

Two-dimensional Vectorial Classification for American University

Ziqian Yang

North China Electric Power University, Huadian-Lu 689, 071000 Baoding, China
peteryangzq@gmail.com

KeyWord: Factor Analysis Two-dimensional Vectorial Classification

Abstract. The Goodgrant Foundation declares that they intend to donate a total of \$100 million to higher education. To supply the investment plan to them, we devise a model aim to select a set of schools making most advantages of funding.

With Factor Analysis method, we gain two principal factors (tuition payment factor P and teaching achievement factor A) to reflect demand level of school. The two principal factors can compose an orthogonal evaluation system. We think medium scored schools have best development potential and schools with explicit discrepancy of A and P have more specific goals of development. So, we regard schools satisfying two conditions above as the results. And we classify selected schools into two categories, PI (lower P and higher A) and AI (lower A and higher P), which need different investment policy.

Introduction

Nowadays a college degree or postsecondary certificate is more important than ever before. National Outcome Goals for Postsecondary Education, Career-Technical Education, and Adult Education focus on improving student achievement to **increase high school graduation, college completion, and educational attainment**[1]. While such goals can be utilized as measurements on specific institution for official evaluation, other metrics measuring institutional performance to prospective students are crucial, considering many students, especially underserved, may not even fully understand which college is within reach. To help them make better choices, the metrics selected to reveal several exemplar institutions that serve students well are **access, affordability and student outcomes**.

One of the advantages of America's vast higher education system is the variety and number of its colleges and universities. And each institution can be approximately divided into various categories such as public or private, profit or non-profit, four-year or two-year referring to its property, size and setting[2].

Factor Analysis Method

Step 1: Determine our own metrics.

While measurements U.S. Department of Education sets forth can distinguish institutional performance for official evaluation, we have to modify the metrics addressing the demand for grant aid corresponding to various institutions.

(1) Get rid of indicators for lack of available data. The primary data source of our devisal is IPEDS [3] data set which do not supply such specific indicator statistics as transfer rate and learning outcomes. Consequently, we foreclose those indicators from the measurement since other available metrics can reflect information obliterated due to certain correlation between indicators.

(2) Take out indicators irrelevant to institution's demand for grant aid. We just remain Net Price Index and Percentage of Pell Students Index with respect to an institutional income and student's demand for grant aid implying the institutional demand. Other indices such as Borrowing and Debt belonging to access and affordability are ignored.

Overall, we devise a new evaluation system for the institutional demand for grant aid with six following indices: Net Price, Percentage, 3-year repayment rate, Completion rate, Median earnings and Proportion of moderate- and high-income.

Step 2: Make factor analysis on indices.

Using above six indices, which are likely to correlate with each other, as measurements may

affect accuracy of investment results. In order to eliminate such correlation, we make factor analysis [4] on these indices corresponding to the four types of institutions.

The results of factor analysis are as depicted in **Table 1**.

Table 1 The results of factor analysis

		Public institution		Private institution	
	Indicators	Factor Score			
		Factor1	Factor2	Factor1	Factor2
Institution for four years	Net Price	0.231	0.075	0.241	0.017
	Percentage of Pell Students	-0.334	0.103	-0.327	0.028
	3-year repayment rate	0.338	-0.096	0.362	-0.144
	Completion rate	0.301	0.068	0.311	-0.02
	Median earnings	-0.033	0.504	-0.053	0.52
	non-low-income students	-0.04	0.501	-0.106	0.549
	Cumulative contribution	0.768		0.785	
Institution for less than four years	Net Price	0.063	-0.434	-0.22	-0.134
	Percentage of Pell Students	-0.048	0.439	0.007	0.437
	3-year repayment rate	0.297	0.003	0.019	0.373
	Completion rate	0.031	0.541	-0.093	0.501
	Median earnings	0.428	-0.04	0.466	-0.09
	non-low-income students	0.455	-0.069	0.47	-0.049
	Cumulative contribution	0.802		0.793	

Analysis:

(1) The former two factors explain all variables by 75% for each kind of school, so we can use two principal factors to represent overall original variables. Additionally, mutually orthogonal principal factors eliminate the correlation with variables.

(2) Both two principal factors make sense in reality. For four-year universities, first principal factor, defined as Payment Factor (P), mainly scores through Net Price, Percentage of Pell Students, 3-year repayment rate and Completion rate, which reflects student payment for tuition and loan; and second principal factor, defined as Achievement Factor (A), mainly scores through Median earnings and Proportion of non-low-income, which reflects the competence after graduation. Furthermore, the two principal factors are in reverse order for two-year colleges.

Two-dimensional Vectorial Classification Method

We devise a two-dimensional vector (P,A) composed of two principal factorsto assessing potential candidate schools. Instead of conventionally formulizing the comprehensive factor score with two principal factors as final index, we draw the A-P distribution of each school regarding A as Y-axis and P as X-axis. Therefore, each school occupies a point (P,A) in frame of reference as shown in **Figure 1**.

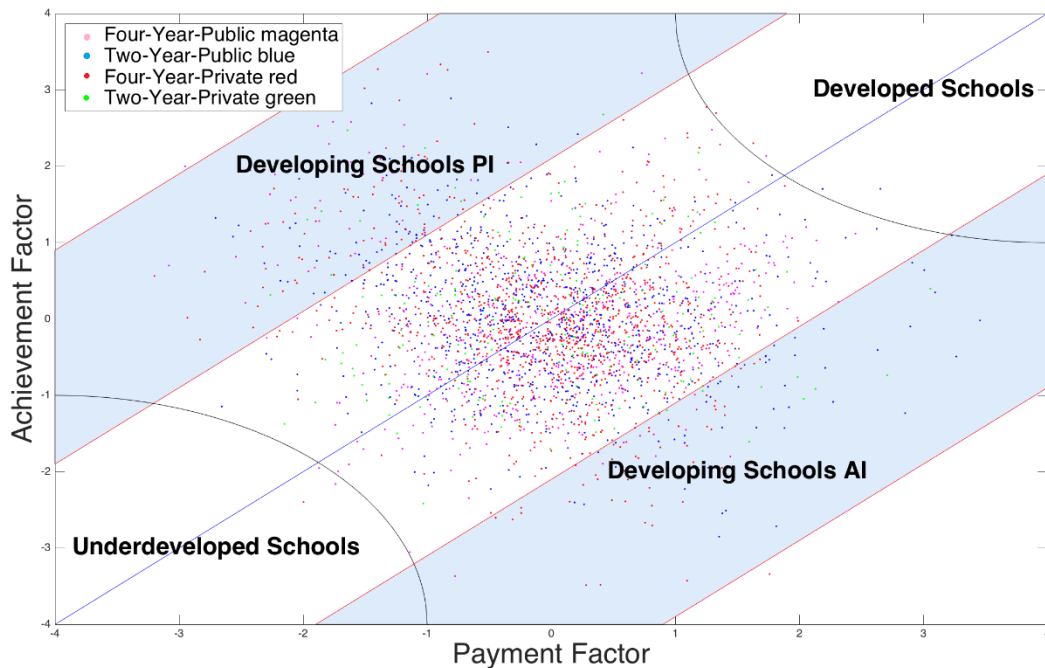


Figure 1 The classification of schools on a two-dimensional performance-based vector

Referring to the identifications of developing and developed country by analogy, we classify colleges into three classes.

- (1) Developed school: both A and P are high;
- (2) Undeveloped school: both A and P are low;
- (3) Developing school: others.

As depicted in **Figure 2**, developed schools and undeveloped schools locate at lower left and higher right quadrants, and others are developing schools.

Several schools are selected randomly in each district to justify our assessment through finding their relevant information. For instance, we choose a school located in developed school district, of which UNITID=231624, P=2.4521, and A=1.1194. After inquiring information of the school, we find that it is College of William & Mary with second longest history in U.S. and first-rate setting of hardware and software; it is a public university with lower tuition and higher P. Overall, assessing College of William & Mary as a developed school is reasonable.

Select schools appropriate to invest.

The principles of investment choice devise that a priority is given to schools with

- (1) Desperate demand for aggrandizing funding and certain development basis ensuring that funding can be effectively utilized and
- (2) Explicit relatively developing disadvantages, that is, obvious effect is achieved by improving weakness first with limited funding.

According to these principles, conclusions can be drawn as:

- Developed schools with better quality of indices have entered saturation of development and are not in urgent demand for our funding;
- Undeveloped schools with worse development basis need heavy investment to attain better level, which is beyond our financial reach;
- Developing schools with major development potential desperately require extra funding aid for rapid expansion, so we tend to invest these schools. What's more, considering the second principle describing a investment tendency to schools with moderate disequilibrium between A and P, we restrict investment goals to schools located at blue areas.

Conclusion

Based on classification and screening method in **Figure 1**, we select 310 developing schools accounting for 10.45% of all candidate schools and consisting of 193 PI and 117 AI. **Table 2** lists

the statistical results of selected schools. The proportion of selected schools belonging to each category is the same with the proportion of selected, which justifies our model's applicability to four school categories. **As a result**, we could invest those schools by which we can get a better effect.

Table 2 The results of screening for Model

	4 year public	4 year private	2 year public	2 year private	Sum
PI	40	77	65	11	193
AI	10	49	49	9	117
Sum	50	126	114	20	310
Total	556	1170	1041	197	2964

Strengths and Weakness

1. In reference to the evaluation system of U.S. Department of Education, we find six applicable metrics for our model, benefiting from avoiding reframing metrics while limiting the ideas and leading inefficient utilization of information supplied. Additionally, we revise the earning index by employment of each race to remove unreasonable elements in metrics.

2. Factor Analysis eliminates indices correlation with each other, accruing the results of our model. However, replacing original metrics with principal factors causes component loss which decreases information integrality.

3. Two principal factors A and P explicitly make sense in reality, improving the credibility of our model. And it is more practical to construct a two-dimensional vector based on two principal factors than creating a comprehensive index, which provides opportunity to make in-depth analysis. The division of potential candidate schools into AI and PI respectively according to their disadvantages benefits targeted investment while the setting of thresholds is slightly subjective.

References

- [1] U.S. Department of Education Strategic Plan for Fiscal Years 2011–2014
<http://www2.ed.gov/about/reports/strat/plan2011-14/draft-strategic-plan.pdf>.
- [2] Seeking Smart Growth in Higher Education: A Look at the Structure of US Higher Education ,
 Past and Future John Aubrey Douglass Translated by XU Dan
 (Center for Studies in Higher Education, University of California - Berkeley , CA 94720 , USA)
 (School of Education , Hunan University , Chang sha , Hunan 410082 , China)
- [3] Information on <http://nces.ed.gov/ipeds/datacenter/Default.aspx>