Analysis of the Influence Factors of the Ability of Graduate Employment Based on the Logistic Regression Model

Xiaowei Xu[1,a] and Weiwei Zhang[2,b]

[1]School of management, Harbin University of Science and Technology, Harbin, 150080, China. 2. China construction bank Heilongjiang branch)
([a]4994129@qq.com, [b]xxw2921@163.com)

Abstract

The problem of the grim employment of university graduate is the hot topic of today's college students and even the whole society in China. For now, the situation of university graduate's employment is becoming increasingly serious, the competition of employment is unprecedentedly intense. With the problem of the grim employment of university graduate continuous outstanding, the influence factor of the employment of university graduate has become a research hotspot to solve the problem of employment. This article used multivariate Logistic regression analysis method to analysis with the data obtained from the colleges and universities. It analyzed the factors that affect the employment of university graduate from the angle of quantitative and qualitative through the estimation and test to the model, put forward suggestions on how to promote graduates employment.

Keywords: Graduate Employment, Influence Factors, Logistic Mode, Empirical Analysis

1. Introduction

The number of college graduates is in high-speed growth with the number of hundreds of thousands each year as a result of college expansion according to national education career development statistical bulletin from 2003 to 2012. The national ordinary higher education enrollment was only about 3.822 million, college students were 11.086 million, and graduates were 1.878 million in 2003. But the data has changed significantly in 2013. The national ordinary higher education enrollment was 6.9983 million, college students were 24.6807 million, and graduates were 6.3872 million in 2013(As shown in Figure 1). New employment opportunities brought by the economic growth is not enough to satisfy the increase of school-age labor employment in our country. It is more unable to meet the employment pressure from structural adjustment. Society can provide a limited number of jobs each year, so, the more the number of graduates, the greater the employment pressure. Large increase in the number of college graduates causes the college graduates employment market supply exceeds demand. It also increases the difficulty of college graduates employment. And the employment situation is more serious.

Many scholars and research institutions have a large number of studies on graduate employment in recent years. Niall O'Higgins (2002) analyzed the internal mechanism of university students' employment from the supply and demand and matching of supply and demand. Association of Universities of Colleges of Canada thought that university should adapt to the demand, enhance the flexibility and adaptability of graduates in order to enhance the employability of graduates and adapt to the change of the new situation. Val Batcher (2002) argues that the change of Britain's career guidance service should incorporated career guidance in the university curriculum design, strengthen college students' employment ability through the career guidance. Hanzhi Zhang (2006), Chen Xiao (2006) and others had an empirical research on the employment cost of the college student which

become increasingly higher from the perspective of labor cost, and put forward the way to reduce labor costs. Xiangquan Zeng(2004) researched graduate employment from the macroscopic environment change. Yige Yang (2002), Jie Shi (2006) and others were put forward the ways to improve the employment rate through the study of the entrepreneurship education of graduates, strengthening employment guidance and improving the social ability. Chan Ma (2007) got the conclusion by analysis that the establishment of job analysis system engineering is a new idea to remit the problem of the grim employment of university graduate. Also some researchers made descriptive positive analysis from graduates, school, employer and government, *etc*. The research above is mainly theoretical argument or empirical description. They cannot obtain the influence degree of various factors on employment; also can't analyzed employment rate of graduates before graduation from the angle of the quantity. Therefore, this article analyzed the factors affecting the graduate employment by establishing a Logistic regression model in order to get valuable conclusion.
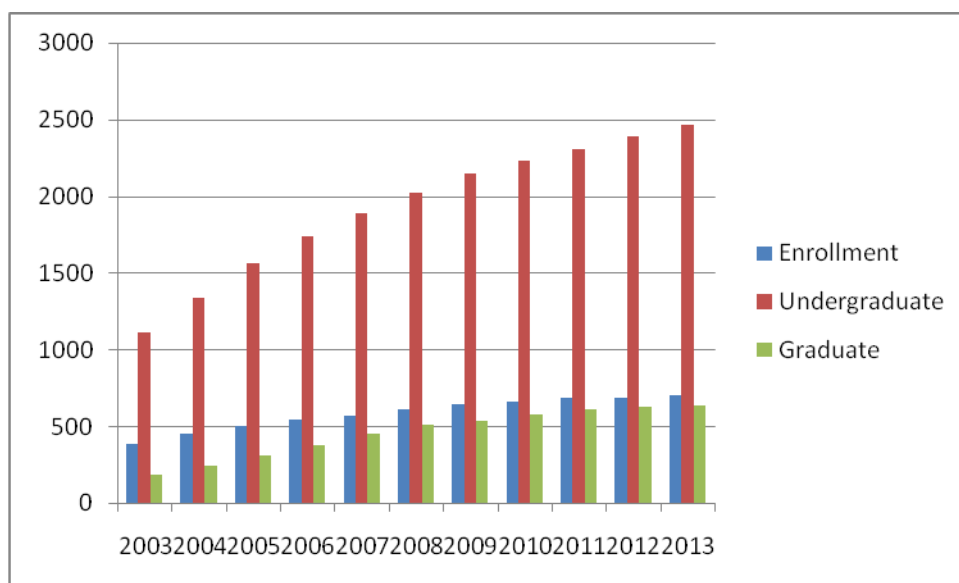

Figure 1. The Number of Enrollment, Undergraduate and Graduate of Ordinary Institutions of Higher Education from 2003 to 2013

2. Research Methods

2.1. Logistic Regression Model

Logistic regression model is one of the most commonly used statistical analysis model on the analysis of the classification of the dependent variable. Logistic regression belongs to the probability model of nonlinear regression. Assumption in the role of the independent variable x_1, $x_2 \cdots x_k$, the probability of an event is p, the ratio of the probability of an event's occurring and the probability that it will not occur is $\frac{p}{1-p}$, then the Logistic regression model is as follows:

$$Ln\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \sum_{k=1}^{k} \beta_k x_i + \varepsilon_i \qquad (1)$$

In the formula (1), β_0 is constant term, β_1, $\beta_2 \cdots \beta_k$ are known as the regression coefficient, error term ε_i is a random variable. $E(\varepsilon_i)=0$, $var(\varepsilon_i) \geq 0$

Among them, $p_i = P(y_i = 1|x_1, x_2, \ldots x$ is the probability of occurrence of

events when the value of independent variable $x_1, x_2 \ldots x_k$ are given. Once we get the samples which constituted by the observed value of independent variable of each case, and have measured value both of happened or not at the same time, we can use Logistic regression models to predict occurrence probability of the event.

When the dependent variable is discrete variable, we usually use the Logistic regression analysis to study the relationship between discrete variables and a set of explanatory variables. In this article, the dependent variable is the current employment situation, 0 and 1are two possible values (employment Y=1, non-employment Y=0). The probability of employment is P, the probability of non-employment is 1-P. This paper focus on the influence of gender, position, political affiliation, household registration, academic record, English level on employment situation of graduates. So the above factors are the independent variable of the model, and the Logistic regression model established preliminary is as follows:

$$Ln\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7 + \beta_8 X_8 + \beta_9 X_9 + \beta_{10} X_{10} + \beta_{11} X_{11} + \beta_{12} X_{12}$$

(2)

Among them, X_1=gender, X_2=position, X_3=political affiliation, X_4=household registration, X_5= academic record, X_6= English level, X_7=computer level, X_8= Hobbies, X_9=Internship experience, X_{10}=Expected employment area, X_{11}=Expected enterprise nature, X_{12}=expected salary. Because of these class variables, we need to introduce virtual variable and coding scheme.

2.2. Index selection and data sources

Employment part of the original data in this instance come from the graduate employment statistics which made by employment guidance & service center of the university in this year. And the data about the academic record and the political landscape of students come from the Students work office of the college. We carried out strict code entry and check work after obtaining the original data. Therefore, data quality has a reliable guarantee.

Table 1. The Description of Explained Variable and Statistical Data

Variable name	Variable declaration	Average value
Explained variable: Whether graduates signing enterprise	1=Yes 0=No	0.84

Table 2. The Description of Explanatory Variable and Statistical Data

Variable name	Variable declaration	Average value
gender	1=male, 0=female	0.73
position	1=student leader,0=non-student leader	0.67
political affiliation	1=communist, 0=the masses	0.76
household registration	1=town, 0=non-town	0.54
academic record	1=above 3.0GPA, 2=2.0-3.0GPA, 3=below 2.0GPA	1.98
English level	1=CET 6, 2=CET4, 3=other	1.43

computer	1=Vvisual FvoxPro, 2=other	0.78
Hobbies	1=Satisfaction, 0=non-Satisfaction	0.51
Internship experience	1=Yes, 0=No	0.54
Expected employment area	1=eastern coastal cities, 2=Large and medium-sized cities in middl, 3= Large and medium-sized cities in middl west	1.92
Expected enterprise nature	1=government agency, 2= public institution, 3=state-owned enterprise,4=Foreign or joint ventures companies, 5=their own business	2.13

2.3. Omnibus Tests

2.3.1 Test of goodness of fit: A goodness-of-fit test methods of Logistic regression model is developed by Homer Simpson (1989), it was record as the HL index. Its statistical formula is as follows:

$$HL = \sum_{g=1}^{G} \frac{y_g - n_g \hat{p}_g}{n_g \hat{p}_g (1 - \hat{p}_g)} \tag{3}$$

Among them, The G represents the number of **distinct categories**, and $G \leq 10$. y_g is the case number in group gth. $\hat{p}_g$ is observed number of gth set of events.

The result of goodness-of-fit test of the model based on HL indexes given by SPSS15.0 is shown in Table 3:

The result of above mentioned indicates that the model passed the goodness-of-fit test based on HL indexes.

Table 3. The Result of the HL Inspection

The iteration steps	Chi-square value	Degrees of freedom	The value of the P
1	4.619	8	0.807
2	3.605	8	0.985
3	3.987	8	0.958

2.3.2. Significance Test of the Model: Using backward step by step in the iteration method to select the independent variables with SPSS15.0, there are a total of three iterations under the confidence level of 0.01, he significance test of each time is shown in Table 4.

Table 4. The Result of the chi-square Test

Iteration Steps		Chi-square value	Degrees of freedom	The value of the P
1	Step 1	364..46	16	0.00
	Module 1	364..46	16	0.00

	Model 1	364..46	16	0.00
2	Step 2	-2.24	2	0.34
	Module 2	362.34	14	0.00
	Model 2	362.34	10	0.00
3	Step 3	-1.25	1	0.26
	Module 3	361.77	13	0.00
	Model 3	361.77	8	0.00

That means each iteration passed the significance test of the model. In addition, the model also passed the test of significance from the perspective of correct classification. When the model only contains constant terms, and without any arguments (hollow module), the rate of correct discrimination of model is 69.6% (Table 5).

Table 5. The Confusion Matrix of Empty Model

Iteration Steps	Observed Value		Predicted value		
			current employment situation		rate of correct discrimination
			non-employment	employment	non-employment
1	current employment situation	non-employment	0	64	0.0
		employment	0	146	100.0
	Comprehensive rate of correct discrimination				69.6

The rate of correct discrimination of model is increased to 76.8% after introducing the independent variable.

2.3.3. The Establishment of the Model: After 3 times of iterative process (the confidence level is.0.01), remove the variable which was not significant such as X_4, X_8, X_9, X_{10}, X_{11}. The result is shown as follow:

Table 6. The Results of Logistic Regression Model Analysis

Variable	Coefficient	Std. Error	z-Statistic	Prob.
X_1	2. 234 780	0. 420 360	6. 312 034	0. 000 0
X_2	4. 491 623	0. 528 762	9. 887 548	0. 000 0
X_3	0. 233 613	0. 425 320	0. 662 321	0. 001 1
X_5	-0. 543 699	0. 327 031	-2. 143 244	0. 002 1
X_6	0. 211 342	0. 312 633	0. 681 367	0. 001 2
X_9	1.867 939	0.548 687	4. 268 263	0. 000 3
X_{12}	-0. 329 458	0.268 391	-1. 231 276	0. 000 1
C	0.612 657	1.640 165	0.654 260	0. 000 2

The probability of which the variable coefficient is zero is all less than 0.01, it is

notable not zero in a = 0.05 level, that means the model pass the test. So, the model established is as follow:

$$Y = 2.23X_1 + 4.49X_2 + 0.23X_3 - 0.54X_5 + 0.21X_6 + 1.87X_9 - 0.33X_{12}$$ (4)

3. Analysis of Result

From the regression coefficient and the change of opportunity ratio, regression coefficient shows that when other factors don't change, the influence of the variable changes on the change of employment rate.

1. Gender factor

Odds ratio of male graduate employment is 1.78 times more than that of women. Graduates employment is a "buyer's market" with now widespread employment pressure, enterprises and social demand for college students may exist gender discrimination. Difficult problem of female university students' employment is still outstanding.

2. The factors of position & internship experience

Position and internship experience have more influence on employment. The coefficient of X2 and X9 are +4.49 and +1.87. This shows that have internship and student cadre experience are positive influence on employment. Graduates who get training from internship and experience of student cadre have employment advantage.

3. The factors of academic record & English level

We use grade point average of the four years as academic record of each graduate, divide it into three levels: above 3.0 means good grades, 2.0-3.0 means medium grades, below 2.0 means poor grades. Regression coefficient of variable of academic record is -0.54. This shows that the graduates whose average study lever of four years is higher than others have no significant advantage when looking for a job. It has no significant influence on graduate employment.

There is no significant difference between graduate who passed CET 4 but not passed CET 6 and graduate who not passed CET 4 in employment from the perspective of English level. In practice, those graduates who not passed CET 4 do not have high expectations to find a good job. So, they will be more positive when they look for work. Due to the impact of these factors, their employment possibilities are not smaller than those who had passed CET 4. The differences between passed CET 4 graduates and not passed CET 4 at employment may reflect on the good or bad work. But the graduates who passed CET 6 have a higher level than not passed CET 6 graduates obviously.

4. The factor of political affiliation

Political affiliation is divided into party member and non-party number. We can clearly found that political affiliation has a significant impact on graduate employment status. Employment occurrence ratio of party member of graduates is 75.69 times more than non-party member. The party member graduates have a big advantage in job searching. In general, graduates who can join the communist party are better than those non-parties, they have certain advantages in finding the job.

5. The factor of expected salary

The coefficient of expected salary(X_{12}) is -0.33. This suggests that expected salary of the graduates show the opposite trend as the increased in require. The behavior of graduates raising their expected salary will affect the employment. They should combine their own and the enterprise situation to put forward the reasonable

salary requirements. The graduates can start with a low salary, gradually get a higher salary through their own efforts in understanding enterprise's overall salary structure and the jobs basic salary.

4. Conclusion

The article established L.o git model based on the comprehensive analysis of the influence factors of graduate employment, get the analysis result o compared with expectations hypothesis n the basis of the model estimation and test as follow:

1. The gender of college students' individual characteristics has obvious effects on ratio of employment opportunities. College graduates, especially the female college graduates should treat the influence of gender difference on employment objectively. They should adjust state of mind, choose the position which can play women's gender advantages, in order to enhance their employment chances.

2. In the learning factors, English level has significant impact on employment rate and its effect is obvious. But the credit grade point is not significant influent the rate of employment. Therefore, college students should pay attention to improve their own quality rather than just grades improved.

3. A certain position for students at school, or take part in many social practice activities, to and pr communication

A certain student cadre experience or social practice activities experience play an important role to graduates to improve the employment experience and develop interpersonal skills. Colleges should strengthen vocational training in the case of widespread experience discrimination of enterprise in the present, and There should be special departments and personnel to do the job market investigation and analysis, in odd to solve the problem of employment information asymmetry between supply and demand.

References

[1] Q. Zhang and W. Zhao, "The cause, social influence and the solution of the graduates difficulty in employment", Liaoning Education Research, vol. 3, no. 75, (2012).
[2] S. Fu and Y. Xiao, "The reasons and countermeasures of the problem of the graduates difficulty in employment", China University Students Career Guide, vol. 5, no. 14, (2013).
[3] X. Shi, "The discussion of the construction of colleges and universities career guidance system", China University Students Career Guide, vol. 34, (2011).
[4] C. Li, "The analysis of the social factors that affect graduates employment", Higher Education Research, vol. 12, no. 20, (2014).
[5] H. Zhang, "The analysis of employment cost of University students", Education and occupation, vol. 4, no. 12, (2013).
[6] Z. Tian, "The empirical analysis of the factors influencing the college graduates employment", Journal of nanjing college of engineering, vol. 6, (2011).
[7] X. He and W. Liu, "Application of regression analysis", Liaoning Education Research, vol. 123, (2011).
[8] G. Jin, "Employment condition and factor analysis of University students", Economic & Trade Update, vol. 8, no. 26, (2014).