

# Do Living Habits Influence the Innovation Ability of College Students? Evidence from Student Information System

Xin Lin and Yizhou Tang\*

School of Economics and Management, Tongji University, Shanghai, China

\*Corresponding author

**Abstract**—Analyzing the relationship between living habits and innovation ability, this paper utilizes a big data approach to analyze information from student system. Firstly the paper established a comprehensive evaluation model to appraise college students' innovation capability based on Analytic Hierarchy Process. To figure out whether living habits have an impact on students' innovation ability, a multiple regression model is carried out. Considering the possibility of endogenous problem, an instrument variable is introduced to the basic model. The 2SLS regression results indicate that good living habits do help in enhancing students' innovation capacity in college, which is convincible since the robustness checks didn't reject it.

**Keywords**—college student; innovation ability; living habits; AHP; 2SLS

## I. INTRODUCTION

Innovation is served as a propeller to the technology development and social progress. And the innovation capacity is a key element to determine the performance of college students and it is also a main ability our universities intend to cultivate.

Numbers of researchers have made effort to evaluate the innovation ability of university students. Li (2014) employed Analytic Hierarchy Process to build evaluation model of college students' ability to innovate. Dong and Rao (2015) combined AHP and Fuzzy Mathematic Theory to build a comprehensive evaluation model of scientific innovation activities for college students. In the problem of fostering college student, scholars had put forward some methods. Liu and et al. (2009), Yao and Liu (2012) argue that institute should establish correct encouraging system, create an environment and platform for innovation and reform education the traditional teaching method [3]. Chen and et al (2013) integrate the advantages of Creativity Study and Extenics and carried out a training model for promoting students' ability to innovate. And the outcomes are proved by practical teaching cases. However, the existing researches don't investigate the internal factors those affect the college students' creativity. By establishing a comprehensive evaluation index of innovation capability, this paper analyzed the casual effect between living habits and innovation capacity, which will provide a new angle for the research of this field.

The rest of this paper is organized as follows. In the next section, database and data processing specifications are

introduced. In section 3, empirical analysis and its results are put forward. Section 4 conducts the robustness check. The conclusion is given in the last section.

## II. DATABASE

### A. Data Processing

In this section, we illustrate the database and our processing procedure. The database we use is gained from the information system of a well-known university in China. It contains huge-sized specific data about each enrolled student in the university. The covering period is from September 2014 to March 2015. After matching and merging the ID number, 37653 students are identified as the observations.

Based on the data acquired, we extract some variables which may be useful in the analyzing below. They are: marks, borrowing amount, numbers of paper publication, numbers of patents, numbers of in and out of the dormitory, consumptions times in school, dummy of local student, dummy of minority and dummy of assistance.

Here, the marks of different courses are all transformed to hundred-mark system. The borrowing amount is the number of borrowing recorded in the library. Based on the numbers of in and out of the dormitory, we can obtain the data of students' living habits. We defined a student go out of the dormitory between 6:00 am to 8:30 am as having good living habit and go out between 23:00 pm to 2:00 am as bad. Then we can calculate the numbers of going out early and going out late. The consumptions times in school indicates the student's consuming records inside the school, such as eating in the college canteens. In addition, we define early morning consumption is occurs between 6:00 am to 8:30 am. Dummy of local student defines whether a student belongs to local people. Dummy of minority indicates a student is a national minority or not. And dummy of assistance is assign to 1 if a student serves as assistance, 0 if not.

### B. Variables Description

- a) ID: student identified number;
- b) cj: course mark;
- c) jynum: borrowing amount;
- d) lwnum: paper publication amount;

- e) zlnum: patent acquired amount;
- f) xfnun: consumption times;
- g) inov: comprehensive index of innovation ability;
- h) zaonum: early morning consumption times;
- i) zaochu: early morning leaving dormitory;
- j) wanchu: midnight leaving dormitory;
- k) sz: dummy of student serve as assistance;
- l) location: dummy of local student;
- m) minor: dummy of national minority student

### C. Basic Assumptions

According to the practical problem and the availability of data, we make the assumptions as follows:

- a) Students leaving dormitory or consuming at the early morning are defined as keeping a good living habit;
- b) Students can make consumptions inside or outside the campus and this two can substitute each other;
- c) The difference between publication papers and patents are limited to quantity, not quality;
- d) All the borrowing conduct is unity. Only the amount differs, no matter what content it is;
- e) Every individual's data is completed during the recorded period without omissions.

## III. EMPIRICAL STUDY

### A. Evaluation Model

In this part, we consider establishing a comprehensive evaluation model to appraise college students' innovation capability. Three main indexes are selected to reflect the ability of innovation. Then the main task is to figure out reasonable weights for these three indexes. Here, the Analytic Hierarchy Process was employed to modeling.

Analytic Hierarchy Process, which is short for AHP, is a structured technique for organizing and analyzing complex decisions, based on mathematics and psychology. It was developed by Thomas L. Saaty in the 1970s and has been extensively studied and refined since then.

Utilizing AHP to construct system model take the following five steps:

- a) Establish hierarchical structure model;
- b) forming decision matrix
- c) calculate the weights
- d) consistency checks

We firstly construct innovation capacity index based on AHP. Based on the data we extracted, we describe students' creativity by three key aspects: course marks, paper issuance and patent admission. Then we set up the decision matrix according to the AHP approach. By comparing the relationship

and significance among the three elements, we consider the course grade is "little important" to creativity while the issue publication is "very essential" and the patent acquired is of "huge importance". Because the marks are largely reflect students' diligence and perseverance. And the issue indicates a student's academic ability. The patent, by no means, demonstrates a student's innovative thinking and creativity. Thus, we obtain the decision matrix about the three indexes below:

$$A = \begin{pmatrix} 1 & \frac{1}{7} & \frac{1}{9} \\ 7 & 1 & \frac{1}{3} \\ 9 & 3 & 1 \end{pmatrix} \quad (1)$$

Where the number of 1, 7 and 9 are used to indicate the marks, paper and patent are little important, essential, extremely important to innovation, respectively. According to the computation, the weights of the three elements are available:

$$W = [0.06 \quad 0.3 \quad 0.64] \quad (2)$$

Then we calculate the consistency index to check the results. The consistency index is  $C.I. = 0.043$ .

According to different numbers of element, Saaty give the reference index value, which is showed in the below:

TABLE I. REFERENCE INDEX VALUE BY SAATY

| n    | 1 | 2 | 3    | 4   | 5    | 6    | 7    |
|------|---|---|------|-----|------|------|------|
| R.I. | 0 | 0 | 0.58 | 0.9 | 1.12 | 1.24 | 1.32 |

Referring to the table above, we can figure out the random conformance rate:

$$C.R. = \frac{C.I.}{R.I.} = 0.074 < 0.1 \quad (3)$$

The random conformance rate calculating here is less than 0.1, which indicates that the result is reasonable.

Here, we can obtain the comprehensive innovation capability index utilizing the weights modeling from AHP. The computing approach is presented as the equations below:

$$inov = 0.06 \times cj + 0.3 \times lwnum + 0.64 \times zlnum \quad (4)$$

### B. Descriptive Statistics

This part presents the descriptive statistics of the data we used. After cleaning abnormal value, 35780 observations are identified. The course marks frequency chart is showed below:

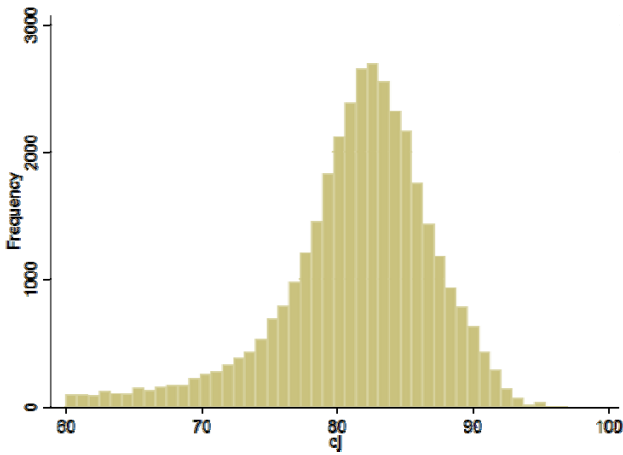


FIGURE I. FREQUENCY OF STUDENTS' MARKS

From the figure, we can find that most frequent marks distribute from 80 to 85. The numbers of marks below 60 and beyond 95 is limited. The marks here are average grade.

Based on the equation (4), the innovation ability index can be calculated. Its frequency distribution is show in the Figure 2. And the main indicators' descriptive statistics are list in the Table 2.

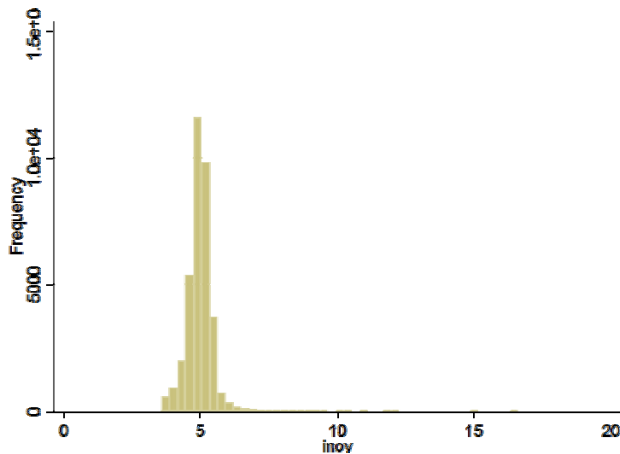


FIGURE II. FREQUENCY OF STUDENTS' INNOVATION ABILITY INDEX

TABLE II. DESCRIPTIVE STATISTICS OF MAIN INDICATORS

| Variable        | Obs   | Mean  | Std. Dev. | Min | Max   |
|-----------------|-------|-------|-----------|-----|-------|
| <i>cj</i>       | 35300 | 81.46 | 5.58      | 60  | 97    |
| <i>lwnum</i>    | 35780 | 0.24  | 0.70      | 0   | 22    |
| <i>zlnum</i>    | 35780 | 0.02  | 0.26      | 0   | 18    |
| <i>inov</i>     | 35300 | 4.97  | 0.45      | 3.6 | 16.61 |
| <i>zaochu</i>   | 35780 | 0.45  | 4.23      | 0   | 184   |
| <i>wanchu</i>   | 35780 | 0.74  | 5.82      | 0   | 150   |
| <i>jynum</i>    | 35780 | 13.85 | 35.74     | 0   | 915   |
| <i>sz</i>       | 35780 | 0.19  | 0.39      | 0   | 1     |
| <i>location</i> | 12992 | 0.33  | 0.47      | 0   | 1     |
| <i>minor</i>    | 13291 | 0.03  | 0.19      | 0   | 1     |

### C. Empirical Results

To analyze the problem of how living habits affect the innovation ability, we established a multiple regression model.

$$\begin{aligned}
 inov_i = & \alpha_i + \beta_1 \cdot zaochu_i + \beta_2 \cdot wanchu_i + \beta_3 \cdot jynum_i \\
 & + \beta_4 \cdot sz_i + \beta_5 \cdot location_i + \beta_6 \cdot minor_i + \delta_i
 \end{aligned} \quad (5)$$

Here the variables of interest are *zaochu* and *wanchu*, they proxy for good living habit and bad living habit, respectively. In order to control other factors' effect on innovation, we consider some control variables. The control variables include borrowing amount, being assistance, being a local student and being national minority. The regression result is list in the first column of Table 3.

Then we consider the problem of endogeneity. Students leaving early in the morning don't necessary stands for a good living habit, which may result in a bias estimation. Because it is possible that student have to get up early or leaving early for other things, such as having fun or going out of the campus. Thus, we consider introduced an instrument variable for the variable of interest. The consuming times in the morning are chosen as the instrument because a student likely consuming early in the campus stands for he is more likely utilizing the campus resources and stay in the campus. Utilizing the consuming times to predict the good living habit can identify the students who are willing to get up early themselves. Then put the forecasting value into the regression can obtain an exogenous estimator. This approach is conducted by two-stage least squares (2SLS). The regression result is listed in the second column of Table 3.

TABLE III. REGRESSION ESTIMATION RESULTS

| Independent Variable | Dependent Variable: <i>perform</i> |                        |
|----------------------|------------------------------------|------------------------|
|                      | Model(1)-OLS                       | Model(2)-IV            |
| <i>zaochu</i>        | -0.0034***<br>(-5.32)              | 0.0149***<br>(3.97)    |
| <i>wanchu</i>        | -0.0053***<br>(-11.67)             | -0.0064***<br>(-12.34) |
| <i>jynum</i>         | 0.0010***<br>(12.13)               | 0.0011***<br>(11.99)   |
| <i>sz</i>            | 0.2238***<br>(24.47)               | 0.1987***<br>(18.55)   |
| <i>location</i>      | 0.0866***<br>(9.19)                | 0.0922***<br>(9.42)    |
| <i>minor</i>         | -0.0551**<br>(-2.42)               | -0.0616***<br>(-2.62)  |
| <i>cons</i>          | 4.9891***<br>(631.37)              | 4.9791***<br>(592.44)  |
| <i>Adj R-squared</i> | 0.0763                             | 0.0155                 |

t statistics in parentheses and \* p<0.10, \*\* p<0.05, \*\*\* p<0.01

From the table above, we can find that all the variables are statistical significance at 0.01 levels. The adjusted R-squared of two models are 0.076 and 0.015, respectively. Because what we focus is the casual effect and the using 2SLS estimation, the R-squared doesn't matter in inference [6], which could be ignored. All the coefficients of the main variables didn't vary too much in the absence of control variables.

In the first column, the coefficient of *zaochu* is negative which may be contradicting with our intuition. The other variables sign is consistence with normal instinct. Thus we consider the variables may be endogenous and he estimator is biased. In the second column, however, the sign of *zaochu* is changed to positive, which is in line with our intuition.

The results indicate that a student with a good living habit will gain better innovation ability while a bad living habit will decrease it. The magnitude of good living habit on innovation is larger than borrowing amount. In addition, being an assistance maybe help in enhancing students' creativity. We explained that a student being assistance can be influenced by what one constantly sees and hears from their supervisor, which may help them cultivate an innovative thinking. The results also show that local student and non-national minority student will have a better performance on innovative activity.

#### IV. ROBUSTNESS CHECKS

In this section, we make the robustness checks of our estimation. Firstly, we justified the necessity of instrument variable. An OLS regression and instrument regression are run separately. Then we conduct the Hausman test to check. The null hypothesis of Hausman test is that all the variables are exogenous. If it is accepted, the estimation of the two approaches is no difference. If it is rejected, we should consider the problem of endogeneity. According to the Hausman test, the p value of the test is less than 0.01, which indicates that the null hypothesis can be rejected with 99% confidence. Thus, the two-stage least squares approach we took is reasonable.

Then we perform the under identification test and weak identification test on the instrument variable to check the reasonability of instrument. the results are listed in the table IV.

TABLE IV. TEST ON THE INSTRUMENT VARIABLE

|   |        |
|---|--------|
| Underidentification test<br>(Anderson canon. corr. LM statistic)  | 382.63 |
| Chi-sq(1) P-val   | 0      |
| Weak identification test<br>(Cragg-Donald Wald F statistic)       | 394.33 |
| Stock-Yogo weak ID test critical values:                          |        |
| 10% maximal IV size   | 16.38  |
| 15% maximal IV size   | 8.96   |
| 20% maximal IV size   | 6.66   |
| 25% maximal IV size   | 5.53   |
| Sargan statistic<br>(overidentification test of all instruments): | 0      |

The test results of our instrument indicate that the instrument variable didn't suffer from the weak identification and under identification. The result we obtained from this study is robustness.

#### V. CONCLUSION

This paper starts with a huge database. After processing and mining of the dataset, the information of college student is utilizing the analyses their living habits performance. The processing work considers using the details as much as possible, which maximize the value of the database. Based on the data extracted, AHP approach was introduced to construct college students' innovation ability evaluation model. Utilizing the innovative index, a multiple regression model is carried out to examine the casual effect between living habit and innovation capacity. The result demonstrated that college students with a good living habit will improve their creativity while bad living habit will be detrimental to it. Taking the issue of endogeneity into account, instrument variable is introduced. The acquiring results are convincible since the robustness checks didn't reject it.

Our results provide a novel direction for the research of college students' innovation ability. In the future works, the database will be mining more specifically and more related factors will be included to explained college students' creativity.

#### REFERENCES

- [1] Li, L, An Innovation Capability Evaluation Model Based on AHP and Training Strategy of College Students, IEEE International Conference on Service Operations and Logistics, and Informatics, 2014, pp. 198-201.
- [2] Dong, YY and Rao, CJ, Comprehensive Evaluation Model and Its Application for Scientific In-novation Activities of College Students, Internation-al Conference on Computer Science and Environmental Engineering, 2015, pp. 817-824.
- [3] Liu, M and et al., Innovation Practice Training Model of College Students in Background of Knowledge Economy, Proceedings of the 2009 Pacific-Asia Conference on Circuits, Communications and Systems, 2009, pp. 51-53.
- [4] Yao YL and Liu, JP, Cultivation of College Students' Innovation Quality in Electrotechnics Ex-periment Teaching, Lecture Notes in Information Technology, 2012(12), pp. 103-107.
- [5] Chen, AL and et al., Study on Innovation Capability of College Students Based on Extenics and Theory of Creativity, Procedia Computer Science, 2013(17), pp. 1194-1201.
- [6] Jeffrey M.Wooldridge, Introductory Econometrics: A Modern Approach, forth ed., China Renmin University Press, Beijing, 2010.