productivity from observed job performance (Altonji & Pierret, 2001; Farber & Gibbons, 1996). These processes imply that mismatch may evolve over the life cycle rather than remain fixed after labor-market entry.

Training participation may facilitate this adjustment by providing occupation-specific competencies and strengthening workers' productive capabilities (Blázquez Cuesta *et al.*, 2025; McGuinness *et al.*, 2018). Training may also serve a compensatory role when skills acquired through vocational or workplace learning enable workers to perform tasks beyond those implied by their formal educational qualifications. Recent evidence shows that some transversal skills (such as language, information technology, management, and social skills) can reduce the risk of overeducation when workers lack job-specific knowledge, suggesting that these skills may play a compensatory role (Blázquez Cuesta *et al.*, 2025). Moreover, signaling theory provides a related explanation: when direct information about productivity is limited, employers initially rely on educational credentials as imperfect signals of workers' productive capacity (Di Stasio, 2017; Spence, 1973). As information about actual performance becomes available, the importance of formal credentials in worker allocation may change.

Demographic characteristics and local labor-market conditions may influence access to vacancies, occupational mobility, employer screening, and worker–job sorting (Khoiruddin *et al.*, 2024; McGuinness *et al.*, 2018). In particular, gender, disability status, marital status, labor-market informality, and geographic location may shape workers' opportunities to obtain jobs corresponding to their qualifications. Regional labor-market structures are also relevant because differences in labor demand, sectoral composition, and the availability of suitable jobs affect the likelihood of both educational overqualification and underqualification (Khoiruddin *et al.*, 2024).

This evidence suggests that mismatch reflects the joint influence of worker characteristics, human-capital formation, school-to-work transition processes, and labor-demand conditions rather than educational attainment alone (Blázquez Cuesta *et al.*, 2025; Cortadas-Guasch, 2024; Khoiruddin *et al.*, 2024). It also provides several mechanisms through which these factors may operate, including imperfect information at labor-market entry, employer learning, workplace skill accumulation, occupational sorting, and differences in local employment opportunities. Based on these arguments, this study examines whether individual characteristics, human-capital investments, school-to-work transition experiences, and spatial characteristics are systematically associated with the mismatch status of upper-secondary graduates. Therefore, the following hypothesis is proposed:

H1. Individual characteristics, human-capital investments, school-to-work transition experiences, and spatial characteristics are systematically associated with

whether upper-secondary graduates are well matched, educationally overqualified, or educationally underqualified.

1.3. Vertical mismatch and worker welfare

Assignment theory predicts that worker welfare depends on the quality of worker–job matching because productivity and compensation are jointly determined by worker qualifications and job requirements (Sattinger, 2012). Educational overqualification may reduce the utilization of acquired knowledge and weaken returns to schooling, whereas educational underqualification may expose workers to tasks exceeding their formal educational preparation. The consequences of mismatch, however, depend not only on formal educational differences but also on the extent to which workers' actual skills can be utilized within their jobs (Kang & Hong, 2025; Yin & Wang, 2026).

Empirical evidence shows that the labor-market consequences of vertical mismatch are not uniform. Educational overqualification is frequently associated with lower earnings relative to adequately matched workers with similar education, although the magnitude of the penalty varies across worker groups and labor-market settings (Green & Zhu, 2010; Santoso *et al.*, 2022). Recent evidence further shows that the consequences of educational mismatch depend on workers' skill proficiency and occupational context, suggesting that formal educational mismatch and actual skill mismatch should not be treated as equivalent (Kang & Hong, 2025). Educationally underqualified workers may also achieve better outcomes when work experience, occupation-specific skills, or workplace training compensate for lower formal education. These findings suggest that the welfare consequences of mismatch depend partly on productive characteristics that are not fully captured by educational credentials.

Different mechanisms may explain why workers in mismatch categories do not always experience poorer labor-market outcomes. Workers may accumulate occupation-specific skills through workplace learning, employer-provided training, and firm-specific human capital, enabling them to perform tasks beyond those implied by their formal educational qualifications (Becker, 1964). Employers may also revise their assessment of worker productivity as they observe on-the-job performance, reducing their reliance on educational credentials and placing greater weight on demonstrated capabilities (Altonji & Pierret, 2001; Farber & Gibbons, 1996). Recent evidence supports the importance of this distinction between formal qualifications and productive skills in explaining heterogeneous outcomes among

mismatched workers (Kang & Hong, 2025). In addition, workers who perform poorly in mismatched jobs may be more likely to leave or move to other occupations, whereas successful matches are more likely to persist (McGuinness, 2006).

These mechanisms also suggest that worker welfare should be evaluated across multiple dimensions rather than earnings alone. In this study, worker welfare is represented by two labor-market outcomes: job quality and earnings. Earnings capture monetary returns to employment, whereas job quality reflects broader aspects of employment, including employment security, social protection, working-time arrangements, and income adequacy. Educational mismatch and job quality are therefore related but conceptually distinct: a worker may be mismatched according to formal educational requirements while still holding a job with relatively better employment conditions. Recent research also highlights the importance of skill utilization in linking educational mismatch to broader employment outcomes, indicating that formal mismatch does not necessarily translate directly into poor job quality (Yin & Wang, 2026).

Based on these arguments, we examine the following two hypotheses:

- H2. Educational overqualification and educational underqualification are associated with job-quality levels that differ from those of well-matched workers.
- H3. Educational overqualification and educational underqualification are associated with earnings that differ from those of well-matched workers.

1.4. Spatial labor-market segmentation

Urban and rural labor markets differ in employment opportunities, occupational structure, and sectoral composition, creating different conditions for worker-job matching. Recent Indonesian evidence shows that regional labor-market characteristics, including urban-rural location and sectoral economic structure, are systematically associated with education-job mismatch (Khoiruddin *et al.*, 2024).

Spatial differences may also shape how qualifications and skills are valued after workers enter employment. Differences in occupational structure and employment opportunities may affect the extent to which mismatched workers can utilize their skills or compensate for formal qualification gaps through experience and occupation-specific competencies. Employer learning may further reduce reliance on formal credentials as firms obtain information about workers' productivity from observed job performance (Altonji & Pierret, 2001; Farber & Gibbons, 1996). These mechanisms suggest that the same

type of educational mismatch may have different welfare consequences across local labor markets.

Hakim and Rosini (2026) also show that the wage associations of vertical educational mismatch in Indonesia vary across spatial conditions, indicating that the returns associated with overqualification and underqualification depend partly on local labor-market context. Spatial segmentation may thus shape both the likelihood of mismatch and its association with worker welfare. Based on these arguments, this study proposes the following hypothesis:

- H4. The associations between vertical education-job mismatch, job quality, and earnings differ between urban and rural labor markets.

2. Data and methods

2.1. Data source and analytical sample

This study uses the August 2023 SAKERNAS, a nationally representative survey covering all 38 provinces of Indonesia. SAKERNAS applies a two-stage stratified sampling design in which census blocks serve as primary sampling units, and households serve as secondary sampling units. The survey provides information on demographic characteristics, education, occupation, employment status, working hours, earnings, training, internship experience, and the use of digital technology. Sampling weights are applied to preserve population representativeness.

The analytical sample is restricted to employed individuals whose highest educational attainment is upper-secondary education, including general senior secondary schools and vocational senior secondary schools. The descriptive sample contains 143,459 workers with complete covariate information. Effective sample sizes vary because mismatch status, job-quality indicators, and monthly earnings are not observed for all respondents, resulting in different estimation samples across models. The multinomial logit model includes 141,558 observations, the Job Quality Index (JQI) regressions include 65,979 observations, and the log monthly income regressions include 126,786 observations.

2.2. Measurement of vertical education-job mismatch

Vertical education-job mismatch is defined as a discrepancy between a worker's educational attainment and the educational level required for the occupation. The classification follows the International Labour Organization framework implemented through ILOSTAT, which links ISCO-08 to ISCED-11. Workers are classified as well matched when their educational

attainment corresponds to the occupational requirement, educationally overqualified when their attainment exceeds the requirement, and educationally underqualified when their attainment is below the requirement. The categories are mutually exclusive. This classification follows the normative approach commonly adopted in the international mismatch literature and provides a transparent benchmark for cross-country comparison.

Let M_i denote worker i 's mismatch status: $M_i = 0$ for a well-matched worker, $M_i = 1$ for educational overqualification, and $M_i = 2$ for educational underqualification. Well-matched workers constitute the reference category in all models. Because this normative classification captures formal educational alignment rather than direct skill use, the results are interpreted as evidence about qualification mismatch, not as a complete measure of skill mismatch.

2.3. Outcome variables and covariates

Worker welfare is assessed using two outcomes: the JQI and log monthly income. The JQI is the unweighted mean of four binary components: employment-related social-protection coverage; contract security; non-excessive weekly working hours, using a 48-h threshold; and income adequacy, defined as monthly income at or above the weighted median within the worker's province-by-urban/rural cell. The index ranges from 0 to 1, with higher values indicating better job quality. Monthly income is expressed in natural logarithms in the regression analysis to reduce right-skewness and permit semi-elasticity interpretations.

The covariates include age and age squared, gender, marital status, disability status, training participation, internship experience, urban residence, formal employment, weekly working hours, and digital technology use. The mismatch model includes individual, human-capital, school-to-work transition, and spatial covariates, whereas the welfare models additionally incorporate job-related characteristics. Province fixed effects are included to account for unobserved regional heterogeneity common to workers within the same province.

2.4. Econometric strategy and robustness checks

The determinants of mismatch are estimated using a multinomial logit model because mismatch status consists of three unordered categories. For $j = 1, 2$, the probability of category j relative to the well-matched base outcome is specified as:

$$\Pr(M_i = j | X_i) = \frac{\exp(X_i' \beta_j)}{1 + \sum_{k=1}^2 \exp(X_i' \beta_k)}, j = 1, 2 \quad (1)$$

Equivalently, the log-odds equation is:

$$\ln[\Pr(M_i = j) / \Pr(M_i = 0)] = \alpha_j + X_i' \beta_j + \delta_{pj} \quad (2)$$

where X_i contains the observed covariates and δ_{pj} denotes province fixed effects. Coefficients are reported for educational overqualification and educational underqualification relative to adequate matching.

The independence of irrelevant alternatives (IIA) assumption is assessed using a Hausman-based diagnostic. Because the diagnostic can be unstable when the covariance-difference matrix is not positive definite, it is treated as specification evidence rather than a sole decision rule. A multinomial probit model, which permits a more flexible correlation structure across outcomes, is estimated as a robustness check.

The associations between mismatch and worker welfare are estimated using weighted linear regressions of the following form:

$$Y_i = \alpha + \beta_1 \text{Overqual}_i + \beta_2 \text{Underqual}_i + \gamma' X_i + \theta' Z_i + \delta_p + \varepsilon_i \quad (3)$$

where Y_i denotes either the JQI or log monthly income; Overqual_i and Underqual_i are mismatch indicators; X_i contains individual and spatial controls; Z_i contains job-related controls; and δ_p denotes province fixed effects. Urban-rural heterogeneity is evaluated by adding Urban_p , $\text{Overqual}_i \times \text{Urban}_p$, and $\text{Underqual}_i \times \text{Urban}_p$, and by estimating separate urban and rural regressions. Sampling weights are applied, and standard errors are clustered at the primary sampling-unit level.

Unobserved ability, motivation, work experience, networks, and labor-market sorting may jointly affect mismatch status and welfare outcomes. The cross-sectional estimates are therefore interpreted as conditional associations rather than causal effects. Robustness analyses include alternative robust standard errors, sequential model specifications, separate urban and rural samples, and a multinomial probit model. A Lee selection correction based on the multinomial first stage is also used as a sensitivity analysis for non-random sorting across mismatch categories. The Lee-corrected estimates rely on functional-form assumptions and are not treated as definitive causal identification.

3. Results

3.1. Descriptive results

Table 1 reports the summary statistics of the main variables used in the analysis, based on 143,459 workers with upper-secondary education with valid covariate information. The sample consists of workers with upper-secondary education and reflects a broad distribution across the working-age population. The average age is 37.24 years (SD = 12.45), indicating substantial variation across

different stages of working life. In terms of demographic composition, male workers account for 65.99% of the sample, while 67.86% are married, suggesting a workforce largely composed of male individuals in more established household stages. Workers with disabilities represent only 2.60% of the sample, indicating a relatively small but potentially vulnerable group.

With regard to school-to-work transition and skill development, internship experience is limited, with only 1.50% of workers reporting such experience. In contrast, participation in training programs reaches 30.12%, indicating a greater reliance on formal training, although work-based transition pathways remain relatively underdeveloped. From a spatial and institutional perspective, 58.82% of workers reside in urban areas, while 46.07% are employed in the formal sector. Average working hours amount to 40.93 h per week, close to standard full-time employment. Additionally, 66.65% of workers report access to or use of digital technology, suggesting relatively high digital penetration in the labor force.

Overall, these statistics indicate a labor market characterized by a predominance of male and married workers, a strong urban concentration, and increasing digital technology use. However, limited internship participation and uneven access to training highlight potential constraints in school-to-work transition mechanisms.

3.2. Factors associated with education–job mismatch

Prior to performing multinomial logit analysis, the Hausman test was conducted to test the IIA assumption (Table A1 summarizes the test outcomes). The evidence is somewhat mixed. The test rejects the IIA assumption when the matched category is omitted, whereas the test for the educationally underqualified category does not reject IIA. For the educationally overqualified category, the test produces a negative chi-square statistic, suggesting that the asymptotic assumptions underlying the test are not fully satisfied; accordingly, this comparison should be interpreted with caution.

These results are regarded as indicative rather than conclusive evidence regarding the validity of the IIA assumption. Given the well-known sensitivity of the Hausman test in multinomial settings, particularly in large samples and complex specifications, the evidence is not considered sufficient to dismiss the multinomial logit framework altogether. To further assess the robustness of the findings, a multinomial probit model, which relaxes the IIA assumption, is also estimated (Table A2). The multinomial probit results display broadly similar

coefficient signs and significance patterns for key variables such as age, gender, training participation, and urban residence, suggesting that the main substantive conclusions remain stable across model specifications. Accordingly, the multinomial logit model is retained as the primary specification because of its parsimonious structure and interpretability, while the multinomial probit estimates are presented as a robustness check to address potential issues related to the IIA assumption.

The multinomial logit estimates (Table 2) show that age is positively associated with the likelihood of both educational overqualification and educational underqualification relative to being well-matched, while the negative and significant squared term indicates a nonlinear relationship. This suggests that mismatch is more prevalent in the early stages of the career and declines with experience, consistent with job search and career mobility dynamics.

Gender is significantly associated with mismatch. Male workers are significantly more likely to be both educationally overqualified and educationally underqualified, pointing to gendered sorting mechanisms in the labor market, potentially driven by occupational segregation or differential access to jobs. In contrast, marital status is not significantly associated with mismatch. Workers with disabilities are less likely to be educationally overqualified, although no significant association is found for educational underqualification.

With respect to human capital, training participation shows contrasting associations with educational overqualification and underqualification. It is associated with a lower probability of educational overqualification but a higher probability of educational underqualification. By contrast, internship experience does not have a significant association, indicating a limited role of work-based transition pathways in shaping mismatch outcomes.

Spatial factors are also important. Workers in urban areas are significantly more prone to experiencing both educational overqualification and educational underqualification, indicating that urban labor markets, despite offering more opportunities, pose greater competition and more complex matching processes.

These findings highlight that education–job mismatch is significantly associated with life-cycle dynamics, gender, human capital characteristics, and spatial labor-market conditions. The coexistence of educationally overqualified and educationally underqualified workers further indicates that mismatch occurs in both directions simultaneously, reflecting not only the underutilization of education but also the allocation of workers to jobs exceeding their formal

Table 1. Descriptive statistics

Variables	Definition	Observations	Mean	Standard deviation
Age	Age of worker in years (continuous)	143,459	37.239	12.450
Age squared	Square of age (continuous)	143,459	1541.746	992.080
Male	Dummy (1 = male worker, 0 = female worker)	143,459	0.660	0.474
Married	Dummy (1 = married, 0 = otherwise)	143,459	0.679	0.467
Disability	Dummy (1 = worker has disability, 0 = otherwise)	143,459	0.026	0.159
Internship	Dummy (1 = attended internship, 0 = otherwise)	143,459	0.015	0.122
Training participation	Dummy (1 = attended certified training, 0 = otherwise)	143,459	0.301	0.459
Urban	Dummy (1 = lives/works in urban area, 0 = rural area)	143,459	0.588	0.492
Formal	Dummy (1 = employed in formal sector, 0 = informal sector)	143,459	0.461	0.498
Hours worked per week	Number of working hours per week (continuous)	143,459	40.934	17.557
Digital technology use	Dummy (1 = uses any digital technology/internet for work, 0 = otherwise)	143,459	0.666	0.471

qualifications. This underscores the multidimensional and structural nature of mismatch, rather than a simple deviation from optimal allocation.

3.3. Education–job mismatch, job quality, and earnings

Table 3 examines the relationship between mismatch status, job quality, and income, while accounting for urban–rural heterogeneity. In the job quality model, age is positively and significantly associated with job quality, with a negative squared term, indicating a concave relationship in which job quality improves with experience but at a diminishing rate over the life cycle. Male and married workers have higher predicted job-quality scores, while disability is linked to lower job quality outcomes. Training shows a strong positive association, whereas internship experience is not statistically significant. Among job characteristics, formal employment and access to digital technology are among the variables most strongly associated with job quality, suggesting the importance of structural labor-market characteristics in relation to job quality.

The associations between mismatch and job quality are heterogeneous. In the baseline specification, one mismatch

category shows no significant association, while the other is positively associated with job quality. However, once interaction terms are introduced, the estimates reveal significant urban–rural differences in the associations between mismatch status and job quality, with the direction and magnitude varying across mismatch categories. This suggests that mismatch in urban labor markets may be accompanied by compensating advantages, such as better working conditions or access to higher-quality jobs despite formal misalignment.

Turning to the income model, age again shows a concave relationship with earnings. Male and married workers earn higher incomes, while disability is associated with a wage penalty. Training is positively associated with income, although the effect is relatively modest, while internship experience is negatively related to earnings. Urban residence, formal employment, and digital technology use are positively associated with income, suggesting the importance of structural advantages.

The associations between mismatch and income vary markedly across categories and spatial contexts. In the baseline model, educational underqualification is positively associated with income, whereas educational

Table 2. Multinomial logit estimates of vertical education–job mismatch

Variables (Base outcome: matched.)	Overqualification	Underqualification
Age	0.042*** (8.40)	0.047*** (6.55)
Age squared	−0.001*** (−11.57)	−0.000*** (−5.80)
Male (=1)	0.917*** (48.86)	0.208*** (7.19)
Married (=1)	−0.021 (−0.88)	0.045 (1.21)
Disability (=1)	−0.156** (−2.60)	−0.123 (−1.47)
Training participation (=1)	−0.124*** (−6.17)	0.424*** (14.65)
Internship (=1)	0.021 (0.27)	0.156 (1.48)
Urban (=1)	0.216*** (10.06)	0.221*** (6.00)
Constant	−1.762*** (−17.71)	−3.250*** (−22.41)
Province fixed effects	Yes	Yes
Observations	141,558	141,558
Pseudo R ²	0.037	

Notes: Entries are reported as coefficients with *t*-statistics in parentheses. Province fixed effects are included in all specifications, reported in Table A3. Statistical significance is denoted as follows: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

overqualification is not statistically significant. In the interaction model, overqualification has a positive coefficient for the rural reference group, but the negative overqualification-by-urban interaction offsets this association in urban areas. Conversely, underqualification is negatively associated with income in the rural reference group, while the positive underqualification-by-urban interaction produces a substantial urban premium. The combined coefficients therefore show that national averages conceal distinct urban and rural patterns.

Overall, the welfare associations are category-specific and context-dependent. Mismatch does not uniformly reduce job quality or earnings, but the positive coefficients observed for some groups should not be interpreted as evidence of efficient initial allocation. They may reflect compensating skills, selective retention, workplace learning, and differences in local job structures.

The robustness analyses support the main pattern of heterogeneity. Alternative standard-error specifications yield substantively similar conclusions (Table A4). Sequential income and job-quality models show that mismatch coefficients change after job characteristics and urban interactions are added, confirming the importance of model specification and spatial context (Tables A5 and A6). Separate urban and rural regressions likewise show that the welfare associations differ across local labor markets (Table A7).

As an additional sensitivity analysis, the Lee selection-corrected model addresses non-random sorting into well-matched, overqualification, and underqualification categories. The first-stage multinomial logit (Table A8) shows that mismatch status is systematically associated with observed individual, human-capital, and spatial characteristics. The correction terms are jointly significant in the income and JQI equations, indicating empirically relevant selection. After correction, overqualification is negatively associated with income and job quality, while the underqualification estimates are less stable and remain context-dependent (Tables A9 and A10). These results suggest that favorable baseline associations partly reflect non-random sorting. Because the correction relies on functional-form assumptions and the data are cross-sectional, the estimates remain sensitivity evidence rather than causal effects.

4. Discussion

4.1. Factors associated with vertical mismatch

The results provide substantial, although not uniform, support for H1, indicating that vertical education–job mismatch among upper-secondary graduates is systematically associated with individual characteristics, human-capital investment, and spatial labor-market conditions. The direction and magnitude of these associations differ between educational overqualification

Table 3. Job quality index and log monthly income regressions

Variables	(1)	(2)	(3)	(4)
	Job quality index	Log monthly income	Log monthly income	Job quality index
Main covariates				
Age	0.012*** (10.99)	0.029*** (16.60)	0.030*** (16.95)	0.012*** (11.01)
Age squared	−0.000*** (−10.39)	−0.000*** (−13.07)	−0.000*** (−13.36)	−0.000*** (−10.41)
Male (=1)	0.038*** (9.45)	0.383*** (56.70)	0.380*** (56.77)	0.038*** (9.43)
Married (=1)	0.042*** (8.78)	0.082*** (10.66)	0.083*** (10.85)	0.042*** (8.82)
Disability (=1)	−0.034* (−2.43)	−0.101*** (−4.50)	−0.100*** (−4.46)	−0.034* (−2.44)
Training participation (=1)	0.068*** (17.24)	0.016** (2.59)	0.017** (2.69)	0.068*** (17.23)
Internship (=1)	0.000 (0.03)	−0.051* (−2.25)	−0.052* (−2.33)	0.001 (0.04)
Urban (=1)	0.009 (1.74)	0.113*** (14.59)	0.121*** (12.39)	−0.009 (−1.30)
Mismatch status and interactions				
Overqualification (=1)	0.007 (1.56)	0.008 (1.32)	0.071*** (6.68)	−0.014 (−1.77)
Underqualification (=1)	0.049*** (8.62)	0.037** (2.98)	−0.160*** (−6.69)	0.030** (2.99)
Overqualification × Urban			−0.088*** (−6.92)	0.028** (3.10)
Underqualification × Urban			0.266*** (9.61)	0.026* (2.17)
Labor and digital covariates				
Formal employment (=1)	0.296*** (62.89)	0.176*** (25.42)	0.181*** (26.27)	0.295*** (61.88)
Hours worked per week	0.000 (1.35)	0.012*** (54.94)	0.012*** (54.77)	0.000 (1.38)
Digital technology use (=1)	0.084*** (16.24)	0.229*** (31.97)	0.230*** (32.08)	0.084*** (16.28)
Constant	−0.378*** (−18.08)	12.632*** (351.53)	12.618*** (352.32)	−0.365*** (−17.23)
Province fixed effects	Yes	Yes	Yes	Yes
Observations	65,979	126,786	126,786	65,979
Adjusted R ²	0.175	0.279	0.282	0.176

Notes: The dependent variables are the Job Quality Index in Models (1) and (4) and log monthly income in Models (2) and (3). Models (3) and (4) include interaction terms between mismatch status and urban residence. Coefficients are reported with *t*-statistics in parentheses. Well-matched workers are the reference category for mismatch status, and rural workers are the reference category for the urban interaction terms. All models include province fixed effects, apply sampling weights, and cluster standard errors at the primary sampling-unit level. Statistical significance: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Robustness checks are reported in Table A4. Source: Authors' calculation from SAKERNAS 2023.

and underqualification, reinforcing the argument that mismatch reflects the interaction between worker characteristics, skill formation, occupational opportunities, and labor-market structures. This interpretation is consistent with recent Indonesian evidence identifying demographic and regional characteristics as important correlates of education–job mismatch (Khoiruddin *et al.*, 2024; Santoso *et al.*, 2022; Wardani *et al.*, 2024).

The positive age coefficient and negative squared term indicate a nonlinear relationship between age and both forms of mismatch. This finding supports the life-cycle component of H1 and is consistent with the search-and-matching and employer-learning mechanisms discussed in Section 1. At labor-market entry, workers may accept imperfect matches because information about occupational

requirements and suitable vacancies is limited (Mortensen & Pissarides, 1999). As experience accumulates, employer learning, workplace learning, and occupation-specific skill acquisition can improve information about worker productivity and occupational fit, facilitating corrective mobility and better worker–job alignment (Altonji & Pierret, 2001; Farber & Gibbons, 1996). Selective retention may also contribute, as workers who perform successfully in mismatched jobs are more likely to remain, whereas unsuccessful matches are more likely to dissolve. The nonlinear age profile may therefore reflect several forms of labor-market adaptation rather than a single adjustment process. Because SAKERNAS is cross-sectional, however, it should not be interpreted as evidence that the same workers move out of mismatch as they age.

Gender differences provide further support for the individual-characteristics component of H1. Male workers are more likely to experience both educational overqualification and underqualification relative to adequate matching. A plausible mechanism is gendered occupational sorting, whereby men and women face different occupational opportunity sets, sectoral concentrations, and mobility patterns. Recent Indonesian evidence also identifies gender as an important correlate of education–job mismatch, although the direction varies across educational groups and definitions of mismatch (Khoiruddin *et al.*, 2024; Wardani *et al.*, 2024). Such variation suggests that gendered mismatch depends on occupational composition and educational level rather than following a uniform pattern across the labor market.

The disability result points to a different selection mechanism. Disability is negatively associated with educational overqualification but not significantly associated with underqualification. This should not necessarily be interpreted as better worker–job matching. One possible explanation is that barriers to recruitment, workplace accessibility, and workplace accommodation restrict the employment opportunities available to workers with disabilities (Ananian & Dellaferrera, 2024). The lower observed probability of overqualification may therefore reflect pre-employment selection rather than more efficient matching. Because the analysis is conditional on employment, disadvantages operating through non-participation or unsuccessful job search are not directly observed.

Training participation provides evidence for the human-capital component of H1, but the contrasting coefficients suggest two complementary mechanisms. Its negative association with overqualification is consistent with an improved-matching mechanism, whereby job-relevant competencies allow workers to obtain positions that make fuller use of their education. Recent Indonesian evidence similarly highlights the role of vocational and labor-market-oriented skill formation in reducing vertical mismatch among young workers (Putranto *et al.*, 2024). By contrast, the positive association with underqualification is consistent with a compensatory human-capital mechanism. Training and workplace learning may provide occupation-specific competencies that enable workers to perform jobs formally requiring higher education. Employers may consequently place greater weight on demonstrated performance and acquired competencies than on educational credentials alone, consistent with employer-learning arguments (Altonji & Pierret, 2001). Thus, training may both improve formal worker–job alignment and enable productive workers to cross formal

qualification boundaries, highlighting the distinction between educational mismatch and actual skill mismatch.

Internship experience, by contrast, is not significantly associated with either form of mismatch, providing less support for the school-to-work transition component of H1. Although internships can provide occupational experience and improve employers' perceptions of applicants' skills and job knowledge, their labor-market value may depend on how the internship is designed and implemented (Jerez-Gómez *et al.*, 2023; Tobback *et al.*, 2024). The insignificant coefficient may therefore reflect heterogeneity in internship characteristics rather than the absence of a potential role for work-based learning.

Finally, urban residence is positively associated with both educational overqualification and underqualification, supporting the spatial labor-market component of H1. Urban labor markets may combine greater occupational diversity with stronger competition and more complex worker–job sorting. Recent Indonesian evidence similarly identifies regional labor-market characteristics as important correlates of mismatch (Khoiruddin *et al.*, 2024; Santoso *et al.*, 2022). The positive urban coefficients may therefore reflect a spatial sorting mechanism: thicker labor markets generate more opportunities but also more complex allocation processes across jobs with different educational requirements. This interpretation provides the basis for the urban–rural heterogeneity in worker welfare examined under H4.

Overall, the findings provide substantial but not complete support for H1. The observed patterns are consistent with complementary mechanisms, including life-cycle adaptation, employer learning, selective retention, occupational sorting, compensatory skill formation, and spatial sorting.

4.2. Mismatch, job quality, and earnings

In terms of worker welfare, the results provide partial support for H2 and H3. The associations between vertical education–job mismatch and worker welfare differ across mismatch categories and between job quality and earnings. Educational underqualification is positively associated with the JQI, whereas educational overqualification is not statistically significant. For earnings, educational underqualification is positively associated with earnings in the baseline specification, while the coefficient for educational overqualification is not statistically significant.

For job quality, the strong positive associations of formal employment and digital technology use show that educational matching represents only one dimension of worker welfare. Educational mismatch concerns the

correspondence between qualifications and occupational requirements, whereas the JQI reflects broader employment conditions. Workers may therefore be educationally mismatched while still holding relatively high-quality jobs. Recent studies also highlight the distinction between educational mismatch, skill utilization, and broader employment outcomes (Kang & Hong, 2025; Yin & Wang, 2026).

The positive association between educational underqualification and job quality may reflect differences between formal qualifications and workers' actual productive capacities. The mismatch measure does not capture practical skills, accumulated experience, occupation-specific competencies, or firm-specific human capital. Workers classified as underqualified may therefore possess skills that compensate for their lower formal educational attainment, while workplace learning may further enable them to acquire competencies required for tasks normally associated with higher levels of education. Recent evidence shows that the consequences of educational mismatch vary with workers' underlying skill proficiency, reinforcing the distinction between formal qualifications and actual skills (Kang & Hong, 2025).

Employer learning and selective retention may reinforce this pattern. As employers observe workers' performance, formal educational credentials become only one source of information about productivity (Altonji & Pierret, 2001; Farber & Gibbons, 1996). Underqualified workers who demonstrate sufficient productivity may remain in jobs with higher formal educational requirements and obtain better employment conditions, whereas workers who cannot compensate for their qualification gaps are more likely to change jobs or occupations. The positive association with job quality may therefore reflect compensating skills, workplace learning, and positive selection among workers who remain in jobs with higher formal educational requirements, rather than a beneficial effect of underqualification itself.

The non-significant association between educational overqualification and job quality points to a different process. Overqualified workers may have relatively better objective employment conditions while simultaneously experiencing underutilization of their knowledge and skills in jobs requiring less education. These opposing features may offset each other in the composite JQI. Recent evidence highlights skill utilization as an important channel through which educational mismatch is related to employment outcomes (Yin & Wang, 2026). The absence of a significant JQI association should therefore not be interpreted as evidence that overqualification is inconsequential, since some of its costs may arise through

skill underutilization or other aspects of work not fully captured by the index.

The earnings results provide partial support for H3. Educational underqualification is positively associated with earnings in the baseline model, whereas educational overqualification does not differ significantly from adequate matching. The positive association for underqualified workers is consistent with the same compensating-skill and selection mechanisms observed for job quality. Workers who enter occupations formally requiring higher education may possess experience, training, or occupation-specific skills that are rewarded in the labor market despite their lower formal qualifications. Employer learning may also allow demonstrated productivity to substitute partly for educational credentials over time. Recent evidence indicates that wage consequences of educational mismatch depend substantially on workers' actual skills and occupational context rather than educational credentials alone (Kang & Hong, 2025).

The absence of a significant earnings difference for educationally overqualified workers may reflect offsetting influences. Their upper-secondary qualifications may provide productive skills that continue to be valued by employers even in occupations formally requiring less education, while the lower educational requirements of those occupations may limit the returns to those qualifications (Kang & Hong, 2025; Sattinger, 2012). The resulting coefficient therefore reflects the net association between these competing forces.

The findings overall provide partial support for H2 and H3. Educational underqualification is positively associated with both job quality and earnings, whereas educational overqualification shows no statistically significant association with either outcome in the baseline specifications. The results are consistent with compensating skills, workplace learning, employer learning, selective retention, and differences between educational qualifications and actual skill utilization.

4.3. Urban–rural heterogeneity

The urban–rural interaction estimates provide support for H4, showing that the associations between vertical education–job mismatch and worker welfare differ across local labor markets. This spatial variation is consistent with the argument developed above that differences in occupational structure, informality, and employer screening can influence the consequences of mismatch. Recent Indonesian evidence also indicates that regional and spatial characteristics are important in explaining education–job mismatch and its labor-market outcomes (Hakim & Rosini, 2026; Khoiruddin *et al.*, 2024).

For job quality, the positive interaction between educational overqualification and urban residence indicates that the positive association is stronger in urban than in rural labor markets. Urban labor markets may offer a wider range of occupations, allowing overqualified workers to find jobs with better working conditions even when their formal education exceeds occupational requirements. These workers may also possess skills that are not fully reflected in formal occupational classifications. Recent evidence shows that the consequences of educational mismatch vary with workers' underlying skill proficiency (Kang & Hong, 2025), while Indonesian evidence indicates that mismatch patterns differ across regional and urban–rural labor market conditions (Khoiruddin *et al.*, 2024).

Educational underqualification shows a different pattern. Its positive interaction with urban residence for job quality suggests that workers whose formal education falls below occupational requirements fare relatively better in urban labor markets.

Greater occupational diversity may provide more opportunities for workers to compensate for formal qualification gaps through experience, workplace learning, and occupation-specific skills. Employers may also place greater weight on demonstrated performance as they gain information about workers' productivity, consistent with employer-learning mechanisms (Altonji & Pierret, 2001; Farber & Gibbons, 1996). Recent evidence further shows that the consequences of educational mismatch depend partly on workers' actual skills and occupational context, supporting this interpretation (Kang & Hong, 2025).

The earnings estimates show a similarly pronounced spatial contrast. Educational overqualification is positively associated with earnings in the rural reference group, but this association is substantially offset in urban areas. One possible explanation is that upper-secondary education carries greater scarcity value in rural labor markets, while the larger supply of similarly educated workers and stronger competition in urban areas may reduce this relative advantage. This mechanism cannot be directly tested with the present data. Indonesian evidence nevertheless shows that the wage associations of vertical educational mismatch vary across spatial conditions, highlighting the role of local labor-market context in shaping the returns associated with mismatch (Hakim & Rosini, 2026).

Educational underqualification displays the reverse earnings pattern. It is negatively associated with earnings in rural areas, while the positive interaction with urban residence reverses this penalty. In rural labor markets, a narrower occupational base and more limited opportunities for specialized training may make formal qualification gaps more difficult to compensate for.

Urban labor markets provide greater opportunities for workers to acquire and demonstrate skills outside formal education, allowing experience, training, and occupation-specific competencies to substitute partly for educational credentials. This interpretation is consistent with recent evidence that the earnings consequences of educational mismatch depend on workers' underlying skills and occupational context (Kang & Hong, 2025).

Selective retention may further contribute to the better outcomes observed among urban underqualified workers. Workers who lack the formally required education but remain in higher-requirement occupations may be positively selected on experience, ability, or other productive characteristics not observed in SAKERNAS. Those who cannot compensate for their qualification gaps may be more likely to move to other jobs, whereas better-performing matches are more likely to persist (McGuinness, 2006).

The results generally support H4, showing that the associations between vertical education–job mismatch, job quality, and earnings differ between urban and rural labor markets. The observed patterns are consistent with differences in the relative value of educational credentials, occupational sorting, employer learning, and selective retention across local labor markets. Better outcomes among some mismatched workers should not be interpreted as evidence that mismatch itself is beneficial; rather, they suggest that the value and utilization of formal qualifications depend partly on the local labor-market context in which worker–job matching occurs.

4.4. Policy implications

The heterogeneity observed across mismatch categories and local labor markets suggests that a standardized “link-and-match” approach may be insufficient. First, national and subnational governments should develop regularly updated labor-market information systems that connect local occupational demand, required competencies, wage distributions, and job-quality indicators. Schools, vocational institutions, career counselors, workers, and employers should be able to use these systems so that education and training decisions respond to local demand rather than national averages alone.

Second, school-to-work transition programs should emphasize verified workplace learning. The non-significant internship coefficient indicates that participation alone is not sufficient; program quality, duration, mentoring, task relevance, and employer commitment must be monitored. Training programs should be evaluated by whether they improve matching, mobility, and productive task use—not merely by the number of certificates issued. Because

training is associated with lower overqualification but higher underqualification, competency-based certification and recognition of prior learning can help employers distinguish demonstrated skills from formal schooling.

Third, policy must address the demand side of the labor market. Expanding productive formal employment, improving contract security, extending social protection, and supporting digital adoption may raise worker welfare more directly than educational alignment alone. In rural areas, strategies should focus on diversifying employment beyond low-productivity activities and improving access to regional employment centers. In urban areas, greater emphasis should be placed on competency-based recruitment, transparent occupational standards, and the use of validated skill assessments by employers.

Finally, interventions should account for unequal sorting by gender and disability. Career guidance, accessible training, anti-discrimination enforcement, reasonable workplace accommodation, and targeted placement services can broaden the range of jobs available to workers who face structural barriers. These measures would reduce mismatch by improving both worker capabilities and access to suitable vacancies.

4.5. Limitations and future research

Several limitations qualify the interpretation of the findings. First, the cross-sectional design does not establish causal effects or reveal whether workers move into or out of mismatch over time. The coefficients should be interpreted as conditional associations. Longitudinal data are required to examine persistence, mobility, and the timing of training or job changes.

Second, the ISCO–ISCED approach is a normative education-based measure of occupational requirements. It may not capture within-occupation differences in tasks, employer-specific requirements, informal competencies, or actual skill use. Some workers classified as underqualified may possess equivalent skills acquired through experience, while some classified as well matched may still experience skill mismatch. Future research should combine education-based measures with direct skill assessments, task indicators, or employer-reported requirements.

Third, outcome availability differs across models, particularly for earnings and the JQI, and may create sample-selection concerns. The Lee correction provides a sensitivity assessment but relies on functional-form and selection assumptions and should not be treated as definitive causal identification. Unobserved ability, motivation, tenure, family networks, firm characteristics, and local labor demand may also influence both mismatch and welfare. Matched employer–employee data, panel

surveys, and credible policy-based identification strategies would strengthen future analysis.

Finally, the JQI is an unweighted index and includes an income-adequacy component. Alternative weighting schemes and specifications that exclude income adequacy should be examined to verify that job-quality results are not mechanically related to earnings. Future studies should also estimate heterogeneous associations by school track, gender, disability, occupation, sector, and career stage.

5. Conclusion

This study shows that vertical education–job mismatch is a substantial feature of the Indonesian labor market for upper-secondary graduates. Individual characteristics, training, and spatial labor-market conditions are systematically associated with whether workers are well matched, educationally overqualified, or educationally underqualified. The welfare associations are not uniform; they vary by mismatch category, outcome, and urban–rural context. Educational mismatch therefore does not necessarily correspond to poorer job quality or lower earnings, highlighting the importance of considering multiple dimensions of worker welfare.

Positive labor-market outcomes among some mismatched workers should not be interpreted as evidence of efficient worker–job allocation. Instead, they may reflect differences between formal qualifications and actual skills, as well as workplace learning, employer learning, selective retention, and other forms of adjustment to initially imperfect matches. By examining the determinants of mismatch alongside its associations with job quality and earnings, this study shows that who becomes mismatched and how mismatch is associated with worker welfare depend partly on the labor-market context in which matching occurs.

Acknowledgments

None.

Funding

None.

Conflict of interest

The authors declare they have no competing interests.

Author contributions

Conceptualization: Devanto Shasta Pratomo, David Kaluge

Formal analysis: Kuku Arisetyawan, Wildan Syafitri

Investigation: Kuku Arisetyawan

Methodology: Kuku Arisetyawan, Devanto Shasta

Pratomo

Writing-original draft: Kukuh Arisetyawan, Devanto Shasta Pratomo

Writing-review & editing: David Kaluge, Wildan Syafitri

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data

The data supporting the findings of this study are available from the corresponding author upon reasonable request, subject to the data-access terms and restrictions imposed by Statistics Indonesia (BPS). The original SAKERNAS microdata may also be obtained directly from Statistics Indonesia through its official data-access services.

References

- Allen, J., & De Weert, E. (2007). What do educational mismatches tell us about skill mismatches? A cross-country analysis. *European Journal of Education*, 42(1), 59-73.
<https://doi.org/10.1111/j.1465-3435.2007.00283.x>
- Allen, J., & van der Velden, R. (2001). Educational Mismatches versus Skill Mismatches: Effects on Wages, Job Satisfaction, and On-the-Job Search. *Oxford Economic Papers*, 53(3), 434-452.
<https://doi.org/10.1093/oep/53.3.434>
- Altonji, J. G., & Pierret, C. R. (2001). Employer learning and statistical discrimination. *Quarterly Journal of Economics*, 116(1), 313-350.
<https://doi.org/10.1162/003355301556329>
- Ananian, S., & Dellaferrera, G. (2024). A study on the employment and wage outcomes of people with disabilities. ILO Working Paper No. 124. International Labour Organization.
- Ariansyah, K., Wismayanti, Y. F., Savitri, R., Listanto, V., Aswin, A., Ahad, M. P. Y., & Cahyarini, B. R. (2024). Comparing labour market performance of vocational and general school graduates in Indonesia: insights from stable and crisis conditions. *Empirical Research in Vocational Education and Training*, 16(1).
<https://doi.org/10.1186/s40461-024-00160-6>
- Becker, G. S. (1964). *Human capital: A theoretical and empirical analysis, with special reference to education*. Columbia University Press.
- Blanck, J. M., Brzinsky-Fay, C., & Powell, J. J. W. (2025). Special NEETs: Institutional Influences on School-to-Work Transitions of Young People with Disabilities in Europe. *Work, Employment and Society*.
<https://doi.org/10.1177/09500170251361814>
- Blázquez Cuesta, M., Pérez Navarro, M. A., & Sánchez-Mangas, R. (2025). Overeducation under different macroeconomic conditions: How Spanish university graduates fit in the labor market? *Journal of Policy Modeling*, 47(2), 448-470.
<https://doi.org/10.1016/j.jpolmod.2024.10.001>
- Botezat, A., Incaltarau, C., Diac, S. A., & Grosu, A. C. (2024). Having your career path decided too early: the effects of high school track on education-occupation mismatch. *International Journal of Manpower*, 45(6), 1171-1190.
<https://doi.org/10.1108/IJM-03-2023-0123>
- Capsada-Munsech, Q. (2017). Overeducation: Concept, theories, and empirical evidence. *Sociology Compass*, 11(10), e12518.
<https://doi.org/10.1111/soc4.12518>
- Choi, S., Li, H., Ogawa, K., & Tanaka, Y. (2024). Secondary vocational education and decent work in Indonesia: differences between urban and rural areas. *International Journal of Training Research*, 22(2), 125-146.
<https://doi.org/10.1080/14480220.2023.2222939>
- Cortadas-Guasch, P. (2024). Analysis of the factors determining educational mismatches: Evidence from the Catalanian employment insertion surveys. *International Journal of Manpower*, 45(10), 77-95.
<https://doi.org/10.1108/IJM-09-2023-0540>
- Di Stasio, V. (2017). Who Is Ahead in the Labour Queue? Institutions' and Employers' Perspective on Overeducation, Undereducation, and Horizontal Mismatches. *Sociology of Education*, 90(2), 109-126.
<https://doi.org/10.1177/0038040717694877>
- Farber, H. S., & Gibbons, R. (1996). Learning and wage dynamics. *Quarterly Journal of Economics*, 111(4), 1007-1047.
<https://doi.org/10.2307/2946706>
- Green, F., & McIntosh, S. (2007). Is there a genuine under-utilization of skills amongst the over-qualified? *Applied Economics*, 39(4), 427-439.
<https://doi.org/10.1080/00036840500427700>
- Green, F., & Zhu, Y. (2010). Overqualification, job dissatisfaction, and increasing dispersion in the returns to graduate education. *Oxford Economic Papers*, 62(4), 740-763.
<https://doi.org/10.1093/oep/gpq002>
- Hakim, D. R., & Rosini, I. (2026). The Effect of Vertical Educational Mismatch on Wages: Evidence from Indonesia's Waged Sector. *Thailand and the World Economy*, 44(1), 176-215. <https://so05.tci-thaijo.org/index.php/TER/article/view/272532>
- Hartog, J. (2000). Over-education and earnings: Where are we, where should we go? *Economics of Education Review*, 19(2), 131-147.

[https://doi.org/10.1016/S0272-7757\(99\)00050-3](https://doi.org/10.1016/S0272-7757(99)00050-3)

Indrawati, S. M., & Kuncoro, A. (2021). Improving Competitiveness Through Vocational and Higher Education: Indonesia's Vision For Human Capital Development In 2019–2024. *Bulletin of Indonesian Economic Studies*, 57(1), 29–59.

<https://doi.org/10.1080/00074918.2021.1909692>

Jerez-Gómez, M. P., Bojica, A., Martínez-del-Río, J., & Karaja, R. (2023). Designing successful internships: Exploring the role of duration, formalization and motivational practices. *Education + Training*, 65(3), 433–453.

<https://doi.org/10.1108/ET-12-2021-0480>

Kang, C., & Hong, S. (2025). Educational mismatch, skills, and wage penalty: Evidence from US heterogeneous labor. *SAGE Open*, 15(4), 21582440251383620.

<https://doi.org/10.1177/21582440251383620>

Khoiruddin, M. A., Setyanti, A. M., Suman, A., Prasetyia, F., & Susilo, S. (2024). Exploring determinants of education-job mismatch among educated workers in Indonesia. *Jurnal Ekonomi Pembangunan: Kajian Masalah Ekonomi dan Pembangunan*, 25(2), 263–281.

<https://doi.org/10.23917/jep.v25i2.23994>

Kim, S.-J., & Choi, S. O. (2018). The effects of job mismatch on pay, job satisfaction, and performance. *Journal of Open Innovation: Technology, Market, and Complexity*, 4(4).

<https://doi.org/10.3390/joitmc4040049>

Lu, Y., Li, X., & Ciocca Eller, C. (2026). Underemployed and Penalized: Education–Occupation Mismatch and Racial/Ethnic Inequality among Highly Educated Workers. *Sociology of Education*, 99(1), 59–81.

<https://doi.org/10.1177/00380407251356275>

McGowan, M. A., & Andrews, D. (2017). labour market mismatch and labour productivity: Evidence from piae data. *Research in labor Economics*, 45, 199–241.

<https://doi.org/10.1108/S0147-912120170000045006>

McGuinness, S. (2006). Overeducation in the labour market. *Journal of Economic Surveys*, 20(3), 387–418.

<https://doi.org/10.1111/j.0950-0804.2006.00284.x>

McGuinness, S., Pouliakas, K., & Redmond, P. (2018). Skills mismatch: Concepts, measurement and policy approaches. *Journal of Economic Surveys*, 32(4), 985–1015.

<https://doi.org/10.1111/joes.12254>

Mimani, P. (2025). *Essays on the role of information frictions in labour market dynamics: A search and matching model perspective* (Doctoral dissertation, European University Institute). European University Institute Research Repository. <https://cadmus.eui.eu/bitstreams/f34854f7-7afd-4306-9d94-10a5561fdc13/download>

Mortensen, D. T., & Pissarides, C. A. (1999). New developments in models of search in the labor market. *Handbook of Labor Economics*, 3, 2567–2627.

[https://doi.org/10.1016/S1573-4463\(99\)30025-0](https://doi.org/10.1016/S1573-4463(99)30025-0)

Putranto, F. G. F., Pratomo, D. S., & Pudjihardjo, M. (2024). Where Did Young Workers Go? The Increase of Youth NEET in Indonesia During Pandemic. *Jurnal Ekonomi Kependudukan dan Keluarga*, 1(2), 143–155.

<https://doi.org/10.7454/jekk.v1i2.1039>

Rocha, A. B., Figueiredo, H., Sá, C., & Portela, M. (2025). Mismatch matters: education and productivity in laggard and frontier firms. *Journal of Productivity Analysis*, 64(2), 189–208.

<https://doi.org/10.1007/s11123-025-00772-4>

Santoso, R. P., Sahadewo, G. A., Sugiyanto, C., & Setiastuti, S. U. (2022). Informal and Formal Wage Differences Based on Cohorts in Indonesia. *Economics*, 10(12), 317.

<https://doi.org/10.3390/economics10120317>

Sattinger, M. (2012). Assignment models and quantitative mismatches. In Prepared for the Expert Workshop” Skill Mismatch and Firm Dynamics: Integrating Skills with the World of Work,” April (Vol. 27). https://melbourneinstitute.unimelb.edu.au/assets/documents/hilda-bibliography/conference-papers-lectures/2012/Sattinger_Assignment_Models_and_Quantitative_Mismatches.pdf

Spence, M. (1973). Job market signaling. *Quarterly Journal of Economics*, 87(3), 355–374.

<https://doi.org/10.2307/1882010>

Tobback, I., Verhaest, D., & Baert, S. (2024). Internships, hiring outcomes and underlying mechanisms: A stated preferences experiment. *De Economist*, 172(1), 25–48.

<https://doi.org/10.1007/s10645-023-09432-0>

Wardani, L. M. I., Syaputra, R. D., Malek, M. D. A., Risnawati, E., Kartikawati, M. S., Sekarini, D. A., Dawanti, R., Putri, P. K., Aprianti, M., & Irmawan, D. (2024). I Am Mismatched, I Am Doing Well: Horizontal Job-Education Mismatches and Workers’ Well-Being. *Dirasat: Human and Social Sciences*, 51(3), 128–142.

<https://doi.org/10.35516/hum.v5i3.408>

Xiao, W., Wei, Y. D., & Chen, W. (2023). Skills mismatch, jobs-housing relationship and urban commuting. *Travel Behaviour and Society*, 33, 100610.

<https://doi.org/10.1016/j.tbs.2023.100610>

Yin, Y., & Wang, C. (2026). Overeducation, vertical skill mismatch, and employment outcomes among Chinese university graduates: The mediating role of skill utilization. *Frontiers in Psychology*, 17, 1859874.

<https://doi.org/10.3389/fpsyg.2026.1859874>

Appendix

Table A1. Hausman test of independence of irrelevant alternatives (IIA) assumption

Omitted Outcome	Chi-square	df	p-value	Interpretation
Matched	44.375	2	<0.001	Evidence against IIA
Overqualification	−0.000	1	–	Test not valid/asymptotic assumptions not met
Underqualification	28.675	49	0.991	

Table A2. Multinomial probit alternative specification

Variables	(1)	Overqualification	Underqualification
	Vertical mismatch (three categories)		
	Matched		
Age	0.000 (0.00)	0.035*** (8.07)	0.034*** (7.02)
Age squared	0.000 (0.00)	−0.001*** (−11.10)	−0.000*** (−6.58)
Male (=1)	0.000 (0.00)	0.767*** (49.61)	0.221*** (11.17)
Married (=1)	0.000 (0.00)	−0.019 (−0.93)	0.028 (1.08)
Disability (=1)	0.000 (0.00)	−0.130** (−2.63)	−0.089 (−1.54)
Training participation (=1)	0.000 (0.00)	−0.100*** (−5.97)	0.274*** (13.66)
Internship (=1)	0.000 (0.00)	0.018 (0.28)	0.101 (1.40)
Urban (=1)	0.000 (0.00)	0.184*** (10.21)	0.171*** (6.96)
Constant	0.000 (0.00)	−1.467*** (−17.38)	−2.408*** (−24.50)
Observations	141,558	141,558	141,558

Notes: Multinomial probit is estimated as an alternative specification to assess robustness to the independence of irrelevant alternatives restriction of the multinomial logit model. Coefficients are reported with *t*-statistics in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A3. Province fixed-effect coefficients (related to Table 2)

Province	(1)	(2)	(3)	(4)	(5)
	Vertical mismatch	Job quality index	Log monthly income	Log monthly income	Job quality index
Aceh	0.000 (0.00)	0.000 (0.00)	0.000 (0.00)	0.000 (0.00)	0.000 (0.00)
North Sumatera	0.190*** (3.77)	0.046*** (4.15)	0.083*** (4.51)	0.084*** (4.57)	0.047*** (4.17)
West Sumatera	0.016 (0.30)	0.038** (3.17)	0.007 (0.32)	0.005 (0.24)	0.038** (3.17)
Riau	0.357*** (6.09)	0.085*** (5.96)	0.226*** (10.77)	0.223*** (10.72)	0.085*** (5.97)
Jambi	0.123 (1.78)	0.069*** (4.62)	0.163*** (6.83)	0.163*** (6.81)	0.070*** (4.65)
South Sumatera	0.019 (0.31)	0.042*** (3.31)	0.044* (1.99)	0.045* (2.02)	0.042*** (3.33)
Bengkulu	0.040 (0.62)	0.089*** (6.18)	-0.018 (-0.75)	-0.017 (-0.69)	0.089*** (6.20)
Lampung	-0.096 (-1.51)	0.028* (2.04)	-0.041 (-1.76)	-0.043 (-1.84)	0.027* (1.99)
Bangka Belitung	0.373*** (5.59)	0.050*** (3.59)	0.269*** (11.88)	0.267*** (11.84)	0.051*** (3.61)
Riau Islands	0.701*** (7.12)	0.208*** (11.75)	0.501*** (14.63)	0.497*** (14.53)	0.208*** (11.78)
Greater Jakarta	0.342*** (5.49)	0.122*** (9.46)	0.487*** (21.38)	0.484*** (21.38)	0.123*** (9.53)
West Java	0.435*** (8.82)	0.183*** (17.40)	0.233*** (12.06)	0.233*** (12.15)	0.183*** (17.40)
Central Java	0.552*** (11.56)	0.128*** (12.78)	-0.028 (-1.57)	-0.031 (-1.73)	0.129*** (12.79)
Yogyakarta	0.259*** (3.79)	0.038* (2.52)	-0.162*** (-6.34)	-0.164*** (-6.50)	0.038* (2.54)
East Java	0.390*** (8.33)	0.162*** (17.01)	0.026 (1.42)	0.024 (1.33)	0.162*** (17.02)
Banten	0.650*** (9.55)	0.248*** (16.11)	0.432*** (17.28)	0.430*** (17.45)	0.248*** (16.15)
Bali	0.161** (2.82)	0.090*** (7.57)	0.195*** (8.73)	0.191*** (8.63)	0.090*** (7.58)
West Nusa Tenggara	0.202** (2.93)	0.193*** (17.28)	-0.162*** (-5.71)	-0.164*** (-5.77)	0.193*** (17.22)
East Nusa Tenggara	0.109 (1.81)	0.116*** (8.57)	-0.294*** (-12.40)	-0.289*** (-12.23)	0.116*** (8.55)
West Kalimantan	0.083 (1.37)	0.046*** (3.71)	0.138*** (6.30)	0.140*** (6.39)	0.047*** (3.75)
Central Kalimantan	0.211** (3.14)	0.118*** (7.49)	0.265*** (10.65)	0.268*** (10.77)	0.118*** (7.53)
South Kalimantan	0.142* (2.32)	0.103*** (7.38)	0.124*** (5.44)	0.123*** (5.45)	0.103*** (7.40)
East Kalimantan	0.265*** (4.02)	0.202*** (16.07)	0.383*** (16.11)	0.381*** (16.06)	0.202*** (16.04)
North Kalimantan	0.229* (2.36)	0.153*** (6.15)	0.215*** (5.75)	0.214*** (5.75)	0.153*** (6.17)

(Cont'd...)

Table A3. (Continued)

Province	(1)	(2)	(3)	(4)	(5)
	Vertical mismatch	Job quality index	Log monthly income	Log monthly income	Job quality index
North Sulawesi	0.465*** (7.61)	0.107*** (8.42)	0.161*** (7.34)	0.163*** (7.46)	0.108*** (8.44)
Central Sulawesi	0.155* (2.36)	0.097*** (6.35)	−0.004 (−0.17)	−0.001 (−0.04)	0.097*** (6.32)
South Sulawesi	0.139* (2.55)	0.056*** (4.79)	0.011 (0.49)	0.011 (0.50)	0.056*** (4.81)
Southeast Sulawesi	0.154* (2.46)	0.112*** (8.70)	−0.029 (−1.20)	−0.029 (−1.20)	0.112*** (8.73)
Gorontalo	0.082 (0.92)	0.161*** (8.68)	−0.091** (−3.15)	−0.088** (−3.07)	0.161*** (8.67)
West Sulawesi	−0.377*** (−4.53)	0.087*** (4.37)	−0.180*** (−4.69)	−0.178*** (−4.67)	0.087*** (4.35)
Maluku	0.132* (2.01)	0.111*** (6.20)	−0.103** (−3.28)	−0.104*** (−3.33)	0.112*** (6.22)
North Maluku	0.578*** (7.77)	0.171*** (11.13)	0.107*** (3.74)	0.104*** (3.62)	0.171*** (11.10)
Papua	0.023 (0.27)	0.147*** (8.51)	0.230*** (8.34)	0.229*** (8.32)	0.147*** (8.52)
West Papua	−0.369*** (−4.72)	0.174*** (9.44)	0.370*** (12.38)	0.373*** (12.57)	0.174*** (9.43)
Observations	141,558	65,979	126,786	126,786	65,979

Notes: Province code 11 is the reference category. Coefficients for province fixed effects are reported with *t*-statistics in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A4. Robustness check using robust standard errors

Variables	(1)	(2)	(3)	(4)
	Log monthly income	Job quality index	Log monthly income	Job quality index
Matched	0.000 (0.00)	0.000 (0.00)	0.000 (0.00)	0.000 (0.00)
Overqualification	0.008 (1.42)	0.007 (1.72)	0.071*** (7.21)	−0.014* (−1.99)
Underqualification	0.037** (3.22)	0.049*** (9.12)	−0.160*** (−6.98)	0.030** (3.18)
Age	0.029*** (17.12)	0.012*** (11.16)	0.030*** (17.47)	0.012*** (11.17)
Age squared	−0.000*** (−13.48)	−0.000*** (−10.54)	−0.000*** (−13.78)	−0.000*** (−10.56)
Male (=1)	0.383*** (58.43)	0.038*** (9.72)	0.380*** (58.37)	0.038*** (9.70)
Married (=1)	0.082*** (11.12)	0.042*** (9.26)	0.083*** (11.30)	0.042*** (9.29)
Disability (=1)	−0.101*** (−4.81)	−0.034* (−2.46)	−0.100*** (−4.77)	−0.034* (−2.46)
Training participation (=1)	0.016** (2.79)	0.068*** (18.93)	0.017** (2.90)	0.068*** (18.90)

(Cont'd...)

Table A4. (Continued)

Variables	(1)	(2)	(3)	(4)
	Log monthly income	Job quality index	Log monthly income	Job quality index
Internship (=1)	−0.051* (−2.49)	0.000 (0.04)	−0.052* (−2.56)	0.001 (0.04)
Urban (=1)	0.113*** (17.60)	0.009* (2.17)		
Formal (=1)	0.176*** (28.65)	0.296*** (73.33)	0.181*** (29.47)	0.295*** (72.09)
Hours worked per week	0.012*** (57.88)	0.000 (1.46)	0.012*** (57.55)	0.000 (1.49)
Digital technology use (=1)	0.229*** (36.08)	0.084*** (20.11)	0.230*** (36.24)	0.084*** (20.13)
Rural			0.000 (0.00)	0.000 (0.00)
Urban			0.121*** (14.10)	−0.009 (−1.48)
Matched × Rural			0.000 (0.00)	0.000 (0.00)
Matched × Urban			0.000 (0.00)	0.000 (0.00)
Overqualification × Rural			0.000 (0.00)	0.000 (0.00)
Overqualification × Urban			−0.088*** (−7.55)	0.028*** (3.48)
Underqualification × Rural			0.000 (0.00)	0.000 (0.00)
Underqualification × Urban			0.266*** (10.08)	0.026* (2.29)
Constant	12.632*** (383.83)	−0.378*** (−19.29)	12.618*** (383.16)	−0.365*** (−18.34)
Observations	126786	65979	126786	65979
Adjusted R ²	0.279	0.175	0.282	0.176

Notes: Coefficients are reported with t-statistics in parentheses. Matched workers are the reference category unless otherwise indicated. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A5. Sequential wage models

Variables	(1)	(2)	(3)
	Log monthly income	Log monthly income	Log monthly income
Matched	0.000 (0.00)	0.000 (0.00)	0.000 (0.00)
Overqualification	−0.008 (−1.29)	0.008 (1.32)	0.071*** (6.68)
Underqualification	0.012 (0.89)	0.037** (2.98)	−0.160*** (−6.69)
Age	0.031*** (16.59)	0.029*** (16.60)	0.030*** (16.95)
Age squared	−0.000*** (−15.65)	−0.000*** (−13.07)	−0.000*** (−13.36)

(Cont'd...)

Table A5. (Continued)

Variables	(1)	(2)	(3)
	Log monthly income	Log monthly income	Log monthly income
Male (=1)	0.438*** (60.19)	0.383*** (56.70)	0.380*** (56.77)
Married (=1)	0.071*** (8.57)	0.082*** (10.66)	0.083*** (10.85)
Disability (=1)	-0.147*** (-6.07)	-0.101*** (-4.50)	-0.100*** (-4.46)
Training participation (=1)	0.033*** (4.97)	0.016** (2.59)	0.017** (2.69)
Internship (=1)	-0.042 (-1.69)	-0.051* (-2.25)	-0.052* (-2.33)
Urban (=1)	0.207*** (24.74)	0.113*** (14.59)	
Formal (=1)		0.176*** (25.42)	0.181*** (26.27)
Hours worked per week		0.012*** (54.94)	0.012*** (54.77)
Digital technology use (=1)		0.229*** (31.97)	0.230*** (32.08)
Rural			0.000 (0.00)
Urban			0.121*** (12.39)
Matched × Rural			0.000 (0.00)
Matched × Urban			0.000 (0.00)
Overqualification × Rural			0.000 (0.00)
Overqualification × Urban			-0.088*** (-6.92)
Underqualification × Rural			0.000 (0.00)
Underqualification × Urban			0.266*** (9.61)
Constant	13.304*** (368.70)	12.632*** (351.53)	12.618*** (352.32)
Observations	126,786	126,786	126,786
Adjusted R ²	0.186	0.279	0.282

Notes: Coefficients are reported with t-statistics in parentheses. Matched workers are the reference category unless otherwise indicated. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A6. Sequential job quality models

Variables	(1)	(2)	(3)
	Job quality index	Job quality index	Job quality index
Matched	0.000 (0.00)	0.000 (0.00)	0.000 (0.00)
Overqualification	−0.041*** (−9.52)	0.007 (1.56)	−0.014 (−1.77)
Underqualification	0.056*** (9.92)	0.049*** (8.62)	0.030** (2.99)
Age	0.012*** (10.10)	0.012*** (10.99)	0.012*** (11.01)
Age squared	−0.000*** (−10.29)	−0.000*** (−10.39)	−0.000*** (−10.41)
Male (=1)	0.033*** (7.86)	0.038*** (9.45)	0.038*** (9.43)
Married (=1)	0.049*** (9.77)	0.042*** (8.78)	0.042*** (8.82)
Disability (=1)	−0.044** (−3.03)	−0.034* (−2.43)	−0.034* (−2.44)
Training participation (=1)	0.082*** (19.91)	0.068*** (17.24)	0.068*** (17.23)
Internship (=1)	0.004 (0.24)	0.000 (0.03)	0.001 (0.04)
Urban (=1)	0.037*** (7.06)	0.009 (1.74)	
Formal (=1)		0.296*** (62.89)	0.295*** (61.88)
Hours worked per week		0.000 (1.35)	0.000 (1.38)
Digital technology use (=1)		0.084*** (16.24)	0.084*** (16.28)
Rural			0.000 (0.00)
Urban			−0.009 (−1.30)
Matched × Rural			0.000 (0.00)
Matched × Urban			0.000 (0.00)
Overqualification × Rural			0.000 (0.00)
Overqualification × Urban			0.028** (3.10)
Underqualification × Rural			0.000 (0.00)
Underqualification × Urban			0.026* (2.17)
Constant	−0.063** (−2.97)	−0.378*** (−18.08)	−0.365*** (−17.23)
Observations	65,979	65,979	65,979
Adjusted R ²	0.095	0.175	0.176

Notes: Coefficients are reported with t-statistics in parentheses. Matched workers are the reference category unless otherwise indicated. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A7. Urban and rural subsample regressions

Variables	(1)	(2)	(3)	(4)
	Log monthly income	Log monthly income	Job quality index	Job quality index
Matched	0.000 (0.00)	0.000 (0.00)	0.000 (0.00)	0.000 (0.00)
Overqualification	−0.012 (−1.69)	0.061*** (5.55)	0.016** (3.24)	−0.020* (−2.30)
Underqualification	0.100*** (7.14)	−0.124*** (−5.34)	0.054*** (7.95)	0.037*** (3.67)
Age	0.028*** (13.10)	0.031*** (11.15)	0.013*** (10.08)	0.006** (2.88)
Age squared	−0.000*** (−10.33)	−0.000*** (−8.93)	−0.000*** (−9.97)	−0.000 (−1.67)
Male (=1)	0.342*** (43.61)	0.474*** (37.48)	0.038*** (8.07)	0.038*** (4.75)
Married (=1)	0.088*** (9.68)	0.080*** (5.82)	0.043*** (7.60)	0.045*** (4.90)
Disability (=1)	−0.107*** (−3.69)	−0.083** (−2.58)	−0.043* (−2.50)	−0.016 (−0.81)
Training participation (=1)	0.016* (2.15)	0.020 (1.72)	0.065*** (14.06)	0.073*** (10.00)
Internship (=1)	−0.090*** (−3.57)	0.016 (0.39)	−0.019 (−1.37)	0.041 (1.18)
Formal (=1)	0.186*** (22.73)	0.162*** (12.99)	0.316*** (53.35)	0.264*** (34.16)
Hours worked per week	0.011*** (42.75)	0.014*** (35.24)	−0.000 (−0.71)	0.001*** (3.95)
Digital technology use (=1)	0.240*** (26.72)	0.212*** (17.82)	0.087*** (13.54)	0.077*** (9.14)
Constant	12.750*** (280.39)	12.496*** (216.19)	−0.378*** (−14.62)	−0.299*** (−7.77)
Observations	77,446	49,340	43,344	22,635
Adjusted R^2	0.257	0.269	0.157	0.207

Notes: Columns (1) and (3) report estimates for the urban subsamples, while columns (2) and (4) report estimates for the rural subsamples. Coefficients are reported with t-statistics in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A8. First-stage multinomial logit for Lee selection correction

Variables	Matched	Overqualification	Underqualification
Age	0.000 (0.00)	0.042*** (8.40)	0.047*** (6.55)
Age squared	0.000 (0.00)	−0.001*** (−11.57)	−0.000*** (−5.80)
Male (=1)	0.000 (0.00)	0.917*** (48.86)	0.208*** (7.19)
Married (=1)	0.000 (0.00)	−0.021 (−0.88)	0.045 (1.21)
Disability (=1)	0.000 (0.00)	−0.156** (−2.60)	−0.123 (−1.47)

(Cont'd...)

Table A8. (Continued)

Variables	Matched	Overqualification	Underqualification
Training participation (=1)	0.000 (0.00)	−0.124*** (−6.17)	0.424*** (14.65)
Internship (=1)	0.000 (0.00)	0.021 (0.27)	0.156 (1.48)
Urban (=1)	0.000 (0.00)	0.216*** (10.06)	0.221*** (6.00)
Constant	0.000 (0.00)	−1.762*** (−17.71)	−3.250*** (−22.41)
Observations	141,558	141,558	141,558
Pseudo R^2	0.037	0.037	0.037

Notes: Entries are coefficients with t-statistics in parentheses. Matched is the reference category in the multinomial logit specification. Statistical significance is denoted as * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A9. Lee selection-corrected models for income and job quality

Variables	(1)	(2)	(3)	(4)
	Log monthly income	Job Quality Index	Log monthly income	Job Quality Index
Mismatch: overqualification	−0.968*** (−11.42)	−0.109** (−2.69)	−0.770*** (−8.70)	−0.101* (−2.32)
Mismatch: underqualification	0.185 (1.47)	−0.025 (−0.45)	−0.226 (−1.74)	−0.028 (−0.51)
Age	0.036*** (18.61)	0.013*** (11.37)	0.036*** (18.59)	0.012*** (11.09)
Age squared	−0.000*** (−15.92)	−0.000*** (−10.77)	−0.000*** (−15.58)	−0.000*** (−10.39)
Male (=1)	0.547*** (33.71)	0.057*** (7.51)	0.528*** (31.22)	0.052*** (6.42)
Married (=1)	0.077*** (9.96)	0.041*** (8.68)	0.079*** (10.30)	0.042*** (8.71)
Disability (=1)	−0.124*** (−5.53)	−0.037** (−2.61)	−0.122*** (−5.45)	−0.036* (−2.56)
Training participation (=1)	−0.007 (−1.12)	0.066*** (16.08)	0.001 (0.08)	0.066*** (16.06)
Internship (=1)	−0.051* (−2.20)	0.001 (0.04)	−0.050* (−2.22)	0.001 (0.04)
Urban (=1)	0.151*** (17.62)	0.013* (2.45)		
Formal (=1)	0.179*** (25.83)	0.296*** (62.89)	0.182*** (26.44)	0.294*** (61.76)
Hours worked per week	0.012*** (54.54)	0.000 (1.08)	0.012*** (54.09)	0.000 (1.22)
Digital technology use (=1)	0.231*** (32.30)	0.085*** (16.37)	0.231*** (32.17)	0.085*** (16.42)
Lee correction: matched	−0.661*** (−12.04)	−0.043 (−1.59)	−0.630*** (−10.70)	−0.014 (−0.49)
Lee correction: overqualification	0.450*** (9.74)	0.086*** (3.61)	0.362*** (7.57)	0.078** (3.11)
Lee correction: underqualification	−0.400*** (−5.99)	0.023 (0.79)	−0.219** (−3.22)	0.027 (0.94)
Urban area			0.183*** (16.07)	−0.008 (−1.00)
Overqualification × urban			−0.128*** (−9.45)	0.032*** (3.35)

(Cont'd...)

Table A9. (Continued)

Variables	(1)	(2)	(3)	(4)
	Log monthly income	Job Quality Index	Log monthly income	Job Quality Index
Underqualification × urban			0.167*** (5.55)	0.025 (1.92)
Constant	12.909*** (324.37)	−0.373*** (−15.83)	12.870*** (320.09)	−0.373*** (−15.76)
Observations	126,786	65,979	126,786	65,979
Adjusted R ²	0.281	0.176	0.284	0.176
Joint <i>p</i> -value Lee correction	< 0.001	0.002	< 0.001	0.001

Notes: Matched workers are the reference category. Columns (3) and (4) include urban interaction terms, with rural workers as the reference group. Entries are coefficients with *t*-statistics in parentheses. Statistical significance is denoted as * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A10. Lee selection-corrected income models by mismatch status

Variables	(1)	(2)	(3)
	Matched	Overqualification	Underqualification
Age	0.044*** (10.83)	−0.043*** (−5.20)	0.037*** (3.41)
Age squared	−0.000*** (−8.22)	0.001*** (6.54)	−0.000** (−2.83)
Male (=1)	0.508*** (7.68)	−1.377*** (−6.99)	0.629*** (11.88)
Married (=1)	0.071*** (6.39)	0.153*** (12.97)	0.036 (1.12)
Disability (=1)	−0.147*** (−4.63)	0.171*** (3.71)	−0.084 (−1.19)
Training participation (=1)	−0.002 (−0.22)	0.432*** (9.76)	0.086 (0.60)
Internship (=1)	−0.028 (−0.95)	−0.040 (−1.23)	−0.093 (−1.09)
Urban (=1)	0.184*** (8.61)	−0.305*** (−7.52)	0.312*** (6.55)
Formal (=1)	0.143*** (14.65)	0.288*** (31.65)	−0.171*** (−6.50)
Hours worked per week	0.009*** (31.92)	0.013*** (37.77)	0.020*** (26.81)
Digital technology use (=1)	0.315*** (30.91)	0.163*** (17.25)	0.332*** (9.50)
Lee lambda: matched	−0.587** (−2.59)		
Lee lambda: overqualification		−4.532*** (−8.69)	
Lee lambda: underqualification			0.227 (0.33)
Constant	12.772*** (179.47)	19.623*** (24.72)	11.554*** (7.57)
Observations	65,097	47,344	14,345
Adjusted R ²	0.248	0.325	0.422

Notes: Dependent variable is log monthly income in all columns. Each column estimates a wage equation for the corresponding mismatch status. Entries are coefficients with *t*-statistics in parentheses. Statistical significance is denoted as * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.