

정규화 및 통합 방법이 순위의 변동성과 순위 역전에 미치는 영향

박 영 선^{*†}

^{*} 서경대학교 경영학부

Effects of Normalization and Aggregation Methods on the Volatility of Rankings and Rank Reversals

Park, Youngsun^{*†}

^{*} Division of Business Administration, Seokyeong University

Abstract

Purpose: The purpose of this study is to examine five evaluation models constructed by different normalization and aggregation methods in terms of the volatility of rankings and rank reversals. We also explore how the volatility of rankings of the five models changes and how often the rank reversals occur when the outliers are removed.

Methods: We used data published in the Complete University Guide 2014. Two universities with missing values were excluded from the data. The university rankings were derived by using the five models, and then each model's volatility of rankings was measured. The box-plot was used to detect outliers.

Results: Model 1 has the lowest volatility among the five models whether or not the outliers are included. Model 5 has the lowest number of rank reversals. Model 3, which has been used by many institutions, appears to be in the middle among the five in terms of the volatility and the rank reversals.

Conclusion: The university rankings vary from one evaluation model to another depending on what normalization and aggregation methods are used. No single model exhibits clear superiority over others in both the volatility and the rank reversal. The findings of this study are expected to provide a stepping stone toward a superior model which is both reliable and robust.

Key Words: Volatility of Rankings, Rank Reversal, Normalization, Aggregation, Outlier

• Received 11 October 2013, 1st revised 24 October 2013, accepted 25 October 2013

† Corresponding Author(yspark@skuniv.ac.kr)

© 2013, The Korean Society for Quality Management

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-Commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. 서론

기관을 평가하여 순위를 결정하는 것은 매우 중요한 일이다. 기관을 평가할 때는 하나의 기준만을 고려하지 않고 여러 가지 기준을 동시에 고려하여 평가하는 것이 일반적이며 이러한 문제를 다기준 의사결정(multi criteria decision making)이라 한다. 기관을 평가할 때 사용되는 기준을 무엇으로 할 것인가와 각각의 기준이 얼마나 중요한지를 고려해 가중치를 얼마로 할 것인가는 평가 결과에 커다란 영향을 미치게 된다. 평가주체에 따라 평가 기준과 가중치는 다른 방식으로 결정되므로 주체별 평가결과는 서로 상이하게 발표되고 있다(Luckman et al., 2010, Saisana, et al., 2011).

기준들은 각각의 척도에 따라 측정되며 이러한 척도들은 서로 상이하므로 측정된 값을 비교 가능한 형태로 전환할 필요가 있다. 이러한 과정을 자료 정규화(data normalization)라고 하는데 그 방법은 다양하다. 기관에 대해 같은 기준과 같은 가중치를 사용하여 평가하더라도 정규화 방법에 따라 그 결과는 차이가 날 수 있다. 정규화 된 자료에 대해 가중치를 고려하여 통합(aggregation)하는 방법에 따라서도 순위의 차이가 날 수 있다.

정규화와 통합 방법에 따라 순위 결정 모형 5가지를 선택한다. 본 연구의 목적은 5가지 모형별로 순위의 변동성(volatility)을 측정하고, 자료에 있는 이상점(outlier)을 제거하였을 때 모형별 순위의 변동성과 순위 역전(rank reversal)이 어떤 영향을 받을 것인지를 밝혀내는 것이다.

본 연구에서 이용한 자료는 영국 CUG(complete university guide)에서 발표한 ‘University League Table 2014’이다. 여기에서 발표된 순위는 9개의 기준을 이용한 것이다. 9개의 기준으로 평가된 값을 사용하여 5개 모형으로 각 대학의 순위를 결정한다. 각 대학의 순위는 5개 모형별로 다르게 나타날 수 있다. 각 대학의 모형별 순위 변동성은 5개 모형에서 결정된 그 대학의 순위에 대한 중앙값(median)과 그 대학의 모형별 순위의 절대편차로 측정한다. 순위 변동성이 작은 모형이 바람직하다고 볼 수 있다.

자료에 있는 이상점을 발견하고 이상점을 갖고 있는 대학을 제거한 후에 각 대학의 모형별 순위 변동성을 측정한다. 이 값을 전체 대학의 모형별 순위 변동성 값과 비교함으로써 이상점이 순위 변동성에 미치는 영향을 파악한다. 두 대학을 비교하였을 때 전체 대학의 순위에서 한 대학이 높은 순위를, 이상점을 제거한 후의 대학 순위에서는 다른 대학이 높은 순위를 차지하는 경우를 순위 역전이라고 한다(Filinov, et al., 2002). 모형별 순위 역전이 얼마나 발생했는지 파악한다. 순위 역전이 적게 발생하는 모형이 가장 바람직하다고 평가할 수 있다.

분석 결과를 통해 모형별로 순위 변동성과 순위 역전을 일으키는 원인이 무엇인지 규명하고 각 모형의 특징을 찾아냄으로써 새로운 모형의 개발에 도움을 줄 수 있을 것이다.

2. 다기준 의사결정

2.1 다기준 의사결정의 개요

다기준 의사결정은 n 개의 기관을 m 개의 기준으로 평가하는 것이다. 평가 기준에 따라 사용된 척도는 ‘클수록 좋은 것’과 ‘작을수록 좋은 것’으로 구분할 수 있다. 예를 들어 대학의 입학성적이 높을수록 높은 평가를 받는 반면, 교수 1인당 학생 수는 적을수록 높은 평가를 받게 된다. 평가 기준과 각 기준에 부여된 가중치는 주어진 것으로 가정하고 다음과 같이 기호를 정의한다.

$A = \{A_1, A_2, \dots, A_n\}$: 평가 대상이 되는 n 개의 기관

$C = \{C_1, C_2, \dots, C_m\}$: m 개의 평가 기준

$X = \{x_{ij}, i = 1, 2, \dots, n, j = 1, 2, \dots, m\}$: 기관 i 의 기준 j 에 대한 평가값

$W = \{w_1, w_2, \dots, w_m\}$: m 개 평가 기준의 가중치, $\sum_{j=1}^m w_j = 1$

$J^+ = \{\text{값이 클수록 좋은 기준 } j \text{의 집합}\}$

$J^- = \{\text{값이 작을수록 좋은 기준 } j \text{의 집합}\}$

$$\mu_j = \frac{\sum_{i=1}^n x_{ij}}{n} : \text{기준 } j \text{의 평가값 평균, } j = 1, 2, \dots, m$$

$$\sigma_j = \sqrt{\frac{\sum_{i=1}^n (x_{ij} - \mu_j)^2}{n}} : \text{기준 } j \text{의 평가값 표준편차, } j = 1, 2, \dots, m$$

다기준 의사결정은 n 개의 기관 중 가장 우수한 하나의 기관을 선택하는 문제도 있고, 기관들의 순위를 결정하는 문제도 있다. 본 연구에서는 모든 기관의 순위를 결정하는 문제로 제한한다. 모든 기관은 각 기준에 따라 서로 다른 척도로 평가되기 때문에 기관들의 순위를 정하기 위해서는 각 기준의 척도를 통일할 필요가 있다. 모든 기준의 척도를 통일화하는 과정을 정규화(normalization)한다고 한다. Tofallis(2012)가 제안한 네 가지 정규화 방법을 고려한다.

2.2 자료 정규화

가. 비율 정규화

비율 정규화는 각 기준에서 최고 실적을 낸 기관의 값을 1로 하고 나머지 기관들은 최고 실적을 낸 기관대비 실적 비율로 나타낸다. 비율 정규화의 결과로 나타난 정규화 값 p_{ij} 는 J^+ 와 J^- 에 따라 다음 eq. 2.1과 같이 나타낸다.

$$p_{ij} = \frac{x_{ij}}{\max_i(x_{ij})}, \quad i = 1, 2, \dots, n, \quad j \in J^+ \quad p_{ij} = \frac{\min_i(x_{ij})}{x_{ij}}, \quad i = 1, 2, \dots, n, \quad j \in J^- \quad \text{eq. 2.1}$$

$\max_i(x_{ij})$: 기준 j 의 평가값 중 최댓값, $\min_i(x_{ij})$: 기준 j 의 평가값 중 최솟값

나. 범위 정규화

범위 정규화는 각각 기준에서 최고 실적을 낸 기관과 최저 실적을 낸 기관의 차이인 범위를 이용하여 정규화 하는 방법이다. 각 기관의 실적에서 최저 실적을 뺀 값을 범위로 나누어 정규화 값을 구하므로 그 값은 0에서부터 1사이의 값을 갖게 된다. 범위 정규화의 결과로 나타난 정규화 값 r_{ij} 는 J^+ 와 J^- 에 따라 다음 eq. 2.2와 같이 나타낸다.

$$r_{ij} = \frac{x_{ij} - \min_i(x_{ij})}{\max_i(x_{ij}) - \min_i(x_{ij})}, \quad i = 1, 2, \dots, n, j \in J^+ \quad \text{eq. 2.2}$$

$$r_{ij} = \frac{\max_i(x_{ij}) - x_{ij}}{\max_i(x_{ij}) - \min_i(x_{ij})} = 1 - \frac{x_{ij} - \min_i(x_{ij})}{\max_i(x_{ij}) - \min_i(x_{ij})}, \quad i = 1, 2, \dots, n, j \in J^-$$

다. z-변형 정규화

z-변형 정규화는 각각 기준의 평균값과 표준편차를 이용하여 정규화 하는 방법이다(Park, et al., 2000) 각 기관의 정규화 값은 그 기관의 실적에서 평균을 뺀 값을 표준편차로 나누어 구한다. 그렇게 정규화 된 값의 평균은 0이고 표준편차는 1이 된다. z-변형 정규화의 결과로 나타난 정규화 값 z_{ij} 는 J^+ 와 J^- 에 따라 다음 eq. 2.3과 같이 나타낸다.

$$z_{ij} = \frac{x_{ij} - \mu_j}{\sigma_j}, \quad i = 1, 2, \dots, n, j \in J^+ \quad \text{eq. 2.3}$$

$$z_{ij} = -\frac{x_{ij} - \mu_j}{\sigma_j} = \frac{\mu_j - x_{ij}}{\sigma_j}, \quad i = 1, 2, \dots, n, j \in J^-$$

라. 비중 정규화

비중 정규화는 각 기관의 실적을 모든 기관의 실적을 더한 값으로 나누는 것이다. 비중 정규화의 결과로 나타난 정규화 값 d_{ij} 는 J^+ 와 J^- 에 따라 다음 eq. 2.4와 같이 나타낸다.

$$d_{ij} = \frac{x_{ij}}{\sum_{i=1}^n x_{ij}}, \quad i = 1, 2, \dots, n, j \in J^+ \quad d_{ij} = \frac{\frac{1}{x_{ij}}}{\sum_{i=1}^n \frac{1}{x_{ij}}}, \quad i = 1, 2, \dots, n, j \in J^- \quad \text{eq. 2.4}$$

2.3 통합 방법

가중치를 고려하여 각 기준의 값을 통합하는 방법도 다음의 두 가지를 고려한다. 하나는 가중 산술평균(weighted arithmetic mean)이고 다른 하나는 가중 기하평균(weighted geometric mean)이다.

가. 가중 산술평균

가중 산술평균은 각 기준의 정규화 값에 그 기준의 가중치를 곱하고 이 값을 모든 기준에 대하여 더하는 방법이다.

모든 기관에 대하여 가중 산술평균을 구하고 이 값을 비교하여 그 기관의 순위를 정하는 것이다. 본 연구에서는 비율 정규화, 범위 정규화, z-변형 정규화, 비중 정규화에 대해 각각 다음 eq. 2.5와 같이 가중 산술평균을 구한다.

$$\begin{aligned}
 P_i^a &= \sum_{j=1}^m w_j \cdot p_{ij}, \quad i = 1, 2, \dots, n & R_i &= \sum_{j=1}^m w_j \cdot r_{ij}, \quad i = 1, 2, \dots, n & \text{eq. 2.5} \\
 Z_i &= \sum_{j=1}^m w_j \cdot z_{ij}, \quad i = 1, 2, \dots, n & D_i^a &= \sum_{j=1}^m w_j \cdot d_{ij}, \quad i = 1, 2, \dots, n
 \end{aligned}$$

나. 가중 기하평균

가중 기하평균은 각 기준의 정규화 값에 그 기준의 가중치를 승하여 이 값을 모든 기준에 대하여 곱하는 방법이다. 모든 기관에 대하여 가중 기하평균을 구하고 이 값을 비교하여 그 기관의 순위를 정하는 것이다. 본 연구에서는 비율 정규화와 비중 정규화에 대해 eq. 2.6과 eq. 2.7과 같이 가중 기하평균을 구한다. 범위 및 z-변형 정규화에 대해 가중 기하평균으로 통합하지 못하는 이유는 범위 정규화 값은 0이 나올 수 있고, z-변형 정규화 값은 (-)가 나오므로 기하평균을 구할 수 없기 때문이다.

$$\begin{aligned}
 P_i^g &= \prod_{j=1}^m p_{ij}^{w_j} = \prod_{j=1}^{j \in J^+} \left\{ \frac{x_{ij}}{\max_i(x_{ij})} \right\}^{w_j} \cdot \prod_{j=1}^{j \in J^-} \left\{ \frac{\min_i(x_{ij})}{x_{ij}} \right\}^{w_j} & \text{eq. 2.6} \\
 &= \prod_{j=1}^{j \in J^+} \left\{ \frac{1}{\max_i(x_{ij})} \right\}^{w_j} \cdot \prod_{j=1}^{j \in J^-} \{\min_i(x_{ij})\}^{w_j} \cdot \prod_{j=1}^{j \in J^+} \{x_{ij}\}^{w_j} \cdot \prod_{j=1}^{j \in J^-} \{x_{ij}\}^{-w_j} \\
 &= C_p \cdot \prod_{j=1}^{j \in J^+} \{x_{ij}\}^{w_j} \cdot \prod_{j=1}^{j \in J^-} \{x_{ij}\}^{-w_j}
 \end{aligned}$$

$$\begin{aligned}
 D_i^g &= \prod_{j=1}^m d_{ij}^{w_j} = \prod_{j=1}^{j \in J^+} \left\{ \frac{x_{ij}}{\sum_{i=1}^n x_{ij}} \right\}^{w_j} \cdot \prod_{j=1}^{j \in J^-} \left\{ \frac{1}{\sum_{i=1}^n \frac{1}{x_{ij}}} \right\}^{w_j} & \text{eq. 2.7} \\
 &= \prod_{j=1}^{j \in J^+} \left\{ \frac{1}{\sum_{i=1}^n x_{ij}} \right\}^{w_j} \cdot \prod_{j=1}^{j \in J^-} \left\{ \frac{1}{\sum_{i=1}^n \frac{1}{x_{ij}}} \right\}^{w_j} \cdot \prod_{j=1}^{j \in J^+} \{x_{ij}\}^{w_j} \cdot \prod_{j=1}^{j \in J^-} \{x_{ij}\}^{-w_j} \\
 &= C_d \cdot \prod_{j=1}^{j \in J^+} \{x_{ij}\}^{w_j} \cdot \prod_{j=1}^{j \in J^-} \{x_{ij}\}^{-w_j}
 \end{aligned}$$

위의 식에서 보듯이 P_i^g 와 D_i^g 는 $\prod_{j=1}^{j \in J^+} \{x_{ij}\}^{w_j} \cdot \prod_{j=1}^{j \in J^-} \{x_{ij}\}^{-w_j}$ 에 각각 상수 C_p 와 C_d 를 곱한 것과 같은 결과이다. 그러므로 다음 eq. 2.8을 이용하여 순위를 구하는 것과 같은 결과가 나타난다. 즉, 비율 정규화와 비중 정규화한 값의 가중 기하평균으로 순위를 구하는 것은 정규화 과정 없이 원래 평가된 값 x_{ij} 로 가중 기하평균을 계산하여 순위를 구하는 것과 같은 결과이다.

$$M_i = \prod_{j=1}^{j \in J^+} \{x_{ij}\}^{w_j} \cdot \prod_{j=1}^{j \in J^-} \{x_{ij}\}^{-w_j} \quad \text{eq. 2.8}$$

P_i^a , R_i , Z_i , D_i^a , 그리고 M_i 를 이용하여 순위를 구하는 것을 각각 모형 1, 2, 3, 4, 5로 정의한다. 모형 1부터 4까지는 정규화 과정을 거친 값으로 가중 산술평균을 구하여 순위를 구하는 반면, 모형 5는 정규화 과정 없이 원래의 값으로 가중 기하평균을 구해 순위를 구한다.

상하이 자오퉁 대학에서 출판하는 발행물인 ARWU(academic ranking of world university)는 모형 1을 사용하고 있다. 모형 3은 CUG, UK's THES(times higher education supplement), 중앙일보 대학종합평가, 교육부의 대학재정 지원 제한 대학 선정 등 많은 기관에서 사용하는 방법이다. UN(united nations)은 HDI(human development index)를 계산할 때 모형 5를 이용한다.

3. 자료 수집 및 분석 방법

3.1 자료 수집

영국 CUG(complete university guide)에서는 매년 영국의 대학순위를 발표하고 있다. 본 연구에서는 CUG에서 발표한 'University League Table 2014'를 이용한다. 이 표에는 124개 대학의 순위가 발표되어 있다. 대학 순위를 결정하기 위해 사용한 9개 기준의 출처, 가중치, 정의는 다음의 <Table 3.1>과 같다.

자료의 출처는 HESA(higher education statistics agency), 학생으로부터의 만족도 설문, 연구 평가 등이다. 영국의 대학 졸업생은 first, upper second, lower second, third, ordinary class의 5단계로 구분하여 학위가 주어지는데, 'good honors'는 전체 졸업생 중 first와 upper second로 졸업한 학생의 비율을 의미한다. 'student:staff ratio'를 제외한 나머지 8개 기준은 그 값이 클수록 좋은 것이다. <Table 3.1>의 기준 중 밑의 2개와 'graduate prospects'는 비율(%)로 측정된 값으로 최고 가능한 값은 100이다. 'student satisfaction'과 'research assessment'는 만점이 각각 5와 4이다. 'entry standards'는 213에서 608, 'student:staff ratio'는 10에서 32.9, 'academic services spending'은 132에서 5,448 그리고 'facilities spending'은 67에서 1,071까지의 값으로 분포되어 있다.

CUG는 이렇게 수집된 값을 z-변형 정규화해 가중치를 고려한 가중 산술평균을 구한 다음 이 값으로 대학순위를 결정한다. 전체 점수(overall score)는 1위 대학에 1000점을 주고 나머지 대학은 1위 대학에 비해 전체 점수를 부여한다. 몇몇 기준('entry standards', 'student satisfaction', 'student:staff ratio', 'good honors', 'graduate prospects')들은 학교의 전공 분야 구성(subject mix)에 따라 조정되지만 구체적으로 어떻게 조정되는지는 알려지지 않았다. 즉, 'entry standards'는 의과대학이 있는 대학의 값이 의과대학이 없는 대학의 값에 비해 높으므로 이를 조정한다.

CUG에서 발표한 124개 대학 자료 중 9개 기준의 값이 모두 조사되지 않고 1개 기준이라도 누락된 대학은 2개가 있다. 이렇게 값이 누락된 2개 대학을 제외하고 나머지 122개 대학을 대상으로 분석한다. CUG에서 발표된 각 대학의 순위에서 2개 대학을 제외한 후 순위를 1위에서부터 122위까지 다시 정렬하였고 이 순위가 i 인 대학을 A_i 로 표시한다.

Table 3.1 Definition from the Complete University Guide

Criteria	Source	Weighting	Definition
entry standards	HESA data for 2011-12	1	The average score for all students at the university.
student satisfaction	The National Student Survey 2012	1.5	The average satisfaction score for all questions except the three about learning resources.
research assessment	The 2008 Research Assessment	1.5	The research assessment and intensity measures were combined in a ratio of 2:1.
graduate prospects	HESA data for 2010-11	1	The number of graduates who take up employment or further study divided by the total number of graduates as a percentage.
student:staff ratio	HESA data for 2011-12	1	The number of students divided by the number of staff.
academic services spending	HESA data 2009-10, 2010-11 & 2011-12.	1	A university's expenditure on library and computing facilities, museums, galleries and observatories was divided by the number of full-time equivalent students.
facilities spending	HESA data 2009-10, 2010-11 & 2011-12.	1	A university's expenditure on student facilities was divided by the number of full-time equivalent students.
good honours	HESA data for 2011-12	1	The number of graduates with first or upper second class degrees was divided by the total number of graduates.
completion	HESA data for 2011-12	1	The percentage of students who were expected to complete their course or transfer to another institution.

3.2 분석 방법

CUG에서 발표된 9개 기준에 대한 평가값을 이용하여 모형 1부터 5까지 각각 대학순위를 결정한다. 이 때 각 기준에 대한 가중치는 모두 같다고 가정한다. CUG는 ‘student satisfaction’과 ‘research assessment’의 가중치를 다른 기준의 1.5배로 주었으나 본 연구는 가중치의 효과를 분석하려는 목적이 아니기 때문에 같은 가중치를 주어도 큰 문제가 없을 것으로 판단된다. 가중치를 결정하는 방법으로 AHP, 요인분석 등이 사용될 수 있다(Yoo, 1994; Leung & Chin, 2004; Kim & Kim, 2013; OECD, 2008).

모형 k 로부터 구한 대학 i 의 순위는 Y_i^k ($i = 1, 2, \dots, 122$, $k = 1, 2, \dots, 5$)로 나타낸다.

CUG는 몇 개의 기준에 대해 각 대학별 전공 분야 구성에 따라 조정하고, 가중치도 다르게 부여하였기 때문에 여기에서 구한 순위는 본 연구의 모형 3에서 구한 순위와 일치하지 않는다. 5가지 모형으로 구한 각 대학순위(Y_i^k)는 다르게 나타낸다. 그러나 진정한 대학순위를 알 수 없으므로 5가지 모형으로 구한 대학순위의 정확성을 평가할 수 있는 판단 기준은 없다.

본 연구에서 고려하고 있는 5개 모형은 각각 그 의미가 있는 모형이다. 그러므로 5개 모형으로 구한 대학 i 의 순위인 $Y_i^1, Y_i^2, Y_i^3, Y_i^4, Y_i^5$ 의 중앙값(\widetilde{M}_i)을 그 대학의 진정한 순위라고 가정하여 모형의 정확성을 평가하려고 한다. 이러한 가정 하에 Y_i^k 와 \widetilde{M}_i 의 절대편차인 V_i^k 는 오차의 정도를 나타낼 수 있다. V_i^k 를 대학 i 의 모형 k 에 따른 변동성으로 측정하는 식은 eq. 3.1과 같다.

$$V_i^k = |Y_i^k - \widetilde{M}_i|, i = 1, 2, \dots, 122, k = 1, 2, \dots, 5 \quad \text{eq. 3.1}$$

V_i^k 는 모형 k 에 따른 대학 i 의 변동성만을 측정한 것이다. 모형 k 의 정확성을 측정하려면 전체 대학에 대한 평균 변동성(\overline{V}^k)을 구해야 한다. \overline{V}^k 를 구하는 식은 eq. 3.2와 같다.

$$\overline{V}^k = \frac{\sum_{i=1}^{122} V_i^k}{122}, k = 1, 2, \dots, 5 \quad \text{eq. 3.2}$$

\overline{V}^k 의 값이 작을수록 모형 k 로부터 구한 대학순위가 진정한 대학순위와의 차이가 적다는 것을 의미하므로, \overline{V}^k 의 값이 최소가 되는 모형 k 가 가장 바람직하다고 판단할 수 있다.

이상점(outlier)이 존재하는 경우 이러한 이상점이 모형에 따른 순위 변동성에 어떤 영향을 미치는지 분석하기 위해 이상점의 존재 여부를 파악한다. 이상점을 찾는 방법도 다양하게 있지만 본 연구에서는 박스-도표(box-plot)를 그려 이상점을 찾으려고 한다.

이상점이 있는 대학을 제외한 나머지 대학에 대해 5가지 모형으로 다시 순위를 정하고 모형 k 의 평균 변동성(\overline{V}_o^k)을 계산한다. 두 가지 경우의 평균 변동성(\overline{V}^k 와 \overline{V}_o^k)을 비교함으로써 이상점이 대학순위의 변동성에 미치는 영향을 파악한다.

모든 대학이 평가 받았을 때와 일부 대학만이 평가를 받았을 때 상대적인 대학순위가 바뀌는 현상을 순위 역전이라 한다. 즉, 모든 대학이 평가 받았을 때 대학 A가 대학 B보다 순위가 높았지만 몇몇 대학이 평가에서 빠졌을 때, 대학 A가 대학 B보다 순위가 낮아진다면 이는 순위가 역전된 것이다. 그러므로 순위 역전이 많이 발생하는 모형은 적합하지 않다고 판단할 수 있다. 전체 122개 대학의 순위와 이상점이 있는 대학을 제외한 대학의 순위를 비교하여 대학 간 모형에 따른 순위 역전이 얼마나 일어났는지 계산한다. 모형별 순위 역전이 일어난 비율을 측정함으로써 모형의 적합성을 평가하고자 한다.

4. 분석 결과

4.1 순위의 변동성

5개 모형의 대학순위 Y_i^k 로부터 구한 각 대학의 변동성 V_i^k 와 모형 k 의 평균 변동성 \overline{V}^k 에 대한 기술 통계량은 다음 <Table 4.1>과 같다.

Table 4.1 Volatility of the Five Models

V_i^k \ Model	Model 1	Model 2	Model 3	Model 4	Model 5
$Max_i(V_i^k)$	8	33	31	29	33
Number of $V_i^k = 0$	74	35	29	32	30
\overline{V}^k	1.0	5.4	4.1	4.2	3.8

<Table 4.1>에서 보듯이 변동성의 최댓값은 모형 1이 8로 가장 작고 모형 2와 5가 33으로 가장 크다. 변동성이 0인 대학의 수는 모형 1이 74개로 가장 많고 모형 3은 29개로 가장 적다. 그리고 평균 변동성은 모형 1이 1.0으로 가장 작으며 모형 2가 5.4로 가장 크다. 모형 3, 4, 5는 각각 4.1, 4.2, 3.8로 비슷한 수준을 나타냈다. 이상의 결과로 판단할 때 모형 1이 변동성을 평가하는 모든 항목에서 우수한 반면 모형 2는 모든 항목에서 좋지 않은 것으로 나타났다. 모형 3, 4, 5는 모든 항목에서 비슷한 결과를 보여주고 있다.

모형 2와 3의 순위에 대한 상관계수는 0.996으로 가장 높다. 다음으로 높은 상관계수는 모형 4와 5의 0.991이다. 모형 1과 모형 4 및 5의 상관계수는 0.98로 그 다음 순으로 높다. 이러한 사실을 근거로 판단했을 때 모형 1, 4, 5가 유사한 순위를 나타내는 그룹으로 구성되고, 나머지 모형 2와 3이 이와는 다른 그룹으로 구성된 것으로 판단된다. 모형 1, 4, 5로 구한 정규화 값은 각 기준의 평가값에 대한 산포도를 고려하지 않는다. 반면에 모형 2, 3은 각 기준의 평가값에 대한 범위와 표준편차를 나누어 정규화 값을 구하는 것으로 산포도를 함께 고려한다. <Table 4.1>에서 모형 1, 4, 5에서 구한 순위가 중앙값이 된 경우는 136개인 반면 모형 2와 3에서는 64개로 작았다. 특히 모형 1에서 구한 순위가 중앙값이 된 경우가 74개로 많았기 때문에 모형 1의 평균 변동성이 가장 작게 나타난 것으로 판단된다.

대학순위의 발표는 대학당국, 학생, 기업 등에 미치는 파급적 효과가 상당히 크다. 현재 많이 사용되고 있는 대학 순위 결정 모형들의 순위 변동성이 크다면 바람직하지 않을 것이다. 그러므로 모형에 따른 순위의 변동성을 초래하게 된 원인이 무엇인지를 규명하는 것은 매우 중요한 일이다. 이를 위해 대학의 모형별 순위 변동성이 큰 대학들은 어떤 특성을 갖고 있는지 먼저 파악해 볼 필요가 있다.

각 대학의 순위 변동성의 평균(\overline{V}_i)은 다음과 같이 eq. 4.1로 구할 수 있다.

$$\overline{V}_i = \frac{\sum_{k=1}^5 V_i^k}{5}, i = 1, 2, \dots, 122 \quad \text{eq. 4.1}$$

\overline{V}_i 의 값이 큰 12개 대학의 자료 특성을 알아보기 위해 다음과 같은 값을 구한다. 각 대학의 9개 기준에 대한 각각의 순위를 구한다. 9개 기준의 순위 중 가장 좋은 순위와 가장 나쁜 순위를 찾고, 가장 나쁜 순위와 가장 높은 순위의 차이인 순위의 범위를 구한다. 또한 기준 9개 순위의 평균과 중앙값 그리고 표준편차를 구한다. 변동성이 큰 대학의 특성을 요약하면 다음 <Table 4.2>와 같다.

Table 4.2 Top 12 Universities with Higher Volatility

i	University Name(A_i)	\overline{V}_i	Rank Median	Rank Average	Rank SD	Rank Range
92	Middlesex	17.2	106	85.9	34.5	106
74	Arts University at Bournemouth	16.2	80	76.8	39.8	119
103	Buckinghamshire New	15.8	109	95.4	35.6	119
106	Bedfordshire	13.2	102	90.6	33.3	116
50	Goldsmiths, University of London	12.6	68	68.7	30.5	88
85	University for the Creative Arts	12.2	85	78.9	32.5	107
75	Chichester	10.6	82	73.0	31.7	107
34	St George's University of London	10.4	28	50.1	46.4	121
78	Bath Spa	9.6	91	82.6	25.2	70
88	Harper Adams	9.6	78	76.6	29.9	99
51	Royal Agricultural University	8.0	82	70.4	41.6	119
91	Central Lancashire	8.0	87	79.0	24.9	76

<Table 4.2>에서 St George's University of London은 'graduate prospects'와 'academic services spending'에서는 1위를 한 반면에 'student satisfaction'에서는 최하위인 122위를 차지하였다. 5가지 모형에서도 각각 3, 11, 33, 14, 33위를 차지하여 순위 평균 50위보다 높은 순위를 차지하였다. 그 이유는 'academic services spending'에서 5,448로 전체 대학의 평균 1,098에 비해 월등히 높게 지출했기 때문이다. 그 밖의 Buckinghamshire New와 Royal Agricultural University도 'facilities spending'에서 다른 대학에 비해 월등히 많은 지출을 하였다. 자료의 이상점이 존재할 때 대학순위의 변동성이 높아질 가능성이 높아질 것으로 예상된다.

4.2 이상점의 영향

이상점의 존재 여부를 판단 할 수 있는 여러 가지 통계적 방법이 있다. z -변형을 하여 그 값의 절대값이 2보다 큰 경우에 이상점이라고 판단하는 근거는 자료가 정규분포를 따른다고 할 때 z 값이 -2에서 +2 사이에 존재할 확률이 95.4%라는 사실이다. 본 연구에서는 박스도표를 그려 이상점을 찾으려고 한다. 박스 도표는 제2사분위수(50%에 해당하는 값, Q_2)을 중심선으로 제1사분위수(25%에 해당하는 값, Q_1)와 제3사분위수(75%에 해당하는 값, Q_3)로 박스를 그린다. 수염은 제1사분위수와 하한값($Q_1 - 1.5*(Q_3 - Q_1)$)을 연결하고 제3사분위수와 상한값($Q_3 + 1.5*(Q_3 - Q_1)$)을 연결하여 그린다. SPSS 18을 이용하여 9가지 기준에 대한 박스-도표를 그려본 결과 다음 <Figure 4.1>과 같다.

<Figure 4.1>에서 상한값과 하한값을 벗어나는 값은 'o'와 그 옆에 i 를 표시하였다. 이러한 이상점은 보통 이상점이라고 한다. '*'와 그 옆에 i 를 표시한 것은 그 값이 $Q_1 - 3*(Q_3 - Q_1)$ 보다 작거나 $Q_3 + 3*(Q_3 - Q_1)$ 보다 큰 것으로 극단적인(extreme) 이상점이라고 한다.

'academic services spending'에서 보통 이상점을 갖는 4개 대학(Middlesex, University College London, London School of Economics, Cambridge)과 극단적 이상점을 갖는 3개 대학(Imperial College London, Oxford, St George's University of London)이 존재한다. 'facilities spending'에서는 보통 이상점을 갖는 2개 대학(Northampton, Durham)과 극단적 이상점을 갖는 3개 대학(Hertfordshire, Royal Agricultural University, Buckinghamshire New)이 존재한다. 그리고 'entry standards'에서는 Cambridge가, 'student satisfaction'에서는

St George's University of London이, 'student:staff ratio'에서는 Bishop Grosseteste가, 'degree completion'에서는 East London이 이상점을 갖는다.

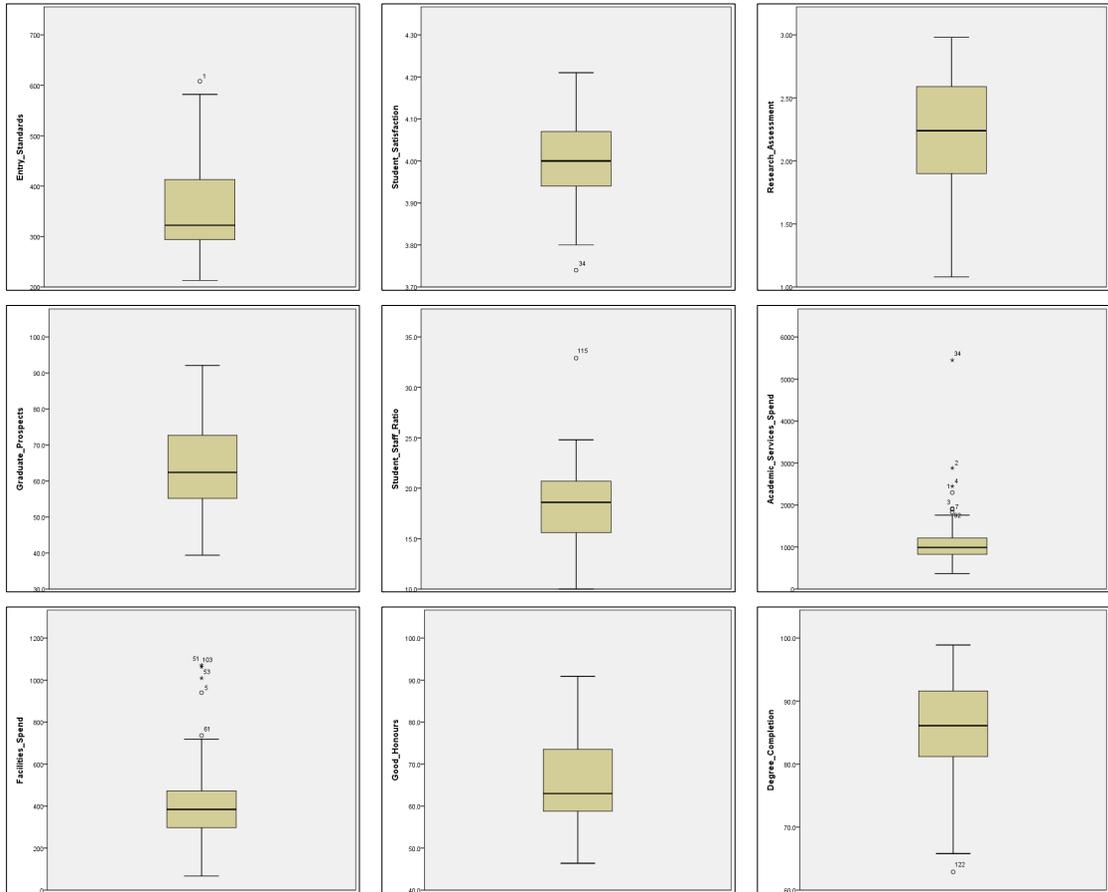


Figure 4.1 Box-Plots for 9 Criteria

'student satisfaction'과 'degree completion'은 하한값을 벗어난 반면 나머지 이상점은 상한값을 초과하였다. 'student:staff ratio'는 상한값을 초과하였지만 이 기준은 클수록 좋지 않은 것이다. 모두 16개의 이상점이 발견되었고, 그 중 Cambridge와 St George's University가 2개의 이상점을 갖기 때문에 하나의 기준에서라도 이상점을 갖고 있는 대학은 모두 14개이다.

122개 대학 중에 이상점이 있는 14개 대학을 제외하고 108개 대학으로 모형에 따른 순위의 변동성을 측정해 본 결과는 <Table 4.3>에 나타나 있다.

Table 4.3 Volatility of the Five Models after Removing Outlier

V_i^k \ Model	Model 1	Model 2	Model 3	Model 4	Model 5
$Max_i(V_i^k)$	4	29	28	13	18
Number of $V_i^k = 0$	77	20	18	40	33
\overline{V}_o^k	0.5	4.8	4.6	1.8	1.8

<Table 4.3>을 <Table 4.1>과 비교해 보았을 때 평균 변동성(\overline{V}_o^k)은 모형 3을 제외하고는 모두 감소하였다. 변동성의 최댓값($Max_i(V_i^k)$)도 5개 모형 모두에서 감소하였다. 이 값은 대학의 수가 122개에서 108개로 적어진 영향도 있겠지만 모형 4와 모형 5는 각각 16위와 15위 차이로 감소하였다. 각 대학별 모형의 변동성 평균(\overline{V}_i)도 8.0을 넘는 대학의 수가 4개로 감소하였다. 이상의 결과를 보았을 때 이상점이 대학별 모형의 변동성에 많은 영향을 미친다는 사실을 확인할 수 있었다.

4.3 순위 역전

이상점이 있는 대학을 제외하고 5개 모형으로 대학순위를 결정하였을 때, 대학순위에 어떤 영향을 미칠지 분석해 볼 필요가 있다. 이상점을 갖고 있는 대학을 제외한 108개 대학의 순위가 전체 122개 대학순위와 어떻게 다른지를 분석하기 위해 전체 122개 대학순위를 108개 대학만으로 한정지어 다시 조정하였다. 즉 1위부터 122위까지의 순위를 14개 대학을 제외하여 다시 1위부터 108위로 조정하였다. 이렇게 조정된 순위(YA_i^k)와 14개 대학을 제외한 108개 대학의 순위(YO_i^k)를 모형별로 비교한 결과는 <Table 4.4>에 있다.

Table 4.4 Rank Change and Reversal of the Five Models

Rank Change \ Model	Model 1	Model 2	Model 3	Model 4	Model 5
# of No Change	20	21	31	75	108
# of Down Change	42	41	40	16	0
# of Up Change	46	46	37	17	0
Max. Rank Change	13	10	9	2	0
Min. Rank Change	-16	-10	-6	-2	0
# of Rank Reversal	204	168	107	20	0
Proportion of Rank Reversal	3.53%	2.91%	1.85%	0.35%	0.00%

순위의 변화는 $YO_i^k - YA_i^k$ 으로 계산하였다. 이 값이 음수이면 이상점을 제거한 후의 순위가 올라간 것이고, 양수이면 순위가 내려간 것이고, 0이면 순위의 변화가 없는 것이다. 이 차이가 최대로 나타난 대학은 순위가 가장 많이 떨어진 대학이고, 최소인 대학은 순위가 가장 많이 올라간 대학을 의미한다. 대학 i, j 를 비교했을 때 $(YA_i^k - YA_j^k) \times (YO_i^k - YO_j^k) < 0$ 의 조건이 만족되면 순위 역전이 일어난 것이다. 두 개의 대학을 서로 비교한 모든 경우의 수는 ${}_{118}C_2$ 이고 그 중 순위 역전이 몇 번 발생하였는지 그리고 그 비율은 얼마인지 측정하였다.

모형 5는 몇 개의 대학이 제외되어도 순위를 계산할 때 사용하는 원래 값이 변하지 않으므로 순위의 변화나 순위 역전이 일어나지 않는다. 모형 4는 순위 변화가 없는 대학의 수가 75개이고 순위가 역전된 비율도 0.35%로 모형 5를 제외한 다른 모형과 비교해 순위가 변화된 대학의 수와 순위가 역전된 비율이 낮았다. 모형 1은 순위가 내려간 대학의 수가 42개로 가장 많았고, 순위가 올라간 대학의 수도 46개로 가장 많았으며 순위가 역전된 비율도 3.53%로 가장 높았다. 순위가 제일 많이 내려간 대학은 13위가 떨어졌으며, 제일 많이 올라간 대학은 16위가 올라갔다. 모형 2도 모형 1과 비슷한 수의 순위 변화가 일어났고 순위 역전도 비슷한 비율로 발생했으나 순위 변화의 폭이 모형 1보다 작았다. 모형 3은 순위가 변하지 않은 대학의 수나 변화의 폭 그리고 순위 역전 비율이 5개 모형의 중간 정도로 나타났다.

모형 1에서 순위 변화가 가장 많이 발생한 두 대학의 모형별 순위 변화를 살펴보면 <Table 4.5>와 같다.

Table 4.5 Rank Change of the Five Models for Two Universities

University		Model	Model 1	Model 2	Model 3	Model 4	Model 5
Goldsmiths, University of London	YO_{Gold}^k		62	53	55	69	80
	YA_{Gold}^k		49	45	49	69	80
	$YO_{Gold}^k - YA_{Gold}^k$		13	8	6	0	0
Falmouth	YO_{Falm}^k		47	63	62	47	48
	YA_{Falm}^k		63	69	67	47	48
	$YO_{Falm}^k - YA_{Falm}^k$		-16	-6	-5	0	0

Table 4.6 $p_{ij}, r_{ij}, z_{ij}, d_{ij}, x_{ij}$ before Removing Outliers

Criteria	Goldsmiths, University of London					Falmouth				
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 1	Model 2	Model 3	Model 4	Model 5
entry standards	0.5888	0.3671	0.0833	0.0084	358	0.4688	0.1823	-0.7905	0.0067	285
student satisfaction	0.9549	0.5957	0.1939	0.0082	4.02	0.9287	0.3617	-0.9230	0.0080	3.91
research assessment	0.8658	0.7895	0.8099	0.0095	2.58	0.7383	0.5895	-0.0643	0.0081	2.2
graduate prospects	0.5885	0.2808	-0.9037	0.0069	54.2	0.5917	0.2865	-0.8764	0.0070	54.5
student:staff Ratio	0.5208	0.5983	-0.2425	0.0075	19.2	0.4184	0.3930	-1.5301	0.0060	23.9
academic services spending	0.1369	0.0739	-0.6365	0.0056	746	0.2768	0.2240	0.7427	0.0113	1,508
facilities spending	0.1643	0.1086	-1.3123	0.0035	176	0.3548	0.3118	-0.1578	0.0076	380
good honours	0.8086	0.6090	0.7473	0.0092	73.5	0.7085	0.4045	-0.1385	0.0080	64.4
completion	0.8554	0.6028	-0.1985	0.0081	84.6	0.8888	0.6944	0.2640	0.0084	87.9
aggregation score	0.6093	0.4473	-0.1621	0.0074	27.348	0.5972	0.3831	-0.3860	0.0079	29.728

모형 1에서 이상점이 있는 대학을 제외하기 전 Goldsmiths, University of London(앞으로 Goldsmiths로 지칭함)은 49위에서 이상점을 제외한 후의 순위는 62위로 13위가 떨어졌다. 반면에 Falmouth는 63위에서 47위로 16위가 상승하였다. 두 대학은 모형 2와 3에서도 마찬가지로 순위 역전이 발생하였다. 이상점이 있는 대학을 제외하기 전

두 대학의 모형별 정규화 값, 평가값 그리고 통합한 값(모형 1부터 4까지는 산술평균, 모형 5는 ‘student: staff ratio’에만 역수를 취한 후 기하평균)을 요약한 결과는 <Table 4.6>에 나타나 있다.

두 대학에 대해 9개 기준의 값을 비교하였을 때 ‘entry standards’, ‘student satisfaction’, ‘research assessment’, ‘student:staff ratio’ 그리고 ‘good honours’는 Goldsmiths가 Falmouth에 비해 우수하였고, 나머지 4개 기준은 Falmouth가 우수하였다. 이러한 차이는 모형 1부터 4까지의 정규화 값에서도 마찬가지로 나타났지만 각 모형에 따라 그 차이의 정도는 다르게 나타났다. 모형 1부터 3까지는 이러한 결과가 반영되어 Goldsmiths가 Falmouth에 비해 높은 순위를 나타냈다. 그러나 모형 4에서는 Falmouth가 ‘academic services spending’과 ‘facilities spending’에서 큰 격차로 우위를 점했기 때문에 나머지 다른 기준에서의 열세를 만회할 수 있어 Falmouth의 순위가 Goldsmiths의 순위보다 높았다. 모형 5에서도 Falmouth가 ‘academic services spending’과 ‘facilities spending’에서 Goldsmiths 보다 값이 2배 이상 큰 반면에 가장 열세에 있는 기준도 80%는 차지한 것이 반영되어 Falmouth의 순위가 Goldsmiths 보다 높게 나타났다.

이상점이 있는 대학을 제외한 후의 결과도 같은 방법으로 <Table 4.7>에 나타났다.

Table 4.7 p_{ij} , r_{ij} , z_{ij} , d_{ij} , x_{ij} after Removing Outliers

Criteria	Goldsmiths, University of London					Falmouth				
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 1	Model 2	Model 3	Model 4	Model 5
entry standards	0.6871	0.4708	0.1873	0.0096	358	0.5470	0.2338	-0.8477	0.0077	285
student satisfaction	0.9549	0.5366	0.1975	0.0093	4.02	0.9287	0.2683	-0.9996	0.0090	3.91
research assessment	0.9149	0.8621	0.8924	0.0108	2.58	0.7801	0.6437	-0.0376	0.0092	2.2
graduate prospects	0.6237	0.2635	-0.9189	0.0080	54.2	0.6272	0.2703	-0.8879	0.0080	54.5
student:staff Ratio	0.6094	0.3902	-0.2828	0.0086	19.2	0.4895	0.0081	-1.8395	0.0069	23.9
academic services spending	0.4243	0.2575	-0.9655	0.0068	746	0.8578	0.8166	1.8236	0.0138	1,508
facilities spending	0.2448	0.1185	-1.6536	0.0042	176	0.5285	0.4497	-0.0312	0.0092	380
good honours	0.8277	0.6392	0.8643	0.0104	73.5	0.7252	0.4245	-0.1225	0.0091	64.4
completion	0.8686	0.5949	-0.1579	0.0091	84.6	0.9025	0.6994	0.3465	0.0095	87.9
aggregation score	0.6839	0.4593	-0.2041	0.0085	27.348	0.7096	0.4238	-0.2884	0.0092	29.728

<Table 4.7>에서 모형 1의 경우 ‘academic services spending’과 ‘facilities spending’의 이상점이 제거된 효과가 크게 나타났다. Falmouth가 Goldsmiths에 비해 ‘academic services spending’에 762만큼 많이 지출했다. Falmouth는 이상점을 제거하기 전에는 762를 최댓값 5,448로 나눈 0.1399의 차이로 정규값이 컸다. 그러나 이상점을 제거하고 난 후의 최댓값은 1,758로 줄어 0.4334의 차이만큼 정규값이 더 커졌다. Goldsmiths가 Falmouth에 비해 ‘entry standards’에서 73이 컸다. Goldsmiths는 이상점을 제거하기 전에는 73을 최댓값 608로 나눈 0.1201의 차이로 정규값이 크게 나타났지만, 이상점을 제거하고 난 후의 최댓값은 521로 줄어 0.1401의 차이만큼 정규값이 더 커졌다. ‘academic services spending’은 큰 폭(0.4334)으로 증가한 반면 ‘entry standards’는 소폭(0.1401)으로 증가한 것이 반영되어 순위 역전이 발생하였다. 모형 2와 3에서도 비슷한 논리로 순위 역전 현상의 원인을 설명할 수 있다. 모형 4에서는 이상점의 제거로 인하여 정규화 값이 증가하였으나 이러한 영향은 이미 반영되었기 때문에 순위 역전이 발생하지 않았다.

이상의 결과로 판단하였을 때, 순위 역전만을 고려한다면 모형 4, 5가 모형 1, 2, 3보다 안정적인 대학순위 결정 모형이 된다.

5. 결론

본 연구는 대학순위를 결정하는 모형 5개를 설정하였다. 5개 모형을 이용하여 순위를 구하기 위한 자료는 최근에 발표된 CUG의 ‘University League Table 2014’이다. 결측 자료가 있는 두 개 대학을 제외하고 122개 대학을 분석 대상으로 하였다.

전체 122개 대학을 대상으로 모형별 평균 변동성을 구한 결과 모형 1이 변동성이 가장 작은 것으로 나타났다. 변동성을 구하기 위해 5개 모형의 중앙값을 이용하였는데 이는 중앙값이 가장 진실 된 대학순위를 나타낼 것이라는 가정 하에서 모형 1이 적합하다는 결론을 내릴 수 있다.

자료에는 이상점이 존재하는데 박스-도표를 이용하여 이상점을 찾아냈고 이상점이 있는 14개 대학을 제외하고 108개 대학을 대상으로 다시 평균 변동성을 구하였다. 마찬가지로 모형 1의 평균 변동성이 가장 작은 것으로 나타났다. 또한 대부분의 모형에서 이상점을 제거한 후에 평균 변동성이 작아진다는 사실도 발견할 수 있었다. 그러므로 자료에 포함되어 있는 이상점을 찾아내고 이를 처리할 수 있는 방법을 모색할 필요가 있다.

전체 122개 대학을 대상으로 한 결과와 이상점이 있는 대학을 제외하고 108개 대학을 대상으로 한 결과의 순위 역전을 모형별로 파악하였다. 모형 5는 순위 역전이 일어나지 않는 모형이다. 모형 1의 평균 변동성은 두 가지 경우 모두 가장 작았으나 순위 역전은 가장 많이 발생하였다. 대학순위 결정에서 많이 사용되고 있는 모형 3은 평균 변동성이나 순위 역전 기준으로 평가하였을 때 모두 우수하지 못한 것으로 나타났다.

지금까지의 연구결과로 판단하였을 때, 변동성도 작고 순위 역전도 적게 발생하는 대학순위 결정 모형은 존재하지 않는 것으로 나타났다. 그러나 이상점이 변동성에 영향을 미치고 있다는 사실은 확인 할 수 있었다.

대학순위는 평가주체에 따라 서로 상이한 결과를 발표하고 있다. 그 이유는 평가기준을 무엇으로 할 것인가? 평가 기준의 가중치는 얼마를 부여할 것인가? 자료의 정규화와 통합 방법은 무엇으로 할 것인가?에 대한 명확한 정답이 없기 때문이다. 본 연구는 자료의 정규화와 통합 방법에 따른 순위 변동성을 파악하였다. 본 연구에서 발견된 연구 결과는 자료에 이상점이 있는 경우와 일부 대학이 평가에 빠지는 경우에도 순위에 크게 영향을 받지 않을 강건한(robust) 그리고 이해관계 집단 모두가 신뢰할(reliable) 모형의 개발에 도움이 될 것으로 기대한다.

REFERENCES

- Complete University Guide 2014. <http://www.thecompleteuniversityguide.co.uk/league-tables/rankings>.
- Filinov, N. B., and Ruchkina, S. 2002. “Ranking of Higher Education Institutions in Russia: Some Methodological Problems.” *Higher Education in Europe* 27(4):407-421.
- Kim, H. J., and Kim, S. W. 2013. “En Empirical Study of Railroad Technology Improvement Using AHP and QFD.” *Journal of the Korean Society for Quality Management* 41(2):301-322.
- Leung, J. P. F., and Chin, K. S. 2004. “An AHP Based Study on Critical Success Factors for the Supply Chain management in Hong Kong Manufacturing Industry.” *The Asian Journal on Quality* 5(2):132-140.
- Luckman, R., Krajnc, D., and Glavič, P. 2010. “University Ranking Using Research, Educational and Environmental Indicators.” *Journal of Cleaner Production* 18:619-628.
- OECD 2008. *Handbook on Constructing Composite Indicators. Methodology and User Guide*. Paris: OECD.

- Park, S. H., Kim, C. H., and Park, J. O. 2000. "A Study on a Standardized System of Selected Subjects for College Entrance Examination." *Journal of the Korean Society for Quality Management* 28(3):124-132.
- Saisana, M., D' Hombres, B., and Saltelli, A. 2011. "Rickety Numbers: Volatility of University Rankings and Policy Implications." *Research Policy* 40:165-177.
- Tofallis, C. 2012. "A Different Approach to University Rankings." *Higher Education* 63:1-18.
- Yoo, H. J. 1994. "A Study on the Success factors of TQM -Through the AHP Analysis of Japanese Companies-." *Journal of the Korean Society for Quality Management* 22(1):33-53.