

The Impact of Distance Education on Employment Structure

Zaohu Zhu

Nanjing Normal University Taizhou College, Taizhou, Jiangsu, 225300, China

ABSTRACT

This study draws on data from the China Family Panel Studies (CFPS) to quantitatively examine the impact of distance education on employment structure. A binary logistic regression model is employed to categorize occupations into high-end (e.g., managerial and professional technical positions) and medium-to-low-end types in order to analyze how participation in distance education influences individuals' occupational choices. The findings indicate that participation in distance education significantly increases the likelihood of entering high-end occupations, with heterogeneous effects across age groups. Robustness tests, including redefining the classification of high-end occupations, further validate the conclusions. This study provides empirical evidence for understanding the relationship between distance education and employment structure optimization, offering policy implications for improving employment quality.

KEYWORDS

Distance education; Employment structure; Occupational choice

1 Research Background

Amid the ongoing wave of global digitalization, distance education has become an indispensable component of modern education systems due to its advantages in breaking spatial and temporal constraints and facilitating convenient resource sharing. Its coverage continues to expand—from skills training in remote areas to academic advancement for urban professionals—resulting in a growing and increasingly diverse group of participants. Meanwhile, the labor market is undergoing profound transformation. Industrial structure adjustment is accelerating, and the demand for workers' skills is becoming increasingly diverse and dynamic. Against this backdrop, examining how distance education reshapes employment structure is not only crucial for individuals' career development but also holds significant practical importance for optimizing national human resource allocation and promoting high-quality economic development.

2 Research Significance

2.1 Theoretical Significance

Existing studies on the relationship between education and employment primarily focus on conventional face-to-face education, while discussions on the employment effects of distance education remain relatively insufficient. Drawing on CFPS data, this study provides an in-depth analysis of the mechanisms linking distance education and employment structure, enriching theoretical research in the field of educational economics and offering new empirical evidence and analytical perspectives for future academic inquiry.

2.2 Practical Significance

For individuals, clarifying the role of distance education in employment helps them make informed use of distance learning resources to enhance their competitiveness and better align with labor market demands in career planning. From the perspective of policymaking, the findings provide reference for governments to optimize the development of distance education and improve employment guidance policies, thereby promoting a more precise match between educational supply and labor market demand, fostering a healthier employment environment, and supporting economic restructuring and upgrading.

3 Literature Review

3.1 Tensions between Practical Development and the Institutional Environment

As distance education in China has expanded rapidly, a growing mismatch between institutional arrangements and practical needs has become increasingly evident. Existing studies indicate that the distance education sector lacks essential supporting institutions and that relevant policies often suffer from limited operational feasibility ^[1]. As a result, local governments face challenges such as ambiguous implementation standards and unequal resource allocation, which indirectly reduce the efficiency with which educational outcomes are translated into labor market gains. Meanwhile, the technological application of distance education exhibits distinct characteristics. Compared with traditional educational models, distance education institutions in China are more inclined to leverage technological advantages by intensifying the use of new media and emerging technologies, thereby integrating the strengths of conventional teaching with those of online learning ^[2]. This technology-oriented development path is reflected not only in innovations such as live-streamed instruction and virtual simulation-based training, but also in the widespread adoption of intelligent services,

including learning data tracking and personalized recommendations. These advances facilitate flexible skill acquisition and are particularly well aligned with the demand for continuous skill upgrading in emerging labor market contexts.

3.2 Alignment between Learner Behavior and Educational Modalities

With the advancement of internet and information technologies, an increasing number of individuals have shown a strong preference for online learning ^[3]. According to the Statistical Report on Internet Development in China, the number of online education users reached 430 million in 2023, with more than 60% engaged in online vocational and skills-oriented learning. This trend is closely associated with the inherent characteristics of distance education. Owing to its ability to transcend temporal and spatial constraints, open and distance education offers learners greater flexibility and adaptability ^[4]. It is particularly suitable for employed individuals seeking to upgrade their skills during fragmented time intervals, thereby positioning distance education as a critical link between labor market supply and demand.

3.3 Bidirectional Mechanisms of Influencing Factors

Policy and technology exert significant influence in the context of distance education ^[5]. From a policy perspective, the level of fiscal support provided by local governments and the degree to which school–enterprise cooperation policies are developed directly affect the quality of educational resource provision, such as the practical relevance of vocational skill courses. From a technological perspective, the application of technologies such as 5G and artificial intelligence not only enhances the learning experience but also enables the dynamic alignment of curricular content with industry demand through the construction of “education–industry” data platforms. However, the synergistic effects between technology and policy have not yet been fully realized. In some regions, excessive emphasis on technological investment combined with insufficient institutional coordination has weakened the employment-enhancing potential of distance education.

3.4 The Dissemination Gap between Theory and Practice

The field of distance education in China is characterized by a notable imbalance in which theoretical research outcomes are well developed, while practical achievements have not been widely disseminated or shared ^[6]. On the one hand, scholars have generated a substantial body of theoretical work on teaching models and learning outcome assessment in distance education, with frameworks such as blended learning theory and connectivist learning theory reaching a relatively mature stage. On the other hand, high-quality practical cases developed by local institutions and enterprises often lack access to national-level dissemination platforms, making it difficult for such experiences to be scaled up and replicated. This disconnect between theory and practice leads to an underestimation of the actual impact of distance education on employment structure and constrains its potential role in optimizing the allocation of human resources.

3.5 Summary

Existing studies generally affirm the technological advantages and employment-enhancing potential of distance education, while also highlighting its shortcomings in institutional and practical dimensions. International research tends to focus on micro-level issues such as technology acceptance and individual skill enhancement, whereas domestic research emphasizes macro-level challenges such as insufficient policy support and disconnection between theory and practice. The imperfect institutional environment and mismatches between technological application and occupational demands constitute key constraints on the effects of distance education on employment structure, providing an important analytical entry point for this study, which integrates China’s institutional context with CFPS data.

4 Data Source and Variable Description

4.1 Data Source

This study utilizes data from the China Family Panel Studies (CFPS), conducted in 2022 by the Institute of Social Science Survey at Peking University. CFPS is a national longitudinal survey project, initiated in 2010, which continuously tracks representative samples across 25 provinces, municipalities, and autonomous regions in China, collecting dynamic data at the individual, household, and community levels. The survey covers major domains including economic conditions, educational activities, demographic changes, physical and mental health, and social attitudes and behaviors. For analytical convenience, CFPS divides the original data into family relationship, household, adult, and child databases. This study is based on the adult database. The analysis sample is limited to adults whose highest educational attainment is a bachelor’s degree. After excluding missing and abnormal observations, the final sample consists of 1,649 individuals.

4.2 Variable Description

4.2.1 Dependent Variable

Employment structure is measured by occupational type, where occupations are dichotomized into high-end and medium-to-low-end categories. High-end occupations are defined as managerial and professional technical positions, representing knowledge- and skill-intensive jobs.

4.2.2 Core Independent Variable

Participation in distance education (dist_edu), a binary variable coded as 1 if the individual has participated and 0 otherwise.

4.2.3 Control Variables

The control variables include gender (female), age (age), age squared (age_sq), self-rated health (health), years of education (edu_years), and parents' education levels (father_edu and mother_edu), in order to account for individual-level heterogeneity.

5 Model Specification

To examine the impact of participation in distance education on employment structure, this study employs a binary logistic regression model. The model is specified as follows:

$$\log \left(\frac{P(\text{high_ooc} = 1)}{P(\text{high_ooc} = 0)} \right) = \beta_0 + \beta_1 \text{dist_edu} + \beta_2 X + \varepsilon$$

$P(\text{high_ooc} = 1)$ represents the probability that an individual engages in a high-end occupation; β_1 denotes participation in distance education; X is the vector of control variables; and ε is the error term.

6 Empirical Analysis

Table 1 Descriptive Statistics

	N	Mean	Variance	Std.Dev.	Min	Max
dist_edu	1649	0.5828	0.2433	0.4932	0	1
female	1649	0.5276	0.2494	0.4994	0	1
age	1649	32.0982	68.9548	8.3031	20	60
health	1642	2.7253	0.7624	0.8732	1	5
edu_years	1649	16.000	0	0	16	16

6.1 Participation in Distance Education (Dist_edu)

The proportion of individuals participating in distance education in the sample is 58.28% (mean = 0.5828), indicating that distance education has achieved a certain level of penetration among the study population. However, nearly 40% of individuals have not participated, suggesting that there remains considerable room for further expansion and promotion of distance education.

6.2 Gender (Female)

The mean value of the gender variable is 0.5276, indicating that females account for 52.76% of the sample, slightly exceeding the proportion of males. Overall, the gender distribution is relatively balanced and broadly consistent with general demographic characteristics.

6.3 Age (Age)

The average age of individuals in the sample is 32.09 years, with a minimum of 20 years and a maximum of 60 years. This indicates that the sample is primarily composed of young and middle-aged individuals, which is consistent with the age distribution of the core labor force in the labor market.

6.4 Health Status (Health)

The mean value of self-rated health status is 2.7253 (where 1 = healthy, 2 = fair, and 3 = unhealthy), suggesting that most individuals report health conditions at the "fair" level or above. Nevertheless, there remains scope for improvement, and health status may exert a potential influence on individuals' employment choices.

6.5 Years of Education (Edu_years)

The mean, minimum, and maximum values of years of education are all 16, indicating a highly concentrated distribution of educational attainment within the sample. This pattern is closely related to the sample selection criterion that restricts the highest educational level to a bachelor's degree, reflecting the relatively high educational foundation of the study population.

6.6 Regression Analysis

This study employs a binary logistic regression model to examine the impact of distance education on employment structure, specifically focusing on the likelihood of choosing high-skilled occupations. The occupational outcome variable is dichotomized into high-skilled occupations (high_occ = 1) and low- and medium-skilled occupations (high_occ = 0). The key explanatory variable is participation in distance education (dist_edu). Control variables include gender (female),

age (age), health status (health), and parental education levels (father_edu and mother_edu), among others. The regression results are reported below.

Table 2 Model Fit Statistics

Indicator	Result	Interpretation
N	1,642	The effective sample size is 1,642, indicating that the data coverage is adequate for empirical analysis.
Wald chi2(7)	25.73	The overall model significance statistic is used to assess whether the key explanatory variable and the control variables are jointly significant.
Prob > chi2	0.0006	The p-value of the overall model significance test is below 0.05, suggesting that the model exhibits strong overall explanatory power and that the key explanatory variable and control variables are jointly significant.
Pseudo R2	0.0116	The pseudo R ² is relatively low (close to zero), indicating that the binary logistic regression model has limited explanatory power for occupational categories and that the results should be interpreted in conjunction with marginal effects analysis.

6.7 Interpretation of Coefficients and Statistical Significance

In a binary logistic regression model, the estimated coefficients capture the effects of explanatory variables on the log odds of choosing a high-skilled occupation relative to a low- or medium-skilled occupation. The direction and magnitude of these effects should be interpreted based on both the sign and the statistical significance of the coefficients.

Table 3

Variable	Coefficient	Standard Error	z-statistic	p-value	95% Confidence Interval
dist_edu	0.2869029	0.1021877	2.81	0.005	[0.0866187, 0.4871871]
female	0.073058	0.101989	0.71	0.480	[0.128783, 0.27490]
age	0.1537453	0.0415151	3.69	0.000	[0.0721812, 0.2353095]
age_sq	0.0018424	0.0005499	3.35	0.001	[0.0029202, 0.0007645]
health	0.0273096	0.0597848	0.46	0.648	[0.144856, 0.0898564]
father_edu	0.132159	0.213497	0.62	0.537	[0.287113, 0.551431]
mother_edu	0.0801352	0.186305	0.43	0.672	[0.285016, 0.445286]
_cons	3.232287	0.9736317	3.32	0.001	[5.14057, 1.324004]

Distance education participation (dist_edu) has a significantly positive effect ($p = 0.005 < 0.05$). Participation in distance education increases the log odds of choosing a high-skilled occupation relative to a low- or medium-skilled occupation, indicating that distance education promotes the upgrading of employment structure toward higher-skilled occupations. Gender (female) does not have a statistically significant effect on occupational choice ($p = 0.480 > 0.05$), suggesting that gender differences in occupational structure are not captured by the model within the sample. Age has a significantly positive effect ($p < 0.01$). A one-year increase in age raises the log odds of choosing a high-skilled occupation relative to a low- or medium-skilled occupation by 0.1537. This suggests that older individuals are more likely to enter high-skilled occupations, possibly due to cumulative effects of work experience and career development. The squared age term (age_sq) has a significantly negative effect ($p = 0.001 < 0.05$). The negative coefficient indicates a nonlinear relationship between age and high-skilled occupational choice, implying that the positive effect of age diminishes beyond a certain point. Health status (health) does not have a statistically significant effect on occupational choice ($p = 0.648 > 0.05$), possibly because the health distribution in the sample is relatively concentrated, with most individuals reporting fair or better health.

Father's education level (father_edu) shows no statistically significant effect ($p = 0.537 > 0.05$), indicating that intergenerational effects of family background are not evident in this model.

Mother's education level (mother_edu) also has no statistically significant effect ($p = 0.672 > 0.05$), consistent with the findings for father's education.

The constant term (_cons) represents the log odds of choosing a high-skilled occupation relative to a low- or medium-skilled occupation when all explanatory variables are set to zero.

6.8 Marginal Effects Analysis

The marginal effect of distance education on the probability of choosing a high-skilled occupation is estimated using the margins, `dydx(dist_edu) atmeans` command. The results are reported below.

Distance education participation (dist_edu) has a significantly positive effect on the likelihood of choosing a high-

Table 4

Variable	dy/dx	Std.Err.	z	P> z	[95% CI]
dist_edu	0.0706091	0.0249581	2.81	0.005	[0.0211521, 0.1189862]

skilled occupation ($p = 0.005 < 0.05$). On average, participation in distance education increases the probability of entering a high-skilled occupation by 7.06 percentage points. The effect of age on high-skilled occupational choice is nonlinear. Older individuals are more likely to enter high-skilled occupations; however, this positive effect diminishes beyond a certain age threshold. Participation in distance education increases the probability of choosing a high-skilled occupation by an average of 7.01 percentage points ($p < 0.01$). Moreover, the 95% confidence interval does not include zero, indicating that the estimated effect is robust and statistically reliable, thereby validating the core hypothesis.

Conclusion: Under average sample characteristics, distance education exerts a significantly positive effect on high-skilled occupational choice. Participation in distance education increases an individual's probability of entering a high-skilled occupation by approximately seven percentage points, providing strong empirical support for the core hypothesis that distance education promotes the optimization of employment structure.

6.9 Heterogeneity Analysis

Marginal Effects Analysis: The Impact of Distance Education on the Probability of Choosing High-Skilled Occupations

Using the margins, `dydx(dist_edu) atmeans` command, we estimate the average marginal effect of distance education (`dist_edu`) on the probability that an individual chooses a high-skilled occupation, thereby providing further insights into the mechanisms through which distance education affects employment structure. The results are reported below.

Table 5

Indicator	Result
Number of obs	1,642
Model VCE	Robust
Expression	Pr(high_occ), predict()
At means	All Control Variables at Their Means

Interpretation: The effective sample size reaches 1,642, indicating adequate data coverage and providing solid support for the reliability of the empirical findings. Robust standard errors are employed to effectively control for potential heteroskedasticity, thereby improving the accuracy and robustness of the estimation results. The analysis focuses on the change in the probability of choosing a high-skilled occupation associated with a one-unit increase in `dist_edu`, allowing for a clear identification of the effect of the key explanatory variable. The marginal effects are evaluated at the sample mean values of the control variables, making the results more representative and generalizable.

6.10 Robustness Check: Redefining High-Skilled Occupations

To examine the sensitivity of the results to the definition of "high-skilled occupations," we adjust the scope of the dependent variable and redefine high-skilled occupations to assess the robustness of the core findings.

6.11 Adjustment of the Dependent Variable Definition

Original definition: `high_occ = (occupation_type == 1 or occupation_type == 2)`, focusing on managerial and professional occupations.

New definition: `high_occ_new = (occupation_type == 1 or occupation_type == 2 or occupation_type == 3)`, which additionally includes clerical and service-related occupations, thereby broadening the scope of high-skilled occupations.

The model is re-estimated using the `logit high_occ_new ...` command, and the estimation results are reported below.

Table 6

Indicator	Result
Number of obs	1,642
Wald chi2(7)	126.70
Pseudo R2	0.1224

The sample size remains identical to that of the baseline model, ensuring the comparability of the results across different definitions. The overall model significance test is highly significant ($p < 0.001$), indicating that the key explanatory variable and control variables jointly exhibit strong explanatory power for occupational choice. The pseudo- R^2 increases to 0.0116 compared with the baseline model. This improvement is attributable to the broader classification of the dependent variable, which enhances the model's explanatory power for occupational choice.

6.12 Coefficient and Statistical Significance of the Key Variable `Dist_edu`

Distance education does not exhibit a statistically significant effect on the newly defined measure of high-skilled occupations ($p = 0.718 > 0.05$), suggesting that the estimated results are sensitive to the definition of high-skilled

Table 7

Variable	Coefficient	std. err.	z	P>z	[95% CI]
<code>dist_edu</code>	0.0547688	0.1514759	0.36	0.718	[0.351656, 0.2421185]

occupations and that the effect of distance education may be subject to certain boundaries.

6.13 Heterogeneous Effects of Other Variables

Age (age) exhibits a positive and statistically significant coefficient of 0.2994529 ($p < 0.05$), indicating that older individuals are more likely to enter the newly defined category of high-skilled occupations. This pattern may be related to the higher representation of clerical and service-oriented occupations within the expanded definition, for which middle-aged workers may be better suited due to accumulated work experience.

Age squared (age_sq) is negative and highly significant, with a coefficient of -0.0075341 ($p < 0.001$), suggesting a diminishing marginal effect of age on occupational choice. In other words, while age initially promotes upward occupational mobility, this positive effect gradually weakens as age increases.

7 Limitations of the Study

This study is subject to several limitations.

Sample selection bias: The sample is restricted to individuals with an undergraduate-level education, which may limit the ability to capture the effects of distance education on individuals with other educational backgrounds. As a result, the generalizability of the findings may be constrained.

Endogeneity concerns: Participation in distance education may be subject to self-selection bias, as individuals with higher ability or stronger motivation may be more inclined to engage in online learning. Future research could address this issue by employing instrumental variable approaches (e.g., regional internet coverage) or propensity score matching (PSM).

Limited variable dimensions: This study does not incorporate macro-level variables such as industry or regional characteristics, which may prevent a comprehensive assessment of how external environments influence occupational choice. Future studies could enrich the analysis by incorporating such variables.

8 Conclusion

From an individual perspective, this study confirms that distance education is an effective means of enhancing career competitiveness, particularly by facilitating access to knowledge- and skill-intensive occupations. The findings provide valuable guidance for individuals seeking to incorporate online education into their career planning strategies.

From a policymaking perspective, targeted optimization of distance education resource allocation is essential. Policymakers should prioritize the development of online courses aligned with managerial and professional occupations, while also exploring blended learning models that combine online instruction with offline practical training to address skill gaps in labor-intensive occupations.

From the perspective of educational institutions, curricula should be designed to align with the skill requirements of different occupational categories. Strengthening the linkage between distance education, professional certification, and labor market demand can improve the effectiveness and real-world applicability of educational outcomes.

References

- [1] Zhao H, Chen L, Wang X K, et al. Development Context and Problem Analysis of Modern Distance Education Policy. *Chinese Journal of Distance Education*, 2021, (08): 12-20,76.
- [2] Xuan H Q, Liu C. Innovative Research on the Teaching Mode of Ideological and Political Theory Courses in Distance Education: A Case Study of the Construction of Ideological and Political Theory Courses at the Open University of China. *Journal of Ideological and Theoretical Education Guide*, 2020, (06): 115-119.
- [3] Li F L. Key Issues and Future Research Priorities in the Economics of Distance Education. *Modern Distance Education Research*, 2019, (01): 53-65.
- [4] Li F L, Xie K. What Prevents Distance Education Learners from Obtaining Degrees: A Perspective from the Credit Bank. *Modern Distance Education Research*, 2016, (02): 75-80.
- [5] Guo Y J, Zhu Z L, Liu S F, et al. Annual Progress Report of Distance Education Research in China (2021). *Journal of Distance Education*, 2022, 40 (05): 21-30.
- [6] Xu Y Q, Chen L. Research Hotspots and Current Status of Teaching Interaction in Domestic Distance Education: Content Analysis and Social Network Analysis Based on Journal Literature from 2012 to 2017. *Chinese Journal of Distance Education*, 2018, (09): 62-72+80.