

# Layoff Factors Analysis: Evidence from 2021 China General Social Survey

Zheng Ji\*, Chongtao Zhao

University of Toronto, Toronto, ON, M5S 1A1, Canada

\*Corresponding author: Zheng Ji, zheng.ji@mail.utoronto.ca

**Copyright:** © 2024 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

**Abstract:** This paper analyzes factors that explain company layoffs, given the individual layoff data in 2021, China. Using the Probit regression model, we find that gender inequality exists in layoffs, an employee's work experience becomes less critical in the company's layoff decisions, and how an employee's health reasons affect work affects its probability of being laid off. Since we consider a significant endogeneity issue with education, using parents' education as an instrumental variable suggests that political status cannot be a significant advantage for employees to lower the chance of being laid off. Moreover, evidence implies that policymakers encouraging the pursuit of higher educational degrees can foster stability in the labor market.

**Keywords:** Layoff; Education; Instrumental variable; Probit model

**Online publication:** August 26, 2024

## 1. Introduction

In recent years, economic recession has harmed the economic conditions of countries worldwide. Most companies also had to adjust their production structures because dual impacts from the demand and supply sides hit them <sup>[1]</sup>. Among many production-side adjustments, the adjustment of labor input, that is, layoffs, has had a substantial social impact. For companies in today's era of advanced technology, labor input is not necessarily required for some positions. During an economic recession, the opportunity cost of hiring capital input becomes smaller than hiring labor input. Therefore, layoffs are a strategy for companies to minimize costs under the new constraints. However, for employees, layoffs are bad news. Labor income is the primary source of financing household consumption.

As we know, improving vocational training and broadening employment channels are crucial to stabilizing the negative impact of layoffs. Based on this, our research will also discuss the relationship between vocational training outcomes and layoffs, especially from the education degree.

This paper investigates the factors that explain layoffs using the Probit regression model according to the 2021 CGSS data. We find statistically significant average partial effects for gender, age, health conditions, and personal political status. Then, we fix the endogeneity issue in education by using parents' political status as an

instrumental variable. The average partial effects of gender, age, health conditions, and educational degree on the probability of being laid off are significant.

Our results show that gender inequality became a severe issue in layoffs in 2021. Also, an employee's working experience becomes less important in the company's layoff decisions <sup>[2]</sup>. Finally, an employee's political status is not a significant advantage in lowering the chance of being laid off, and a higher educational degree leads to a lower chance of being laid off. Hence, our results can provide some references for Chinese policymakers on promoting employment stability and social equity. It can also provide policymakers with a perspective on whether to encourage education and reduce the injustice of layoffs in the workplace caused by gender and age disparities.

## 2. Literature review

Previous micro-econometrics research consistently links layoff and re-employment. One of the most common strategies is using the duration model to estimate the time of temporary layoff. However, only a few studies discuss how factors such as gender, education, or income may affect employees' chances of being laid off by companies. A similar research was done by Okatenko in 2010 <sup>[3]</sup>, who focused on the impact of the reason for the layoff on unemployment duration. In this article, to fix the endogeneity issue of the reason for layoff, the author uses a joint model to estimate the difference between the impact of economic reasons for layoff and personal reasons for layoff <sup>[4]</sup>. The finding shows that the length of layoff is significantly shorter for economic reasons than personal reasons, especially for those workers with less than a high school degree.

Some researchers discuss layoffs under special circumstances. Ne'eman and Maestas analyze how the pandemic has affected disability employment <sup>[5]</sup>. The article compared laid-off trends for people with and without disabilities during the pandemic, both overall and by occupational category. The authors used linear probability models to estimate percent changes in employment-to-population ratios and identify differences between disabled and non-disabled employment within specific occupational categories. Through econometrics analysis, the finding of this article shows that people with disabilities experienced being laid off by companies that were proportionately like those experienced by people without disabilities, mainly concentrated in teleworkable, essential, and non-frontline occupations.

The research above helps us identify possible factors that may lead to layoffs, such as health conditions, educational degree, and non-personal factors <sup>[6]</sup>, so we will consider these factors when we use a probit regression model to estimate how characteristics of one person explain the chance of being laid off. To the best of our knowledge, our paper is the first to use microdata to analyze whether employees are being laid off as the dependent variable and the characteristics of the employees as the independent variable <sup>[7]</sup>. We also realized that there may be endogeneity in education, which leads to the incorrect estimation of the causal relationship between whether a person is fired and their educational degree. However, the level of education may be related to other factors such as household income, regional differences, and personal attitudes <sup>[8]</sup>. Therefore, we decided to eliminate endogeneity by using parents' education as an instrumental variable. This paper is one of the few empirical studies on layoffs that use instrumental variables innovatively.

## 3. Data

### 3.1. Data source and processing

Our paper analyzes individual-level data from the 2021 0.7% Census, following the comprehensive seventh national census in 2020, which surveyed China's population of approximately 1.4 billion. The data is a subset

of the China General Social Survey (CGSS)—China’s most extensive national academic survey initiative since 2003. For this study, we concentrate on 2021, specifically targeting respondents aged 18 and above from mainland China. With a multi-stage stratified sampling technique, researchers gather data through interviews, covering vital demographics such as education, health, employment status, income, family dynamics, parental occupation, income levels, and Communist Party membership.

Among the dataset, some people have just graduated from college and never worked, and their unemployment status cannot be considered layoffs<sup>[9]</sup>. Hence, we removed people who had never had a job from our analysis. Another thing we need to notice is retirement. We exclude those who are unemployed due to reaching retirement age from this dataset (i.e. individuals born before 1960). According to China’s current employment policy, many people over 60 who are legally allowed to retire can extend their retirement until they are 65. Therefore, we treat people over 65 as retired to minimize non-sampling bias. For the remaining observations, we consider those who have previously worked but are currently unemployed and have been looking for work as a layoff status.

### 3.2. Variables and their descriptive statistics

**Table 1.** Summary statistics

| Variable      | (1)   | (2)   | (3)          |
|---------------|-------|-------|--------------|
|               | Mean  | Std   | Observations |
| If_Employment | 0.72  | 0.45  | 5,740        |
| Gender        | 0.44  | 0.50  | 5,884        |
| Age           | 44.14 | 13.66 | 5,884        |
| Ethnic        | 0.92  | 0.26  | 5,884        |
| FMSTA         | 0.15  | 0.36  | 5,884        |
| Education     | 4.58  | 2.68  | 5,872        |
| Self          | 4.26  | 1.79  | 5,757        |
| Health        | 8.23  | 2.25  | 5,868        |
| Pol_Status    | 0.10  | 0.30  | 5,884        |

Our sample offers insights into the socioeconomic characteristics of a diverse group of individuals (**Table 1**). The primary dependent variable for employment status (If\_Employment) implies that approximately 72% of participants have yet to experience layoffs. In terms of gender, coded as one means the employee is male, and coded as 0 means the employee is female. A mean of 0.44 means 44% of the employees in our sample are male. The average age for our sample is approximately 44 years old. This average age is a good representation of the group we want to study since it includes those with enough work experience who will continue to work for a long time for the rest of their lives. For ethnic groups, “1” means the “Han” ethnic, which is the majority in Chinese society, and “0” means other ethnic. In our sample, 92.36% of participants have Han ethnicity. Parental Communist Party membership, denoted by the variable FMSTA, this variable implies whether at least one of the parents of an employee has a Chinese Communist Party Membership. Coded as 0 means that none of the parents is a member of CCP, and 1 means at least one of the parents has the identity. The mean value of 0.15 in FMSTA indicates that CCP membership among parents is relatively uncommon. Personal education level (Education) is measured on a scale from 1 to 10, with the average educational degree at 4.58, indicating

a range from no formal education to postgraduate studies. [Notes: The scale encompasses various educational milestones: “1: No formal education; 2: Literacy class completion; 3: Primary school education; 4: Junior high school education; 5: Technical school education; 6: Vocational high school education; 7: Ordinary high school, or secondary school; 8: College education (adult education); 9: University undergraduate education (formal higher education); 10: Graduate education and beyond]. Self-evaluation (Self) is a self-rating scale from 0 to 10, with an average score of 4.26, suggesting a moderate self-assessment among participants. Health-related leave from work frequency (Health), another variable rated from 1 to 10, averages 8.23, indicating generally low absenteeism due to health issues. Lastly, personal political status (Pol\_Status) is binary-coded, coded with 1 representing the employee has membership in the CCP, and coded with 0 means the person is not membership in the Communist Party.

## 4. Methodology

Here, we discuss whether the assumptions of the binary outcome model can be better met and the validity of the use of instrumental variables.

### 4.1. Binary outcome model

Our research analyzes what personal characteristics lead to an employee’s layoff. Therefore, our core model will be a binary outcome model. In the model, the value of the dependent variable, If\_Employment, is 0 or 1. The outcome with zero means that a person has been laid off, and one means that someone has not. For the selection of binary outcome models, most empirical studies will use the Probit model or the Logit model. By comparing the two models, we choose the Probit model since it better fits our data from four primary considerations: normality, interpretability, robustness, and computational feasibility.

### 4.2. Instrumental variable

An instrumental variable is used in empirical studies to address the endogeneity problem. Our research considered a strong endogeneity problem with the explanatory variable of education, so we chose parents’ education as an instrumental variable<sup>[10]</sup>. Researchers formed the parents’ education (FMEDU) by the following procedures: First, the education level of an employee’s mother and father was rated separately. This rating is the same as the previous rating rules for personal education described in **Section 3.2**. After that, we calculate the weighted average between one’s father’s education rating and mother’s education rating. We name this variable: parents’ education as “FMEDU”, which represents the average educational level of parents. To ensure one’s parents’ education is a valid instrumental variable, we must consider exogeneity and relevance.

- (1) Exogeneity: To ensure exogeneity is satisfied, we need to ensure that parents’ education does not directly affect one’s chance of being laid off, and it must affect the chance of being laid off by affecting one’s education. Furthermore, exogeneity also requires us to ensure no covariance between the parents’ education and other error terms that may affect the chance of one’s layoff. We can assume here that one’s parents’ education is not correlated with other characteristics, such as age, sex, income, or political status of employees. In summary, Assumption 1 (Independence):

$$COV[Parents' education, u_i] = 0$$

- (2) Relevance: To ensure the relevance of parents’ education and personal education is satisfied, we must ensure that the correlation between parents’ education and education is statistically and economically significant. Intuitively, one’s education and parent education should have a positive correlation. As parents receive a higher level of education, they also realize the importance of education level for personal



development. Parents with higher educational degrees will also provide more significant financial and mental support for their children's higher education levels. In summary, Assumption 2 (Relevance):

$$COV[Personal\ education, Parents'\ education] > 0$$

## 5. Empirical analyses and findings

### 5.1. Average partial effect from the probit regression model

Since we are interested in estimating partial effects of how different factors affect the choice probability of whether an employee is fired by the company or not, the baseline model would be a probit regression model <sup>[10]</sup>:

$$y^* = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \dots + \theta_i x_i + u_i$$

$$y_i = \begin{cases} 0, & y^* < c \\ 1, & y^* \geq c \end{cases}$$

Where  $y_i \in \{0, 1\}$  and  $y_i = 0$  if the individual is laid off by the company, otherwise  $y_i = 1$ .  $u_i$  represents unobserved terms.  $\theta_i$  measures the partial effect of the first explanatory variable  $x_i$  on the latent dependent variable  $y^*$ :  $\frac{\partial E[y^*|x_i]}{\partial x_i} = \theta_i$ . Also, the partial effect of  $x$  on choice probability  $P(y^*=1|x_i)$  is:

$$\frac{\partial P(y^*=1|x_1)}{\partial x_1} = \frac{\partial G(\theta_0 + \theta_1 x_1 + \theta_2 x_2 + \dots + \theta_i x_i)}{\partial x_1} = g(\theta_0 + \theta_1 x_1 + \theta_2 x_2 + \dots + \theta_i x_i) * \theta_1$$

is not constant, it shares the same sign as  $\theta_j$ . We set our explanatory variables  $x_1$  to  $x_8$  as those variables we described in Table 1. We check the multicollinearity problem by using the correlation matrix between variables (Refer to **Appendix 2**). We don't find a significant multicollinearity problem between our explanatory variables.

The probit regression model has the form:

$$\widehat{If}_{Employment} = 0.94 + 0.03Ethnic + 0.47gender - 0.02Age - 0.003Education + 0.2Pol_{Status} + 0.06Health + 0.01Self - 0.02FMSTA$$

After we have the probit regression model, we summarize the average partial effects of explanatory variables on the probability of being laid off in **Table 2**.

**Table 2.** Average partial effect on the probability of being laid off (layoff status [If\_Employment] as dependent variable)

|                                   | (1)        | (2)    |
|-----------------------------------|------------|--------|
|                                   | APE        | SE     |
| Ethnic                            | 0.009      | 0.022  |
| Gender                            | 0.143 ***  | 0.012  |
| Age                               | -0.007 *** | 0.0005 |
| Education                         | -0.001     | 0.003  |
| Own Political status (Pol_Status) | 0.061 ***  | 0.021  |
| Health condition (Health)         | 0.017 ***  | 0.003  |
| Self-assessment (Self)            | 0.003      | 0.003  |
| Parents' political status (FMSTA) | -0.007     | 0.016  |

Notes: \* $P < 0.1$ , \*\* $P < 0.05$ , \*\*\* $P < 0.01$ ; Column (1) reports the average partial effects of explanatory variables on the chance of being laid off, holding other factors constant. Column (2) reports delta method standard error, which represents some robustness.

We interpret those variables with the average partial effect (APE) statistically different from zero for 2021 labor markets in China. For Gender, APE = 0.143 indicates males are, on average, 14.3 percentage points more likely than females not to be laid off by their companies, holding all other factors constant. For age, APE = -0.007 represents that as an employee's age increases by five years, the probability of still working, which means not being laid off, decreases by approximately 3.5 percentage points, controlling for other things constant. The APE = 0.061 for Political status means that, on average, employees who are members of the Chinese Communist Party are 6.1 percentage points more likely to avoid being laid off by the company than employees who are not members of the Communist Party. Finally, in the setting of this article, the extent to which employees are affected by their health problems at work is divided into five rating levels: 2, 4, 6, 8, and 10. APE = 0.017 implies that for an employee whose health-related impact on work is rated 4, compared to an employee whose health-related impact on work is rated 2, which can also be understood that for each higher rating level, on average, the probability of not being laid off by the company is 3.4 percentage points higher.

From the table, the APE of ethnicity, education, self-assessment, and parents' political status is not statistically significantly different from zero, so we cannot ensure these factors are correlated with the chance of being laid off for employees. However, we believe there is an endogeneity problem in education, which causes the average partial effect of education on the chance of being off is not statistically significant.

## 5.2. Average partial effect from the probit regression model with instrumental variable

Now, we believe there is one endogenous regressor  $x_1$ , which means now  $u_i$  and  $x_1$  are not independent, say  $u_i | x_1 \sim N(x_1, 1)$ . This will lead to incorrect estimation for the Average Partial Effect of  $x_1$ :  $\frac{\partial P(y^*=1|x_1)}{\partial x_1} = g(\theta_0 + (\theta_1 + \delta)x_1 + \theta_2x_2 + \dots + \theta_ix_i) * (\theta_1 + \delta)$ . The quantity is different from  $g(\theta_0 + \theta_1x_1 + \theta_2x_2 + \dots + \theta_ix_i) * \theta_1$

only when  $\delta=0$ .

We use an instrumental variable to solve the endogeneity issue. The first stage is to construct a regression between the endogenous regressor  $x_1$ . In our setting,  $x_1$  represents personal education, and  $z_1$  represents parents' education. We find the structural equation is:

$$\widehat{Education} = 3.27 + 0.74parents' education$$

Where the estimates are tested statistically significant at a 0.01 significance level from the regression output (Refer to **Appendix 3**). This also proves our assumption about relevance in **Section 4.2. (2)**.

The main hypothesis test is the overidentification test <sup>[11]</sup> to ensure the instrumental variable validity. Our test focuses on two variables: the father's education and the mother's education. Since the test results show the  $P$ -values are big enough (refer to **Appendix 4**), we cannot reject the null hypothesis that all instrumental variables are valid, which means both the father's education and the mother's education are uncorrelated with the error term in the structural equation. Based on this, we can use parents' education that contains both the father's education and the mother's education as a valid instrumental variable, since it satisfies the exogeneity restrictions. The overidentification test result also shows our independence assumption in **Section 4.2. (1)** is satisfied.

**Table 3.** Average partial effect of education on the probability of being laid off with IV (layoff status [If\_ Employment] as dependent variable)

|  | (1)                    | (2)                    |
|--|------------------------|------------------------|
|  | APE                    | APE with IV            |
| Ethnic                                 | 0.009<br>(0.022)       | -0.004<br>(0.023)      |
| Gender                                 | 0.143 ***<br>(0.012)   | 0.136 ***<br>(0.012)   |
| Age                                    | -0.007 ***<br>(0.0005) | -0.006 ***<br>(0.0009) |
| Education                              | -0.001<br>(0.003)      | 0.02 **<br>(0.01)      |
| Personal political status (Pol_Status) | 0.061 ***<br>(0.021)   | 0.019<br>(0.027)       |
| Health Condition (Health)              | 0.017 ***<br>(0.003)   | 0.014 ***<br>(0.003)   |
| Self-assessment (Self)                 | 0.003<br>(0.003)       | 0.004<br>(0.003)       |
| Parents' political status (FMSTA)      | -0.007<br>(0.016)      | -0.021<br>(0.018)      |

Notes: \* $P < 0.1$ , \*\* $P < 0.05$ , \*\*\* $P < 0.01$ ; Column (1) reports average partial effects of explanatory variables on the chance of being laid off, holding other factors constant, in the probit regression model. Column (2) reports the average partial effects of explanatory variables on the chance of being laid off, holding other factors constant, in the probit regression model with parents' education as an instrumental variable for education. Delta method standard errors are included in parentheses.

## 6. Discussions and conclusions

### 6.1. Conclusions

Based on the interpretations of average partial effects from the probit model, this article has the following conclusions:

First, when other factors remain unchanged, females are more likely to be laid off by the company than males. Meanwhile, this probability is statistically and economically significant regardless of whether we use parents' education as an instrumental variable for individual education. This conclusion shows that gender inequality existed in China's job market layoffs, negatively affecting the Chinese government's aim to promote social and employment equality. We believe this issue deserves the government's attention and, if necessary, through policies and guidance for companies to ensure that female groups will not suffer unfair layoffs from companies in the job market.

Secondly, the data shows that the probability of being laid off increases as an employee becomes old, controlling all other variables. However, many theories in labor economics predict that an employee's value to the company increases as the working age increases. In other words, as an employee's age increases, the employee should have a lower probability of being laid off. This implies that in the recent trend of layoffs, more work experience may no longer improve an employee's competitiveness, thus reducing the probability of being laid off. With the increasing development of technology, perhaps the irreplaceability of positions will gradually become an essential criterion for a company to measure whether an employee will be laid off.

Furthermore, the impact of employees' political identity on layoffs is uncertain. In the probit model, the effect of this factor is significant. However, after we use instrumental variables, we find that the factor of this effect becomes no longer statistically significant. This difference shows that an employee's membership in the CCP does not determine whether it will affect the probability of being laid off, which also reflects social equity. For ethnicity, the APE is not statistically different from zero. This implies that different ethnicities do not directly affect the chance of being laid off, which is a good representation of Chinese social equity in that there is no racial discrimination in the company's layoff decisions.

Finally, following the utilization of instrumental variables, education has a statistically significant impact on the probability of an employee being laid off. A higher educational degree diminishes the likelihood of an individual being laid off. This result underscores the positive influence of advancing educational attainment and vocational skills on an employee's career paths. Moreover, this study suggests that pursuing higher education will ensure a solid foundation for future career paths while mitigating the probability of future layoffs <sup>[12]</sup>. As the economy recovers from the downturn, policymakers should consider encouraging students to pursue higher educational degrees through supporting educational policies and subsidies, especially for students who find it hard to afford the tuition for higher educational degrees. This policy change or subsidy provision serves the dual purpose of fostering stability in the labor market while decreasing the adverse societal effects of company layoffs <sup>[13]</sup>.

## 6.2. Discussions

To reflect some limitations in our research, in this article, we assume that people who previously had jobs but are now unemployed and are searching for new jobs are employees who have been laid off. However, there would be a possible case that employees identified as being laid off quit their jobs voluntarily. Even though it is unreasonable for one to quit the job, because searching for a new job has an extremely high cost, we cannot eliminate this possibility. Hence, we need updated data to make our conclusions more precise <sup>[14]</sup>.

For further research, we would like to use the duration model to predict the partial effect of personal characteristics on the time of being laid off <sup>[15]</sup>. The duration model is a more formal econometric model describing layoff-related topics. However, we need help on finding the dataset that satisfies the assumption and necessary conditions of the duration model.

## Acknowledgment

Special thanks to Professor Yuanyuan Wan and TA Chenyue Liu's instructions throughout the writing of this paper.

## Disclosure statement

The authors declare no conflict of interest.

## References

- [1] Okatenko A, 2010, The Impact of the Reason for Layoff on the Subsequent Unemployment Duration. *Annals of Economics and Statistics*, 99/100: 285–315. <https://doi.org/10.2307/41219168>
- [2] Blau FD, Kahn LM, 1981, Causes and Consequences of Layoffs. *Economic Inquiry*, 19(2): 270–296.
- [3] Røed K, Nordberg M, 2003, Temporary Layoffs and the Duration of Unemployment. *Labour Economics*, 10(3):

381–398. [https://doi.org/10.1016/S0927-5371\(02\)00118-5](https://doi.org/10.1016/S0927-5371(02)00118-5)

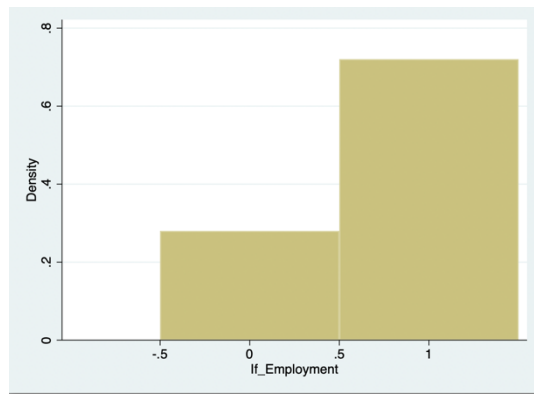
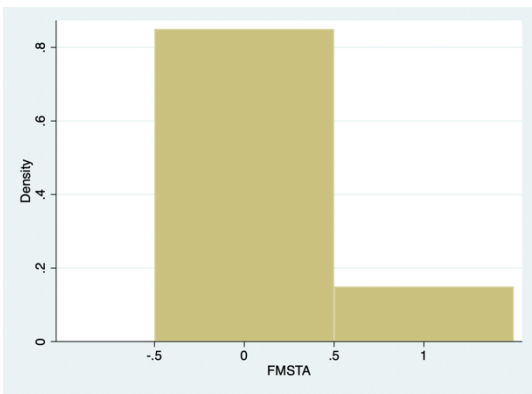
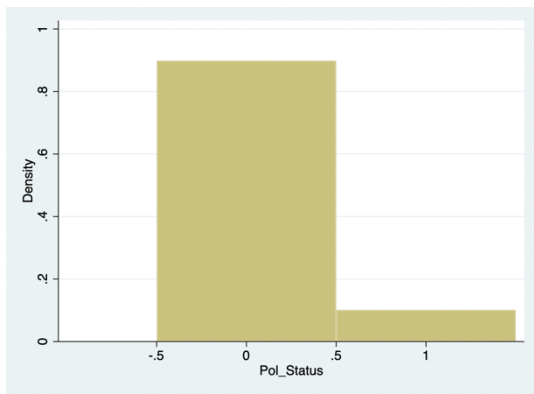
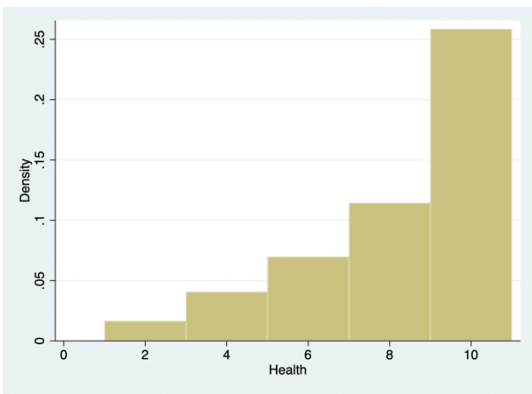
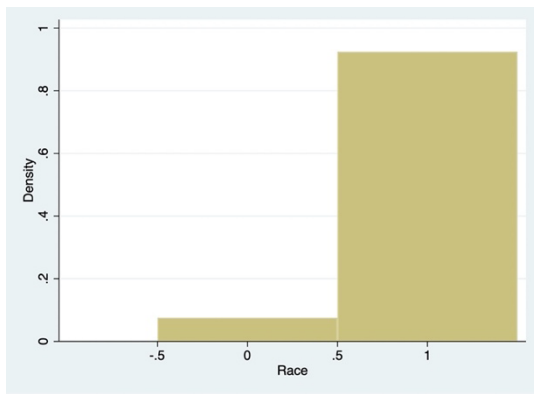
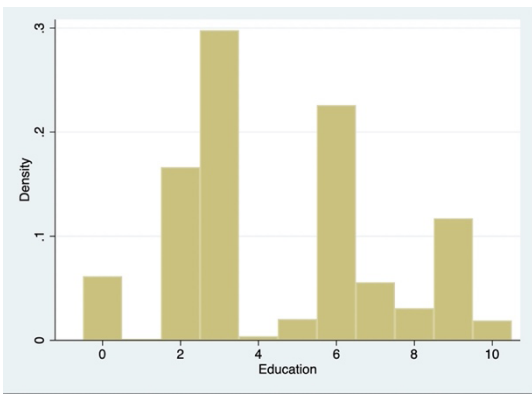
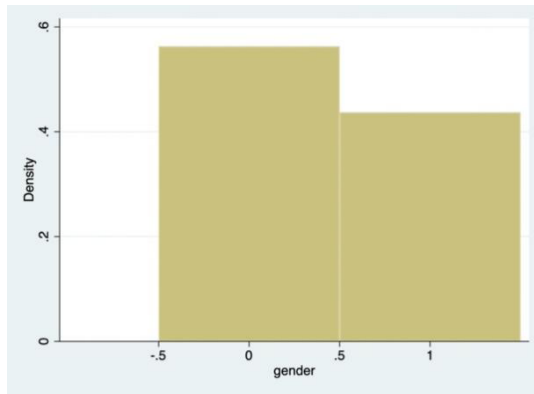
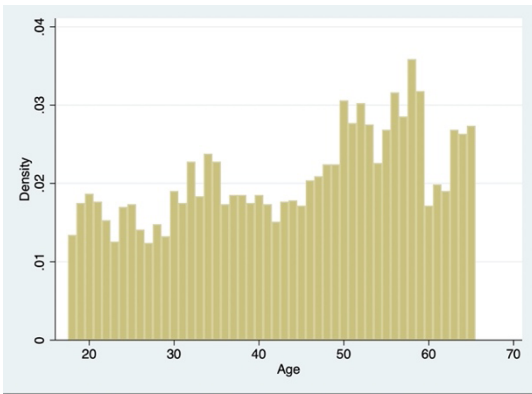
- [4] Chebolu S, 2021, Analysis of Reasons for Layoffs by Technology Startups during Covid-19 Pandemic. *Shanlax International Journal of Management*, 8(S1): 58–61. <https://doi.org/10.34293/management.v8iS1-Feb.3758>
- [5] Ne’eman A, Maestas N, 2023, How has COVID-19 Impacted Disability Employment? *Disabil Health J*, 16(2): 101429. <https://doi.org/10.1016/j.dhjo.2022.101429>
- [6] Cornfield DB, 1983, Chances of Layoff in a Corporation: A Case Study. *Administrative Science Quarterly*, 28(4): 503–520. <https://doi.org/10.2307/2393006>
- [7] Pan W, Ost B, 2014, The Impact of Parental Layoff on Higher Education Investment. *Economics of Education Review*, 42: 53–63. <https://doi.org/10.1016/j.econedurev.2014.06.006>
- [8] Couch KA, Fairlie RW, Xu H, 2020, Early Evidence of the Impacts of COVID-19 on Minority Unemployment. *J Public Econ*, 192: 104287. <https://doi.org/10.1016/j.jpubeco.2020.104287>
- [9] Goldhaber D, Strunk KO, Brown N, et al., 2016, Lessons Learned from the Great Recession: Layoffs and the RIF-Induced Teacher Shuffle. *Educational Evaluation and Policy Analysis*, 38(3): 517–548.
- [10] Joungtrakul J, 2019, Thai Higher Education in Crisis: Lay Off of University Staff on the Way. *Journal of Business Administration and Social Sciences Ramkhamhaeng University*, 2(1): 1–16.
- [11] Cameron AC, Trivedi PK, 2005, *Microeconometrics: Methods and Application*. Cambridge University Press, Cambridge. <https://doi.org/10.1017/CBO9780511811241>
- [12] Dias FA, 2021, The Racial Gap in Employment and Layoffs during COVID-19 in the United States: A Visualization. *Socius*, 7. <https://doi.org/10.1177/2378023120988397>
- [13] Sunarsi D, Jasmani, Suryani NL, et al., 2020, Covid-19 Pandemic Analysis toward Productivity Giving Layoffs Effect in the Company of Industrial Sector around South Tangerang. *Prosiding ICoISSE*, 1(1).
- [14] Liu S, Zhang X, Wan Y, 2023, Retirement Spillover Effects on Spousal Health in Urban China. *Journal of Family and Economic Issues*. <https://doi.org/10.1007/s10834-023-09935-7>
- [15] Wooldridge JM, 2019, *Introductory Econometrics: A Modern Approach*. 7th Edition. Cengage Learning, Massachusetts.

**Publisher’s note**

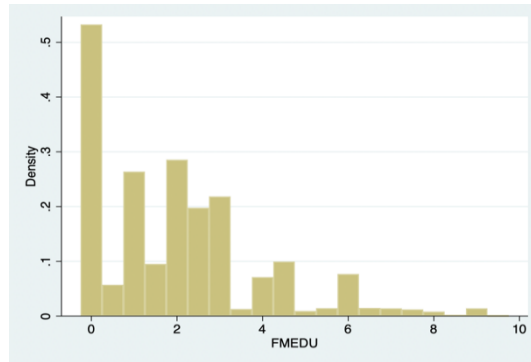
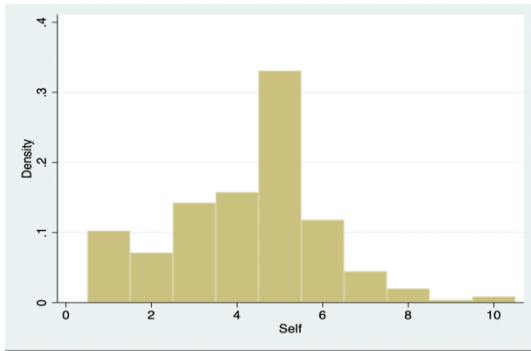
Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

# Appendix

## Appendix 1. Histogram for explanatory variables







## Appendix 2. First stage regression

```
> regress Education FMEDU
```

| Source   | SS         | df    | MS         | Number of obs | = | 5,413   |
|----------|------------|-------|------------|---------------|---|---------|
| Model    | 10380.6675 | 1     | 10380.6675 | F(1, 5411)    | = | 2624.16 |
| Residual | 21404.8264 | 5,411 | 3.95579862 | Prob > F      | = | 0.0000  |
| Total    | 31785.4938 | 5,412 | 5.87315111 | R-squared     | = | 0.3266  |
|          |            |       |            | Adj R-squared | = | 0.3265  |
|          |            |       |            | Root MSE      | = | 1.9889  |

| Education | Coefficient | Std. err. | t     | P> t  | [95% conf. interval] |
|-----------|-------------|-----------|-------|-------|----------------------|
| FMEDU     | .7413838    | .0144726  | 51.23 | 0.000 | .7130116 .769756     |
| _cons     | 3.269273    | .0509566  | 64.16 | 0.000 | 3.169377 3.369168    |

## Appendix 3. Correlation matrix

```
. correlate gender Age Race Education Pol_Status Health FMEDU
(obs=5,399)
```

|            | gender | Age     | Race   | Educate~n | Pol_St~s | Health | FMEDU  |
|------------|--------|---------|--------|-----------|----------|--------|--------|
| gender     | 1.0000 |         |        |           |          |        |        |
| Age        | 0.0108 | 1.0000  |        |           |          |        |        |
| Race       | 0.0174 | 0.0408  | 1.0000 |           |          |        |        |
| Education  | 0.1216 | -0.4982 | 0.0488 | 1.0000    |          |        |        |
| Pol_Status | 0.1384 | 0.0366  | 0.0091 | 0.2730    | 1.0000   |        |        |
| Health     | 0.0356 | -0.2648 | 0.0290 | 0.2289    | 0.0545   | 1.0000 |        |
| FMEDU      | 0.0362 | -0.5384 | 0.0190 | 0.5716    | 0.1100   | 0.1960 | 1.0000 |

## Appendix 4. Overidentifying restrictions test

```
. estat overid
```

Tests of overidentifying restrictions:

Sargan (score) chi2(1) = .351151 (p = 0.5535)  
 Basman chi2(1) = .350569 (p = 0.5538)

## Appendix 5. Wald test result

. test Education

( 1) [If\_Employment]Education = 0

chi2( 1) = 5.35  
Prob > chi2 = 0.0207