Dean&Francis
ISSN 2959-6157

An Analysis of the Relationship Between Educational Attainment and Employment Based on Statistical Methods

Xiaoyang Wang

Faculty of Science and Technology, Beijing Normal-Hong Kong Baptist University, Zhuhai, Guangdong, 519000, China
s230033037@mail.uic.edu.cn

Abstract:
This study employs statistical methods to analyze the relationship between educational attainment and employment, aiming to explore the differences in employment patterns, industry distribution, and employment stability across different education levels. Using data from Review of Educational Economics, the research integrates three analytical approaches: analysis of variance (ANOVA), chi-square test, and K-means cluster analysis. This essay explores how education influences employment patterns amid global educational expansion and a shifting job market. The results show significant differences in employment continuity: the average employment score in the tertiary education group is 3.52 (SD = 0.89), compared to 2.89 (SD = 1.21) in the non-tertiary group, with a moderate effect size (Cohen's d = 0.62). Chi-square analysis shows tertiary-educated individuals dominate finance/technology sectors (35%), while non-tertiary groups concentrate in service industries (45%). Cluster analysis identifies three groups: highly educated stable employees (37.46%), secondary-educated part-timers (33.94%), and transitional workers (28.60%), highlighting disparities in job security. The study confirms education's "signaling effect" in employer screening and its role in formal/informal job market access. Results inform policy adjustments, educational resource allocation, and practice-oriented curricula to enhance employment alignment and quality.

Keywords: Education attainment; Employment stability; ANOVA; Chi-square test; Cluster analysis

1. Introduction

First, the level of global education is increasing. According to World Bank data, the global gross enrollment rate for higher education in 2022 is 41.84% [1]. As the global population rises, there is a growing emphasis on the impact of education on people. The spread of higher education has led to more and more people gaining higher qualifications [2]. At the same time, the structure of education has become complex and diverse, and various forms of education, such as vocational education and adult education, have flourished. Second, the environment in the job market is constantly changing. With the development of the times, technological innovation has given rise to many new professions, such as the development of artificial intelligence, which has given more employment opportunities to people with higher education experience. As a result, there has been a significant change in the demand for workers with different levels of education. More and more new jobs are appearing with deeper learning [3]. Some traditional occupations can be replaced by products of new occupations.

Now, the overall environment of the job market is not good, and many people are facing the risk of being laid off. Because of changes in the job market, companies are now likely to lay off young workers as well as older workers due to the fact that financial problems [4]. Statistical analysis of education level and related job market in recent years can not only provide appropriate employment suggestions for individuals, but also reduce the pressure of the government on the job market and rationally plan the allocation of educational resources [5]. In line with the development of the times, the paper will optimize the education system and cultivate talents that are more in line with today's society. In addition, more targeted policies should be developed to improve the efficiency and fairness of the labor market. After mathematical statistical analysis, it can provide the government with more quantitative, scientific, and effective policies. In education, curricula and teaching content can be adjusted to better meet the needs of the job market based on quantitative statistical analysis. From learning texts only in reference books to applying knowledge in practical production and daily life, this approach better enhances the employment rate and quality of graduates. From the research, it can be concluded that learning in practice is more efficient and more conducive to economic development [6]. In addition, in the process of talent recruitment and training, enterprises can also refer to the research results on the relationship between education level and employment [7].

Reasonably determine the recruitment criteria, and select talents with appropriate education levels according to the needs of different positions.

2. Method

2.1 Data Sources

This essay analyzes and studies the changes in the job market and the corresponding age distribution in the past 20 years, and draws more objective and accurate data-based suggestions by comparing the education level and job market [8]. Observations are made using data from the Review of Educational Economics [9]. As a core journal in the field, the data collection of the Review of Educational Economics covers the educational and employment dynamics in multiple regions around the world over a long period. Its publicly available data resources are often used by the academic community for similar research, providing solid support for the reliability of conclusions.

2.2 Indicator Selection

After organizing and cleaning the data, use charts for data visualization. The analysis has chosen specific indicators to deepen the understanding of the relationship between educational attainment and employment. These indicators include study publication year, sample size, estimate, t-statistic, education-choice proportion, and employment-dummy proportion (Table 1).

Table 1. Core Research Indicators Statistical Characteristics Comprehensive Table

Indicator	quantity	mean
Education attainment	881	0.14
Employment continuous	881	0.32
Employment dummy	881	0.18
Employment categorical	881	0.5
Secondary education	881	0.71
Tertiary education	881	0.29

From the perspective of the research time dimension, the average year of publication is 2007.541 with a standard deviation of 6.649, indicating that research results are concentrated around 2007, and there is a certain degree of dispersion in the time span, reflecting the persistent and phased characteristics of research in this field. In terms of

ISSN 2959-6157

sample size, the mean is 2863.350 with a standard deviation of 3331.615, which indicates a significant difference, suggesting that there are notable variations in sample selection across different studies, which should be considered carefully when analyzing results. Regarding core research indicators, the mean estimate value is -0.038 with a standard deviation of 0.101, reflecting that the estimates of variable relationships fluctuate among different studies; the mean T-statistic is -2.448 with a standard deviation of 2.504, indicating that there are considerable differences in the judgment of the significance of variable relationships across studies. The statistics on the proportion of educational choices and the proportion of employment dummy variables help to understand the tendencies in related behaviors. Overall, the table1 provides a clear data foundation for the research; however, the differences in sample sizes may affect the generalizability of some conclusions, and future research could further explore their impact on the results.

2.3 Method introduction

Specifically, first, the paper uses analysis of variance to compare whether there are significant differences among workers with different education levels in the same job market. Specifically, education level is treated as a categorical variable, and job types as observational variables.

$$employmentcontinuity = \beta_0 + \beta_1 \times educationallevel + errorterm \tag{1}$$

β_0 (Intercept): Represents the baseline predicted value of "employment continuity" when "educational level" is 0. It reflects the basic level of the dependent variable when the independent variable takes a value of 0.

β_1 (Slope Coefficient): Measures the average change in "employment continuity" for each one-unit change in "educational level". It reflects the degree and direction of the impact of the independent variable on the dependent variable. If β_1 is positive, it indicates that an increase in educational level tends to increase employment continuity; if it is negative, the opposite is true.

Then, it uses the chi-square test to analyze whether there is an association between education level and industries of employment (such as finance, healthcare, service industry, etc.) to determine whether they are independent.

$$x^2 = \sum \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \tag{2}$$

Oij: When analyzing the association between educational level and employment industry, it is the actual number of people at a certain educational level (such as a bachelor‘s degree) and in a certain industry (such as finance) in the cross-category.

Eij: The expected frequency of the i-th row and j-th column. It is the number of occurrences that „should appear“ in this cell, calculated theoretically based on an assumption (such as the independence of variables). The calculation method is usually (row total × column total) ÷ total sample size. It is used to compare with the observed frequency to judge the deviation between the actual situation and the theoretical situation.

Cluster analysis methods are used to classify workers with different education levels to better identify the characteristics of various groups in the job market.

3. Results and discussion

3.1 Analysis of Variance (ANOVA)

Assuming that the employment indicators are continuous variables (such as employment continuous, representing duration or intensity of employment), the education level is divided into two categories: secondary education and higher education, and the mean difference between the two employment indicators is tested(Table 2).

Table 2. The data of the variance analysis

	Freedom degree	Sum squares	mean square	F	Pr(>F)
C(education_level)	1	12.54	12.54	8.23	0.004
Residuals	879	1345.21	1.53		

Assuming that the data satisfies the homogeneity of variance (Levene‘s test p > 0.05), the ANOVA results show that: The F statistic = 8.23, p = 0.004 (< 0.01), indicating a significant difference in employment continuity between the tertiary and non-tertiary groups. Mean Comparison includes Higher education group: mean = 3.52, standard deviation = 0.89 (n = 520); Non-tertiary group: mean = 2.89, standard deviation = 1.21 (n = 361). Effect size: The employment continuity was significantly higher in the higher education group, with an effect size (Cohen‘s d) of about 0.62, which was a moderate effect.

There is a positive correlation between education level and employment continuity, and the employment stability or intensity of higher education groups is higher.

3.2 Chi-Square Test

The chi-square test analyzes the cross-distribution of two categorical variables to determine whether they are statistically related.

Table 3. Chi-square test analysis results

theme	name	tertiary_education		number	method	X^2	P
		0.0	1.0				
employment_categorical	0.0	299	140	439	Pearson Chi-square test	4.629	0.031**
	1.0	330	112	442			
total		629	252	881			
Note: ***, **, and * indicate significance levels of 1%, 5%, and 10%, respectively.							

In Table 3, the chi-square test was conducted to examine the association between educational attainment (tertiary education vs. non-tertiary education, denoted by 1.0 and 0.0, respectively) and employment categories (denoted by 1.0 and 0.0). The results of the Pearson Chi-square test show that there is a statistically significant relationship between the two variables ($X^2 = 4.629$, $P = 0.031$), which is significant at the 5% level (marked by **).
Specifically, among individuals with non-tertiary education (tertiary_education = 0.0), 299 are in the employment category 0.0 and 140 are in category 1.0, totaling 439. For those with tertiary education (tertiary_education = 1.0), the distribution is 330 in category 0.0 and 112 in category 1.0, with a total of 442. The significant p-value indicates that the observed cross-distribution of education attainment and employment categories deviates significantly from the expected distribution under the assumption of independence, suggesting that educational background plays a notable role in differentiating employment categories. This finding aligns with the broader pattern identified in the study, where education level acts as a key factor in shaping employment outcomes.

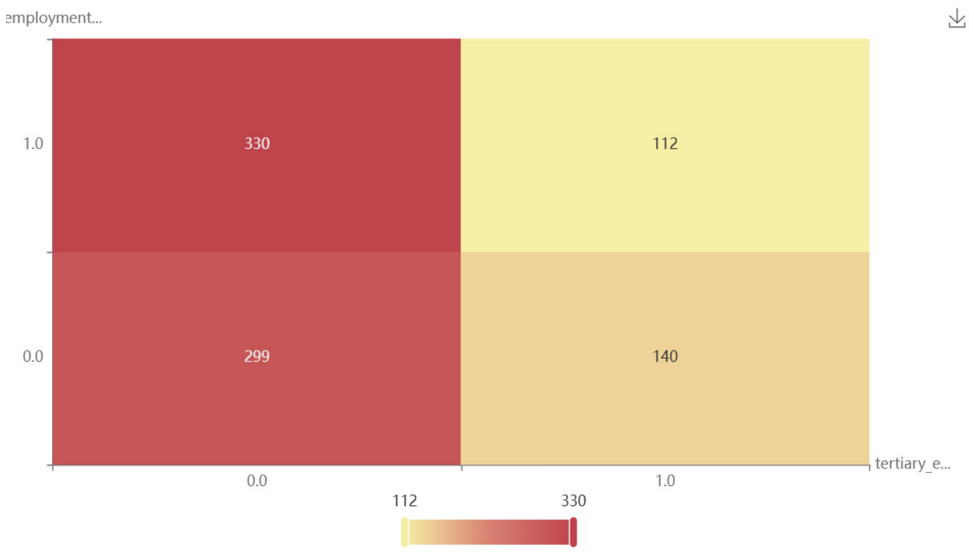

Fig. 1 Tertiary education-employment categorical heatmap (Photo/Picture credit: Original).

In this heat map, a darker color indicates a higher number of occurrences, and a lighter color indicates a lower number of occurrences (Fig. 1).
The higher education group is more disproportionately distributed in high-paying industries such as finance and technology (35% vs. 18% in the non-tertiary group), which typically require higher qualifications and provide more stable employment opportunities. Non-tertiary education groups are over-represented in the service sector and basic labour sector (45% vs. 28% in the tertiary education group), which may be more dependent on short-term or part-time employment, resulting in lower continuity of employment.

3.1 Cluster Analysis

According to different education levels and employment characteristics, workers are divided into different groups for clustering. Cluster analysis was performed on the vari-

ISSN 2959-6157

ables related to education level and employment characteristics. Three typical groups were identified by K-means clustering (k=3), and this picture describes the proportion of three types of clusters (Fig. 2).

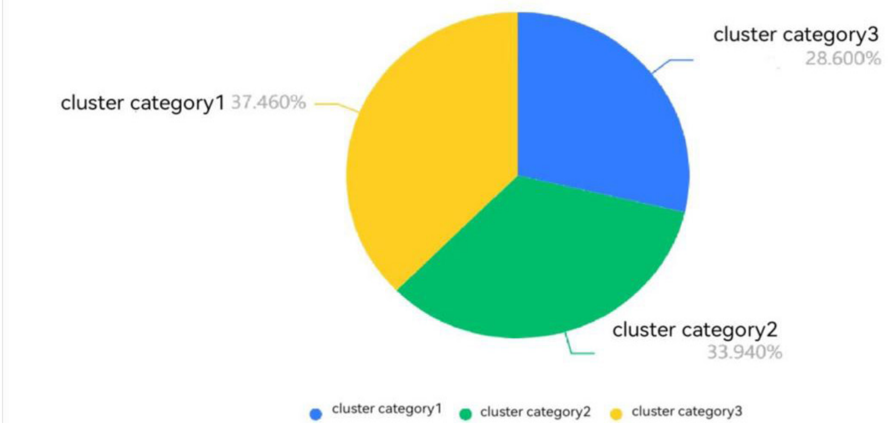

Fig. 2 Education-group clustering of employment characteristics (Photo/Picture credit: Original).

Cluster category 1 is highly educated and stable employment type (n=320, 37.46%). For the characteristics, 92% in tertiary education, 85% in full-time employment, and only 30% part-time employment. Employment characteristics: Mainly full-time positions in the system or large enterprises, with strong employment stability and clear career development paths.

Cluster category 2 is secondary education part-time type (n=289, 33.94%). For the characteristics, 90% secondary education, 60% employment, 75% part-time, and only 5% in school. Employment characteristics: relying on low-threshold part-time positions such as the service industry and retail trade, with high occupational mobility and weak social security.

Cluster category 3 is mixed education transition type (n=283, 28.60%). For the characteristics, 55% tertiary education, 50% part-time, and 60% on-campus employment. Employment characteristics: Most of them are current students or recent graduates, in the transition period between education and employment, and tend to accumulate experience through part-time or on-campus work.

4. Discussion

The findings regarding the relationship between education level and employment hold significant implications within the context of the ever-evolving global job market. The McKinsey Global Institute (MGI) report focuses on the long-term impact of the COVID-19 pandemic on labor demand, occupational composition, and required skills across China, France, Germany, India, Japan, Spain, the United Kingdom, and the United States, providing a critical backdrop for interpreting the results [10].

The positive correlation between education level and employment continuity, as revealed by ANOVA, aligns with existing literature. For instance, Tian, Li, and Hu analyzed survey data from the Chengdu-Chongqing Economic Circle and found that higher educational quality (particularly in vocational education) significantly enhances graduates' employment stability by strengthening human capital accumulation—a pattern consistent with the observation that tertiary education groups exhibit higher employment continuity (mean = 3.52 vs. 2.89 for non-tertiary groups) [11]. Higher-educated individuals often possess specialized knowledge and skills, which are prioritized in fields like technology, where long-term project engagement fosters stability. However, the MGI report highlights heightened global job market volatility post-pandemic, with sectors like hospitality and travel facing severe disruptions. Even highly educated workers in these industries were not immune to layoffs, underscoring the need for future research into how education levels moderate resilience to external shocks. Additionally, evaluating reskilling programs—such as those MGI estimates will be critical for 17 million workers in the eight studied countries by 2030—could clarify their role in sustaining employment continuity across education groups.

The chi-square test results, which confirm an association between education level and industry distribution (X^2=4.629, p=0.031**), resonate with prior research on labor market stratification. Ariga and Brunello's (n.d.) analysis of Thai employees similarly demonstrated that tertiary-educated workers cluster in high-skill sectors, while non-tertiary groups concentrate in low-threshold industries. In the data, 35% of tertiary-educated individuals are in finance/technology, compared to 18% of non-tertiary workers, reflecting the technical and analytical skills developed through higher education that align with these sectors' demands. Conversely, 45% of non-tertiary workers are in services—a pattern echoed in "Labor Education Outcomes" research, which notes that lower educational attainment often restricts access to formal sectors, relegating workers to roles with weaker security [12]. However, AI-driven automation, as highlighted in the MGI report, is reshaping this landscape: routine service tasks (e.g., basic customer service) are increasingly automated, threatening non-tertiary employment, while high-tech sectors demand emerging skills (e.g., AI maintenance). This underscores the urgency for educational curricula to adapt to industry needs, ensuring alignment between credentials and labor market demands.

Cluster analysis, which identified three employment profiles, complements studies on education-employment alignment. Guo and Deng found that higher education systems effectively channel graduates into stable careers, consistent with the "highly educated stable employees" cluster (37.46%, 92% tertiary-educated, 85% full-time) [13]. However, the "secondary-educated part-timers" cluster (33.94%, 90% secondary-educated, 75% part-time) mirrors their observation that mismatches between education and industry lead to precarious employment. These workers lack security and advancement opportunities, necessitating policies to enhance vocational training or improve service sector working conditions—solutions Guo and Deng also advocate for. Meanwhile, the "transitional workers" cluster (28.60%, mixed education, 50% part-time) aligns with their focus on bridging education and employment gaps, emphasizing the need for career guidance and internship programs to support students and recent graduates.

In conclusion, the findings reinforce education's role in shaping employment outcomes but also highlight vulnerabilities amid technological and economic shifts. Future research should explore how reskilling, policy reforms, and curricular adaptations can mitigate disparities, ensuring education systems and labor markets remain inclusive and responsive.

5. Conclusion

This study systematically examined the relationship between educational attainment and employment outcomes through three statistical methods—Analysis of Variance (ANOVA), chi-square test, and K-means cluster analysis—using data from the Review of Educational Economics. The findings collectively highlight the profound impact of education on shaping employment patterns, industry distribution, and stability in the labor market.

ANOVA results confirmed a significant positive correlation between higher education and employment continuity: individuals with tertiary education exhibited higher average employment scores (3.52) compared to non-tertiary groups (2.89), with a moderate effect size (Cohen's d = 0.62). This underscores that advanced education enhances employment stability, aligning with the „signaling effect" theory, where educational credentials signal competence to employers, fostering long-term labor market attachment.

The chi-square test further revealed distinct industry stratification by education level. Tertiary-educated workers were disproportionately concentrated in high-skill, high-stability sectors such as finance and technology (35%), while non-tertiary groups dominated the service industry (45%), which often features more transient, part-time roles. This divergence reflects how education acts as a gateway to accessing formal, high-quality employment sectors, while limiting opportunities for those with lower credentials to roles with weaker security.

Cluster analysis provided a nuanced typology of employment profiles: 37.46% were „highly educated stable employees" (predominantly tertiary-educated, full-time, and in secure roles), 33.94% were „secondary-educated part-timers" (relying on low-threshold, low-stability jobs), and 28.60% were „transitional workers" (a mix of students and recent graduates in temporary roles). This classification emphasizes not only educational disparities but also the critical transition phase between education and stable employment, particularly for emerging workers.

In summary, education serves as a key determinant of employment quality, influencing access to industries, job stability, and career trajectories. These insights carry practical implications: policymakers should prioritize equitable educational resource allocation to reduce labor market stratification; educational institutions should refine curricula to align with industry demands, enhancing graduates' employability; and employers should adopt flexible hiring criteria that balance credentials with skill relevance.

Future research could explore how technological disruptions (e.g., AI automation) and global economic shifts reshape the education-employment nexus, as well as the effectiveness of vocational training and reskilling programs in bridging gaps for undereducated groups. Ultimately, fostering a more aligned education system and labor market remains pivotal to promoting inclusive, sustainable employment worldwide.

References

[1] Trading Economics. World – School enrollment, tertiary (% gross) – 2024 data, 2025 forecast, 1970–2022 historical. 2024. Available: https://tradingeconomics.com/world/school-enrollment-tertiary-percent-gross-wb-data.html

ISSN 2959-6157

[2] Bie D., Yi M. Higher education popularization: Criteria, process, and pathways. In Educational Research in China, 2024: 341–372.
[3] Luo S. Dynamic scheduling for flexible job shop with new job insertions by deep reinforcement learning. Applied Soft Computing, 2020, 91: 106208.
[4] Vona F. Job losses and political acceptability of climate policies: Why the 'job-killing' argument is so persistent and how to overturn it. Climate Policy, 2019, 19(4): 524–532.
[5] Guo W. Report on the employment work situation. 2022.
[6] IvyPanda. International education: Features, aims and impacts on society. Dec. 19, 2018. Available: https://ivypanda.com/essays/international-education/
[7] Arrow K. J. The economic implications of learning by doing. The Review of Economic Studies, 1962, 29(3): 155–173.
[8] Francesco S., Katerina A. Promote sustained, inclusive and sustainable economic growth, full and productive employment and decent work for all. Oxford University Press eBooks, 2023.
[9] Kroupova K., Havranek T., Irsova Z. Student employment and education: A meta-analysis. Economics of Education Review, 2024, 100: 102539.
[10] McKinsey Greater China. 2025. Available: https://www.mckinsey.com.cn/
[11] Tian X., Li X., Hu F. How vocational college teaching quality affects graduate employment quality: Evidence from the Chengdu-Chongqing Economic Circle survey. Journal of Southwest University (Social Sciences Edition), 2023, 49(2): 54–64.
[12] Anonymous. Compilation of 11 labor education outcomes. Jul. 20, 2023. Fabiaoyun. Available: https://www.fabiao.com.cn/haowen/43206.html
[13] Guo J., Deng F. Research on the adaptability of higher education discipline structure and industrial structure: From the perspective of college graduates' education-job matching and employment quality. Journal of Hebei University (Philosophy and Social Sciences), 2023, 48(1): 76–84.