

# The Accurate Identification of College Students in Financial Hardship Based on Growth Background

An Shuting, Chen Yuxin, Dai Wanping

College of Biomedical Engineering, South-Central Minzu University, Wuhan, PR China

## Email address:

An\_Shooting@mail.scuec.edu.cn (An Shuting), chen\_xiaoer@qq.com (Chen Yuxin), d980531@163.com (Dai Wanping)

## To cite this article:

An Shuting, Chen Yuxin, Dai Wanping. The Accurate Identification of College Students in Financial Hardship Based on Growth Background. *International Journal of Statistical Distributions and Applications*. Vol. 3, No. 2, 2017, pp. 1-6. doi: 10.11648/j.ijstd.20170301.11

**Received:** March 16, 2017; **Accepted:** March 23, 2017; **Published:** April 14, 2017

---

**Abstract:** Funding college students in financial hardship is one of the most important works in the reform and development of Chinese higher education. In this paper, in order to enhance efficiency and establish uniform standards in identifying the college students in financial hardship, an index system is proposed considering the background of students. This index system is used to measure his/her family wealth index and, applying the binary Logistic regression model, the validity of the indexes is examined. The TOPSIS model is then utilized for the estimation of the family wealth index, and the data from South-Central Minzu University is used to perform the empirical analysis. The results show that the accuracy rate is 71.43%, and the sensitivity analysis indicates that the proposed model is robust.

**Keywords:** Students in Financial Hardship, Binary Logistic Regression Model, Family Wealth Index, Topsis Model

---

## 1. Introduction

Funding college students in financial hardship is one of the most important works in the reform and development of Chinese higher education. Identifying the students in financial hardship effectively is a prerequisite for funding work. The Ministry of Education of the People's Republic of China released "The Guiding Opinions on the Identification of College Students in Financial Hardship" in June 2007, which made a clearly illustration for the identification. And in December 30, 2016, the Ministry of Education issued "The Notice of the Ministry of education of the people's Republic of China on Further Enhancing and Regulating the work of Identifying College Students in Financial Hardship", which put forward further demands on the identification work in funding students.

Nowadays, universities identify students in the following levels of financial hardship: "High", "Middle", "Low" as well as "None of above" when they enrolled at September, according to "The Questionnaire of Family Financial Situation", which is provided by students themselves. After that, the level of hardship will be adjusted according to the performance of students in every term. Various methods, such as poor proof method, transverse comparison method, consumption level method, and low-income allowances

methods are used to identify the eligible students [1]. Since these methods are lack of theoretical context, accuracy and controversy are difficult to guarantee in practice. According to the data from South-Central Minzu University (SCMZU), authors in [2] found that the family income has no significant influence when determining the students who should be gained the National Grants. This phenomenon which violates the original intention of national policy showing that, as family income of single resident can be hardly obtained in China, the rationality and efficiency of identify the eligible students are greatly limited [3-4]. Therefore, many scholars turn to estimate the family wealth index according to the variables which are non-income but closely related to the income. In [5-7] the family wealth index model were established, and the the family income of some developing countries were estimated by using different weighting methods, while in [8] the data from some provinces in China were investigated indicating that, using such variables (i. e. closely related to the income), the family income can be easily estimated and can reduce reduce the cost under the premise of ensuring accuracy.

SCMZU has received numerous rewards by Ministry of Education due to the prefect performance in funding students in financial hardship. In order to ensure the authenticity of the identification, the process of "Application by student -

Evaluation in class - Re-examination by counsellor - Re-examination by college - Final judgment by the sector of student affairs” has been taken in SCMZU. Since each department will verify those materials one by one, and the different colleges identify the students independently, although we can ensure the authenticity, the large workload and different standards in different colleges are troubling the sector of student affairs in SCMZU.

In this paper, we aim to enhance the work efficiency and establish uniform standards between different colleges in identifying college students in financial hardship. The paper is organized as follows: Section 2 constructs an index system, considering the background of students, and checks its validity. Section 3 establishes TOPSIS model to estimate the family wealth index, and the empirical analysis is performed with data taken from SCMZU. Finally, the conclusions of the

results are given in Section 4.

## 2. Index System of Growth Background

To ensure the fairness and improve the efficiency, we should choose the indexes which are closely related to family income, which are also easy to collect. According to our preliminary investigation, we revise and expand the indexes discussed in [8], considering the subjective and objective factors, constructing the index system such as to include: family background, family residence, infrastructure, parents, raising pressure, personal consumer electronics, student loan (National student loan or Student-Home-Based-Loan). We shall call it as the “student background” index system. Table 1 provides the definitions and the corresponding values of this index system.

**Table 1.** The index system of growth background.

First index	Second index	The meaning of the index (value)
Family Background	Orphan	Students in high level financial hardship (directly identified).
	Minimal assurance	Students in high level financial hardship (directly identified).
	The disabled	Students in high level financial hardship (directly identified).
	None of the above	Continue the investigation.
Family Residence	Eastern Region	At and above the county level (4); below the county level (3);
	Central Region	At and above the county level (2.5); below the county level (2);
	Western Region	At and above the county level (1.5); below the county level (1).
	Hill or Plain;	Yes (0), No (1).
Infrastructure of Residence	No electricity or no road	Yes (0), No (1).
	Father's education level	Fatherless (0); Primary and below Primary(1); Junior (2); Senior (3); College or above (4).
The situation of parents	Mother's education level	Fatherless (0); Primary and below Primary(1); Junior (2); Senior (3); College or above (4).
	Father's occupation	Fatherless(0); Farmers or Unemployed (1); Staff, labor, service worker, retired (2); Civil service and self-employed (3); Manager, private entrepreneurs and government official (4)
	Mother's occupation	Motherless(0); Farmers or Unemployed (1); Staff, labor, service worker, retired (2); Civil service and self-employed (3); Manager, private entrepreneurs and government official (4)
	Father's physical status	Serious disease, disabled or fatherless(0); Chronic disease (2); In good health (4).
	Mother's physical status	Serious disease, disabled or fatherless(0); Chronic disease (2); In good health (4).
Raising pressure	The number of unemployed teenagers in the family	The non-negative value of 6 minus the number of unemployed teenagers (minimum value is 0)
The number of personal consumer electronics	The number of personal consumer electronics	Computers, smartphone worth more than ¥2,000, digital cameras.
Student loan	student loan or not	Yes (0), No (4).

We consider the binary Logistic regression model to test whether the index system can be used to measure the family wealth index or not. We have investigated a total of 5,054 students at SCMZU, in which 631 students that fall in the indexes of “orphans”, “minimal assurance” and “disabled” were identified directly as the students in high level financial hardship. In the remaining 4,423 students, 1,289 are

artificially identified as students in financial hardship.

The results I of binary Logistic regression model are shown in Tab. 2. It can be seen that the *P*-value of infrastructure and raising pressure are greater than 0.05, which means that they cannot be used to distinguish students in or not in financial hardship under the confidence level of 95%.

**Table 2.** The results I of binary logistic regression model (via SPSS).

Variables in the Equation	B	S.E.	Wald	df	Sig.	Exp (B)
Family residence	-0.154	0.04	14.942	1	0.000	0.857
Infrastructure	-0.03	0.07	0.184	1	0.668	0.97

Variables in the Equation	B	S.E.	Wald	df	Sig.	Exp (B)
Father's education level	-0.125	0.129	0.938	1	0.000	0.634
Father's occupation	-0.562	0.06	86.586	1	0.000	0.57
Father's physical status	-0.333	0.068	23.876	1	0.000	0.717
Mother's education level	-0.258	0.05	26.409	1	0.000	0.772
Mother's occupation	-0.375	0.074	25.932	1	0.000	0.687
Mother's physical status	-0.354	0.073	23.467	1	0.000	0.702
Raising pressure	-0.04	0.028	2.023	1	0.155	0.961
Personal consumer electronics	-0.501	0.102	24.145	1	0.000	0.606
Student loan	-0.423	0.023	324.308	1	0.000	0.655
Constant	6.189	0.462	179.124	1	0.000	487.225

The results II of binary Logistic regression model are shown in Tab. 3. It can be seen from Table 3 that all the P-value of indexes are less than 0.05, which means that they can be used to distinguish students in or not in financial

hardship under the confidence level of 95%. So we will use the nine indexes in Tab. 3 as the final indexes to continue the work.

**Table 3.** The results II of binary logistic regression model (via SPSS).

Variables in the Equation	B	S.E.	Wald	df	Sig.	Exp(B)
Family residence	-0.151	0.04	14.47	1	0.000	0.86
Father's education level	-0.085	0.092	0.854	1	0.000	0.537
Father's occupation	-0.559	0.06	86.297	1	0.000	0.572
Father's physical status	-0.33	0.068	23.422	1	0.000	0.719
Mother's education level	-0.257	0.05	26.136	1	0.000	0.774
Mother's occupation	-0.376	0.074	26.079	1	0.000	0.687
Mother's physical status	-0.349	0.073	22.863	1	0.000	0.706
Personal consumer electronics	-0.506	0.102	24.718	1	0.000	0.603
Student loan	-0.424	0.023	326.571	1	0.000	0.655
Constant	5.86	0.391	224.822	1	0.000	350.794

It is needed here to emphasize that, although the binary Logistic regression model can be used also to identify students in financial hardship, the returned value of the model falls into the interval  $[0, 1]$ , which can reflect the probability of being a student in financial hardship. Since the level of financial hardship are classified as “High”, “Middle”, and “Low”, the use of binary Logistic regression model by itself is not enough. To get more accurate identification results the TOPSIS model shall be considered and applied in the following section.

### 3. Topsis Model: Estimate the Family Wealth Index

#### 3.1. Topsis Model

TOPSIS is an effective multi-index evaluation model. It constructs the positive ideal solution and the negative ideal solution of the evaluation, calculates the relative closeness coefficient of each object to the ideal object, and ranks the objects according to the relative closeness coefficient. The steps of TOPSIS model are as follow [9]:

(1) Determining the normalized matrix. We use  $x_{ij}$  to represent the value of indicator  $j$  of student  $i$ , and  $X = (x_{ij})_{m \times n}$  to represent the data matrix, where  $i = 1, 2, \dots, m$  represents the students, and  $j = 1, 2, \dots, n$  represents the indexes. We use  $G = (g_{ij})_{m \times n}$  to represent the normalized matrix, and

$$y_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^m x_{ij}^2}} \quad (1)$$

(2) Constructing weighted normalized matrix. Supposing the weighted vectors of indexes are  $w = [w_1, \dots, w_n]^T$ , then we have the weighted normalized

$$c_{ij} = w_{ij} \cdot y_{ij} \quad (2)$$

(3) Determining the positive and negative ideal solution. Suppose  $c_j^*$  and  $c_j^0$  represent the index  $j$  of positive ideal solution  $C^*$  and negative ideal solution  $C^0$ , respectively, since all indexes in this paper are efficient index, i.e. the larger the index is, the richer the students may be, then we have

$$\begin{cases} c_j^+ = \max_i c_{ij}, \\ c_j^- = \min_i c_{ij}. \end{cases} \quad (3)$$

(4) Calculating the distance between each object and the positive/negative ideal solution. We use  $d_i^*$  to represent the Euclidean distance between student  $i$  and positive ideal solution, i.e.

$$d_i^+ = \sqrt{\sum_{j=1}^n (c_{ij} - c_j^+)^2} \quad (4)$$

In the same way, we use  $d_i^0$  to represent the Euclidean

distance between school  $i$  and negative ideal solution, i.e.

$$d_i^- = \sqrt{\sum_{j=1}^{10} (c_{ij} - c_j^-)^2} \quad (5)$$

(5) Calculating the relative closeness. Define the relative closeness of student  $i$  to the positive ideal solution

$$f_i = 100 \times \frac{d_i^0}{d_i^+ + d_i^0} \quad (6)$$

Obviously, the larger the  $f_i$  is, the richer the student  $i$  may be, and we call the relative closeness as family wealth index.

### 3.2. Determining the Weight

In the second step of TOPSIS model, we need to determine the weight of each index. There are various methods to determine the weight, and we use coefficient of variation method.

The basic thought of the coefficient of variation method is that, if the volatility of an index is larger, then this index can distinguish the differences between samples, so the weight of this index should be larger, and vice versa [10]. For the normalized matrix, the coefficient of variation of index  $j$  is

$$V_j = \frac{\sigma(y_j)}{\mu(y_j)} \quad (7)$$

where  $\sigma(y_j)$  and  $\mu(y_j)$  are the standard deviation and the average value of index  $j$ , respectively.

Then the weight of index  $j$  is

$$w_j = \frac{V_j}{\sum_{j=1}^n V_j} \quad (8)$$

### 3.3. Results and Basic Analysis

The data used in this paper are provided by the sector of student affairs in SCMZU. We consider the data of 5,054 students from 20 colleges, in which 631 students falling into indexes “orphans”, “minimal assurance”, or “disabled”, are identified directly as the students in high level of financial hardship.

Combining equation (1)-(9), we estimate the family wealth indexes of the remaining 4,423 students, and divide them into three categories according to the nature of their colleges, as shown in Tab. 4.

Table 4. The basic statistical results.

Category	Total number of students	Maximum	Upper Quartile	Lower Quartile	Minimum	Average Value	Standard Deviation
Natural science	1868	79.02	61.57	47.67	32.81	53.33	8.78
Social science	2205	88.65	62.13	49.71	31.45	54.18	9.11
Arts	350	74.42	62.22	50.24	34	55.41	8.65
Total	4423	88.65	60.88	48	31.45	53.99	9

From Tab. 4 we can know that:

- The family wealth indexes are in the interval of [31.45, 88.65], and have an obvious gradient, which shows that our model is suitable in identifying students in financial hardship.
- The students with minimum and maximum family wealth indexes are both in social science colleges, and the social science colleges have the largest standard deviation. These phenomenon indicate that the students in social science colleges have serious polarization in family wealth.
- The students with the maximum average and the minimum standard deviation of family wealth index are both in arts colleges, which shows that the family financial condition of students in arts colleges is relatively good and balanced, which accords with subjective cognition.

### 3.4. The Definition of Accuracy Rate

We divide all the students into four levels according to their financial condition: high, middle, low, and none of above. In the traditional artificial identification, a total of 4,423 investigated students are identified as in high, middle, low level and none of above at the ratio of 229: 561: 499: 3134.

According to the ratio in the model identification we rank the students according to their family wealth index and thus students from 1 to 229 are identified as in high level of financial hardship, students from 230 to 790 in middle level, students from 791 to 1,290 in low level, and the remaining students as in none of above index.

We use  $h_1(i)$  and  $h_2(i)$  to represent the level of financial hardship of student  $i$  by artificial identification and TOPSIS model respectively, and  $h_1(i)$  (or  $h_2(i)$ ) = 0, 1, 2, 3 stands for levels of high, middle, low, and none of above respectively.

We define the accuracy rate as

$$P = 1 - \frac{\sum_{i=1}^m \chi_A(h_1(i) - h_2(i))}{m + m_0} \quad (9)$$

where  $m_0$  is the number of students identified as in high level financial hardship directly,  $m$  is the number of the remaining students.  $\chi_A(h_1(i) - h_2(i))$  is an indicator function, which is defined as

$$\chi_A(h_1(i) - h_2(i)) = \begin{cases} 0, & h_1(i) = h_2(i) \\ 1, & h_1(i) \neq h_2(i) \end{cases} \quad (10)$$

### 3.5. The Results of Accuracy Rate

The comparison result (except those 631 students identified

as in high level directly) of artificial and model identification are shown in Tab.5.

**Table 5.** The comparison result of artificial and model identification.

		Model identification				Column total
		High	Middle	Low	None of above	
Artificial identification	High	48	82	43	56	229
	Middle	88	172	120	181	561
	Low	48	128	94	229	499
	None of above	45	179	242	2668	3134
Row total		229	561	499	3134	4423

Tab. 5 shows that

- Our model identifies 229 students as in high level of financial hardship, in which 48, 88, 48 and 45 students are identified as in high, middle, low and none of above level in artificial identification, respectively.
- Our model identifies 561 students as in middle level of financial hardship, in which 82, 172, 128 and 179 students are identified as in high, middle, low and none of above level in artificial identification, respectively.
- Our model identifies 499 students as in low level of financial hardship, in which 43, 120, 94 and 242 students are identified as in high, middle, low and none of above level in artificial identification, respectively.
- Our model identifies 3,132 students as in none of above level, in which 56, 181, 229 and 2,668 students are identified as in high, middle, low and none of above level in artificial identification, respectively.

Then the accuracy rate is

$$P = 1 - \frac{1444}{5054} = 71.43\%$$

The result shows that the accuracy rate of our model is 71.43%, which means that there are, on average, 71.43 students identified accurately per 100 students. Therefore, our model is of great reference value for identifying students in financial hardship. We propose that the following steps should be carried out in identifying students in financial hardship in practice.

- Get the initial level of financial hardship by estimating

the family wealth index according to our model.

- Accept application for review from students. According to the conclusion, there are only less than 30\% of the students will apply for review.
- Identify reviewed students artificially, since education is the work of humanity, cannot be replaced by model completely, this step is also important.
- Obtain the final result of identification.

To sum up, our model can improve the efficiency by about 70%, and can unify the standards in identifying between different colleges.

### 3.6. Sensitivity Analysis

There exist subjective factors in determining the value of indexes in Tab. 1. In order to test that whether our model is robust, we change the weights slightly to recalculate the accuracy rate. The following six methods have been adopted.

- All the values are multiplied by 0.9.
- The value of family residence is multiplied by 0.9.
- The value of parents' educational level is multiplied by 0.9.
- The value of parents' occupation is multiplied by 0.9.
- The value of the number of personal consumer electronics is multiplied by 0.9.
- The value of whether loan or not is multiplied by 0.9.

The results of sensitivity analysis are shown in Tab. 6. We can see that, when the weight changed by 10%, the relative changing rate of accuracy rate is less than 7%, which indicate that our model is robust.

**Table 6.** The results of sensitivity analysis.

Method	Original Method	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Accuracy rate (%)	71.43	71.43	69.75	70.32	70.64	68.75	72.49	66.45
Relative change rate (%)	0	0	2.35	1.55	1.11	3.75	1.48	6.97

## 4. Conclusion

In this paper, we investigated the identification of students in financial hardship. In the first part, the index system of students' background is constructed and the TOPSIS model is utilised to estimate the family wealth index of students. The data from South-Central Minzu University were used for the empirical analysis. The result showed that the accuracy rate of TOPSIS model is 71.43%. The sensitivity analysis revealed

that our model is robust, which indicates that our model can be used in the identification.

In real world applications, our model can be used to get the initial level of students in financial hardship. According to the initial level, colleges accept the application for review from students, identify reviewed students artificially and obtain the final result of identification. The empirical analysis showed that our model can improve the efficiency by about 70%. Moreover, the subjectivity of artificial identification is excluded in our model.

In addition to the growth background of students, the consumption of students in university is also the main factor to estimate the family wealth index. How to adjust the level of financial hardship according to the consumption of students (according to the data offered by Campus-ID card) will become our next focus.

## Acknowledgments

The authors are thankful to the sector of student affairs in SCMZU (especially Mrs. Li hong and Mr. Wei Dajiang), who provides the data and support for our research. This work is supported by “the Fundamental Research Funds for the Central Universities”, South-Central University for Nationalities (CSQ17005).

---

## References

- [1] He F. (2011): “The Existing Problems and Strategies in Identification of College Students in Financial hardship”, *Ideological and Political Education Research*, Vol. 27, No. 1, pp. 132-134. (In Chinese).
- [2] Wang M., Wang S. (2013): “Econometric Analysis of the State Grant Assessment Factors: A Case Study on Poisson Regression Model Based on the National Institutes of Hubei”, *Journal of Zhengzhou Normal Education*, Vol. 2, No. 2, pp. 84-87. (In Chinese).
- [3] Yang X. (2009): “Thinking and Practice on the Model of Identifying College Students with Financial Difficulties”, *Ideological and Theoretical Education*, Vol. 33, No. 2, pp. 79-82. (In Chinese).
- [4] Jin X. (2009): “Implementation of Double Aid to ensure the sustainable development of Funding College Students”, *Chinese Higher Education*, Vol. 53, No. 3, pp. 70-71. (In Chinese).
- [5] Jensen E. (1996): “The fertility impact of alternative family distribution channels in Indonesia”, *De-mography*, Vol. 33, No. 1, pp. 153-165.
- [6] Dargent-Molina P., James S., Strpoatz D., Savitz D. (1994): “Association between maternal education and infant diarrhea in different household and community environments”, *Social Science and Medicine*, Vol. 38, No. 2, pp. 343-350.
- [7] Layte R., Maitre B., Nolan B., Whelan C. (2001): “Persistent and consistent poverty in the 1994 and 1995 waves of the European Community Household Panel”, *Review of Income and Wealth*, Vol. 47, No. 4, pp. 427-450.
- [8] Tian Z., Yuan L. (2010): “Empirical Analysis on Identifying Needy College Students with Non-income Indicators”, *Peking University Education Review*, Vol. 8, No. 2, pp. 145-157. (In Chinese).
- [9] Si S., Sun X. (2011): “Mathematical Modeling Algorithm and Its Application”, *National Defense Industry Press*. (In Chinese).
- [10] Jia J., He X. (2015): “Statistics (sixth edition)”, *China Renmin University Press*. (In Chinese).