Principal Component Analysis: Development of Educational Index Equation for Secondary Schools

Habiba Khan[a,*], Syed Feroz Shah[a], Faheemullah Shaikh[a], Asif Ali Shaikh[a], Ariba Shan[a]

[a]Mehran University of Engineering & Technology, Jamshoro, Pakistan.

Abstract

This article examines the contribution of selected educational indicators by implementing principal component analysis (PCA). This methodology helps us to generate two secondary school educational index (SSEI) equations of Pakistan for the period 2000-2016 using fourteen indicators. This study further initiates to analyze the relationship of enrolment and repeaters with our indicators. Thus, this article has made two key contributions: (1) the construction of the Sindh secondary school educational index equation (SSEI), and (2) the participation of selected educational indicators, which educational policy-makers can use for strengthening educational quality. Results indicate that availability of washroom is the highest weighted indicator in rural female, and rural male with a percent of 9.381 and 10.017, whereas the lowest weighted indicator is repeaters in rural female (0.759%) and teacher-institution ratio in rural male with 2.052 percent.

Keywords: *Principal Component Analysis; Secondary schools; Educational development.*

1. Introduction

Education is a fundamental principle for development and growth. According to the Universal Declaration of Human Rights and the United Nations Convention on the Rights of the Child, access to education is a fundamental human right and a beneficial approach towards development [1]. One of the important reasons for downfall may be the poor quality of education since it provides the base for both social and economic development. To enhance individual personality or to develop a country, education is regarded as a critical factor. Asian development bank issued a report which indicates that: In 2014-2015, 80% and 32% of Pakistan rural and urban women are illiterate whereas, the urban and rural male illiteracy rate was 18% and 46%. Also, Pakistan's lower secondary net enrolment rate is meager compared to other countries like Sri Lanka, Bangladesh, India, Bhutan revealed by Pakistan Social and Living Standards Measurement survey. Pakistan Education Statistics points out that from 2008 to 2016, teachers having professional qualifications like B.Ed and M.Ed faced a decline of 77% to 59% at the lower secondary level in Sindh. In 2016,22.9, million children of Pakistan are out-of-school and among them, 12.2 and 10.7 million and female and male [2]**.**

In 2015, a survey was conducted in rural Sindh, revealing that there are 24%,12%, and 64% of government boys, girls, and co-education schools. It was reported that 24% of children were out of school, which is lower than (27%) the previous year[3]. Overall assessment of Government in Pakistan economic survey 2018-19 suggests that the key performance indicators for improving education are enrolment rate, number of institutes, and teachers [4].

A recent study shows that improvement in school infrastructure provides a significant increase in literacy rate and quality enhancement, as these are strongly correlated. The finding also reveals an awful situation in Sindh that is more than 60% of schools have no electricity, above 40% are without toilet facilities. Approximately 50% of schools have no drinking water[5].In 2019, the Asian development bank revealed that there had been some progress in teacher education as the Sindh government shows interest in developing policies and institutions for teacher training [2].

Pakistan provides many educational facilities at different levels of education, but still, the quality of education is the first and foremost challenge for Pakistan. The above facts and figures inspire the present study to examine the contribution and in-depth understanding of different factors enhancing Sindh secondary school education.

Interpretation often becomes problematic when we deal with a larger dataset. Different techniques are used to cope with such issues; PCA is the oldest and widely used technique. The principal component analysis is the technique used to reduce the dimension of the large dataset by keeping most of the crucial information of the data. PCA converts high-

*Corresponding author. Tel.: +92-336-0356860; fax: NA
E-mail address: habibakhan33@yahoo.com

dimensional datasets to low dimensions by transforming old variables into a new set of principal components. PCs arise due to the transformation of original correlated variables into uncorrelated variables, indicating that PCs generated are orthogonal and present a linear combination of original correlated variables [6].

2. Literature Review

Many investigations are conducted to improve the quality of education, as education plays a crucial role in our lives and society. The outcome of such studies evaluated from different approaches indicates various significant factors that influence distinct areas of education. These results have played an essential role in the enhancement of educational institutions. Below is the work done by researchers for the quality of education in different aspects.

School attendance and dropout rates are affected by the quality of school infrastructure to a great extent. Low quality of infrastructure shows lower school attendance and more excellent dropout rates at elementary and secondary school using different factors such as Daily attendance of a student, Students dropout, Buildings, Structural repair, Staff, Student background, Student to teacher ratio, Average teacher experience, Enrolment, Mobility, Poverty, Teacher interest [7]. Results show that the used infrastructure factors affect the students' enrolment and academic performance and thus need to be maintained by authorities to get satisfactory performance[8]. On the other hand, this research study reveals that the achievement and enrolment of students possess a positive relationship with teacher-related variables [9]. However, this study suggests that the relationship of student satisfaction with facilities and exam results has minor yet significant [10]. This study reveals that the teacher's performance is significantly influenced by the work environment [11]. Moreover, researchers also assessed that schools should reduce the extra responsibility of teachers as it has a significant negative impact on student performance [12]. A general review was conducted on different countries at a primary level of education. This study shows that these countries have no suitable quality of education in primary schools due to poor teaching skills and practices [13]. Another study reported that school infrastructure and educational resources strongly influence education performance at secondary school [14].

3. Data and Methodology

The principal component analysis aims to reduce the data size by extracting the most decisive information and further looking over the arrangement of the observations and indicators[6]. Principal components have two essential features: there is no correlation among them, and secondly, there is always a decline in the variance of the first principal component to the last. This study uses the principal component analysis technique to develop the two Sindh secondary school educational index (SSEI) equations because of its competence and data dimensionality reduction approach by calculating eigenvalues and eigenvectors correlation matrix and PC loadings from the rotated component matrix. This research depends on data assembled from Pakistan Education Statistics (2000-2016) of Public Sector, Secondary Schools of Sindh by selecting the most suitable indicators in the construction of index equation that helps inspect secondary schools. These indicators are x_1=the total number of institutions (INS), x_2=total number of enrolment (ENROL), x_3=total number of teachers (TCH), x_4=total number of teachers having professional qualification (TPQ), x_5=the total number of teachers with academic qualifications (TAQ), x_6=the total number of repeaters (RP), x_7=availability of building (BL), x_8=availability of electricity (EL), x_9=availability of drinking water facility (DW), x_{10}=availability of washrooms (WR), x_{11}=availability of boundary wall (BW), x_{12}=satisfactory building condition (BC), x_{13}=student-teacher ratio (STR), x_{14}=teacher-institution ratio (TIR).

Initiating the methodology by first normalizing the indicators. For the normalization of data Min-Max is technique is used. This normalization method scales the data from zero (minimum) to one (maximum). The mathematical formula of this feature scaling technique is:

$$\boldsymbol{Z_{i,t}} = \frac{\boldsymbol{x_{i,t} - \min(x_{i,T})}}{\boldsymbol{\max(x_{i,T}) - \min(x_{i,T})}} \quad (1)$$

Where, $Z_{i,t}$ is the normalized factor, i= 1 to 14, t= corresponding year i-e 2000-2016, T= the total observation period. Further, the calculation of eigenvalues and eigenvectors are determined by using the characteristics equation:

$$|\boldsymbol{Q} - \boldsymbol{\lambda I}| = \boldsymbol{0} \quad (2)$$

Where Q is the correlation matrix, λ is the obtained eigenvalues, I is the Identity matrix. Now, the Coefficients of each indicator in the overall linear combination are calculated by using the rotated component matrix ([15]) as:

$$\boldsymbol{\beta_i} = \frac{\boldsymbol{C_{ij}\lambda_j}}{\boldsymbol{\sum_{m=1}^{14} \lambda_m}} \tag{3}$$

Where, β_i= coefficient of each indicator, C_{ij}= rotated component matrix of the i^{th} indicator (row) on j^{th} (column), $i = row\ 1\ to\ 14$ (indicator) of the rotated component matrix, j= largest value of the extracted principal component(s) for each indicator, m= total number of eigenvalue,λ_j= eigenvalue of j^{th} rotating component.

Finally, the weight of the SSEI equation is calculated by equation (4) **(**[16]**)** as:

$$\boldsymbol{(SSEI)_p} = \frac{\boldsymbol{\sum_{i=1}^{14} X_i(\sum_{i=1}^{14} \beta_i)}}{\boldsymbol{\alpha}} \tag{4}$$

where,

$$\boldsymbol{\alpha} = \sum(\boldsymbol{\sum_{i=1}^{14} \beta_i}) \tag{5}$$

p= Sindh public sector gender-wise, i.e., rural male (RM) and rural female (RF). Eq.(5) is defined as the summation of coefficients obtained from Eq.(3) to calculate the weight of each indicator in Eq.(4).

4. Results and Discussions

4.1. Rural Female Model

The first two components are selected based on corresponding eigenvalues, which explain nearly 87% of the total variation, as shown in Table 3A and Table 4A (appendix A.1). Thus the $\boldsymbol{(SSEI)_{RF}}$ equation of RF is given below:

$(SSEI)_{RF}$=0.092942INS+0.092895ENROL+0.091041TCH+0.091788TPQ+0.041704TAQ+0.007588RP+ 0.037281STR+0.045565TIR (6)

As seen in appendix A.1, INS, TCH, TPQ, BL, EL, DW, WR, BW have positively high correlation whereas BC and STR have moderate correlation and the remaining indicators hold a weak correlation with enrolment. RP is negatively weak but correlated with all indicators.Fig.1(a) shows the historical SSEI and the share of each normalized indicator during 2000-2016 of a rural female, which is derived from the annual values of normalized indicators and the total weight calculated in Eq.(6).

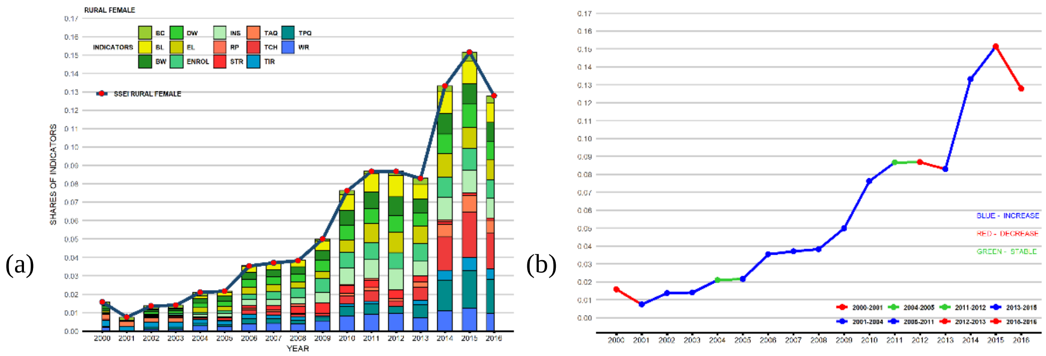

Fig.1. (a) Contribution of indicators 2000-2016 (b) Trends of SSEI rural female 2000-2016.

Fig.1(b) represents different trends that occur throughout the period, which explains the variations in the SSEI. There are three decreasing (2000-2001,2012-2013,2015-2016) and three increasing periods i-e 2001-2004,2005-2011,2013-2015 whereas there is lesser stability in 2004-2005,2011-2012 as clearly shown in Fig.1(b). Here, SSEI (rural female) examination is taken under two different years, 2000 and 2016, with a peak in 2015 and the lowest contribution of indicators can be seen in 2001(Fig.1(a)). In 2000, the share of TIR, TAQ, WR, BC, BW was more that is 0.33%,0.25%,0.20%,0.19%,0.18% whereas, TIR is 0.26%, TAQ is 0.25%, BC is 0.19% having the highest share in 2001 while ENROL, TCH, TPQ, BL, EL, DW, WR, BW, STR have no contribution. In peak year (2015), TCH, TPQ, DW, BL, INS, WR, ENROL were the dominated indicators with 2.45%,2.07%,1.25%,1.24%,1.24%,1.23%,1.20% contribution. At the end of the observed year, TCH, TPQ, EL, INS, BL, ENROL were the prevailing indicators with a share of 1.91%,1.84%,1.10%,1.08%,1.08%,1.01% but showed an overall decline in the share as compared to the previous year.

4.2. Rural Male Model

Likewise, the SSEI equations of the RM model are given below, which are constructed using their respective extracted eigenvalues and rotated component matrix.

$$(SSEI)_{RM}=0.099393INS+0.093544ENROL+0.088101TCH+0.097926TPQ+0.032237TAQ+0.021113RP+0.098974BL+0.100013EL+0.099659DW+0.100165WR+0.099899BW+0.023507BC+0.02494STR+0.020524TIR \quad (7)$$

(a) (b)

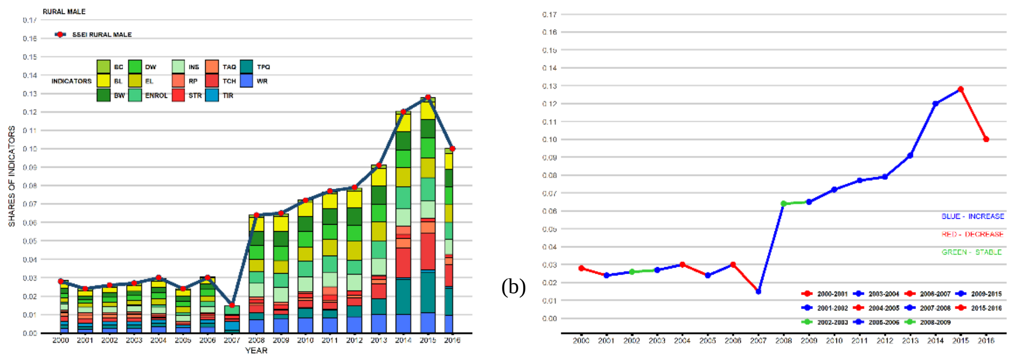

Fig.2. (a) Contribution of indicators 2000-2016 (b) Trends of SSEI rural male 2000-2016

Table 2B (see appendix B.2,) reveals that all indicators have robust correlations except BC, TIR possesses moderate correlation, and TAQ is weakly correlated with enrolment. All indicators are weakly correlated with repeaters. Fig.2(a) shows the historical SSEI and the share of each normalized indicator during 2000-2016 of RM, which is derived from the annual values of normalized indicators and the total weight calculated in Eq.(7). Fig.2(b) represents different trends that occur throughout the period, which explains the variations in the SSEI. There are four decreasing (2000-2001,2004-2005,2006-2007,2015-2016) and five increasing periods i-e 2001-2002, 2003-2004, 2005-2006,2007-2008,2009-2015 whereas there are less stable period which is 2002-2003,2008-2009 as clearly shown in Fig.2(b). SSEI (RM) examination is taken under two different years, 2000 and 2016, with a peak in 2015 and the lowest contribution of indicators can be seen in 2007(Fig.2(a)). In 2000, the share of INS, TCH, BL, EL, DW, WR was more that is 0.30%,0.29%,0.29%,0.26%,0.25%,0.24%,0.24% whereas, TIR is 0.48%, ENROL is 0.45%, TCH is 0.13%, TPQ is 0.14% having the highest share in 2007 while INS, BL, EL, DW, WR, BW, BC have no contribution. In peak year (2015), TPQ, TCH, ENROL, DW, EL were the dominated indicators with 2.21%,2.00%,1.28%,1.10%,1.05% contribution. At the end of the observed year, TCH, TPQ, DW, EL, BW were the prevailing indicators with a share of 1.19%,1.47%,0.97%,0.96%,0.96% but showed an overall decline in the share as compared to the previous year.

5. Comparative Analysis of Secondary School Educational Indices

Fig. 3 compares the SSEI of UM and RM. The historical SSEI of RM observed a four year decline in RM while three years in RF. Also, more growth can be seen in RF compared to RM (eleven years and ten) along with stability in 2002-2003 and 2008-2009 in RM whereas 2004-2005 and 2011-12 in RF. Apart from this, 2007 is recorded as the lowest year in RM, with a value of 0.0151 while 2001 is recorded as the lowest year in RF, with a value of 0.0075. Also, in 2005, the overall share of indicators in both RF and RM are very near i-e 0.021 and 0.024.

Table 1: lowest to highest SSEI weight of each indicator in (%)

	RF	RM
1	0.759RP	2.052TIR
2	3.166BC	2.111RP
3	3.728STR	2.351BC
4	4.17TAQ	2.494STR
5	4.556TIR	3.224TAQ
6	9.104TCH	8.81TCH
7	9.179TPQ	9.354ENROL
8	9.29ENROL	9.793TPQ
9	9.294INS	9.897BL
10	9.328BW	9.939INS
11	9.336BL	9.966DW
12	9.345EL	9.99BW
13	9.363DW	10.001EL
14	9.381WR	10.017WR

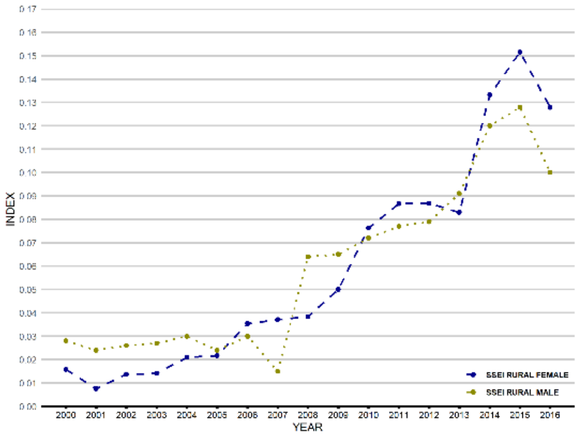

Fig.3. Comparison of Rural SSE Indices 2000-2016

Fig.3 reports two significant peaks: 2014 and 2015 with a value of 0.1332 and 0.1515 in RF and 0.12 and 0.128 in RM. Clearly, in RF, there was less contribution of indicators in the early years, i-e 2000-2005; however, in later years from 2010 to 2016, RF shows the highest contribution with a decline in 2013 compared to RM. Notice that in 2016, the overall share of indicators of RF is more compare to RM but decline in both model. After 2008-2009, RM showed improvement in share of indicators as compared to RM. There is a continuous improvement after 2001 till 2012 in RF, unlike in RM, it has been observed from 2007 till 2015 as shown in the figure.

6. Conclusion

This research executes PCA methodology to construct two secondary school educational index (SSEI) equations of above mentioned indicators. Table 5 represents that WR, EL, BW are in RM whereas WR, DW, EL, BL are in RF. Moreover, an increasing trend can be seen in RM from 2008 until 2015, while 2013 shows a slight decline after 2001 in RF. Results indicate that in RM and RF, enrolment possesses a positive and significant relationship with all the indicators except TAQ. In contrast, repeaters have an insignificant and negative correlation.

7. Policy Recommendations

The outcomes of this research give awareness about the performance of demographic educational indicators and thus, recommends some policy implementations for the satisfactory development in the educational field such as:

- Improvement in school infrastructure or building should be one of the main targets as it enhances the students' interest and thus increases the enrolment rate.
- The results highlight the dynamics of gender inequality in school education and thus are beneficial to formulate policies to eliminate it.
- To avoid the undesirable outcomes, policy-makers should be familiar with the relationship among the indicators.

Acknowledgement:

The authors are wishing to acknowledge the Directorate of Post Graduate Studies Mehran University of Engineering and Technology, Jamshoro, Sindh, Pakistan and greatly acknowledge the provided comments, suggestions, and for their support to carry on this research.

References

[1] C. Simmons and C. Hawkins, "Learning for All," *Teach. Comput.*, pp. 213–236, 2020, doi: 10.4135/9781473919785.n7.
[2] A. D. Bank, *School education in Pakistan: A sector assessment*, no. June. 2019.
[3] S. I. Astuti, S. P. Arso, and P. A. Wigati, "済無No Title No Title No Title," *Anal. Standar Pelayanan Minimal Pada Instal. Rawat Jalan di RSUD Kota Semarang*, vol. 3, pp. 103–111, 2015.
[4] R. O. Nodari, "Agriculture: Risk assessment for Brazil's GM bean," *Nature*, vol. 479, no. 7373, p. 299, 2011, doi: 10.1038/479299e.
[5] M. A. Siddique, A. A. Tagar, Z. A. Khoso, and H. K. Tagar, "Role of Infrastructure to improve quantity and enhance quality of School Education in Sindh Province of Pakistan," *Adv. Soc. Sci. Res. J.*, vol. 6, no. 3, pp. 198–206, 2019, doi: 10.14738/assrj.63.6289.
[6] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley Interdiscip. Rev. Comput. Stat.*, vol. 2, no. 4, pp. 433–459, 2010, doi: 10.1002/wics.101.
[7] D. Branham, "The wise man builds his house upon the rock: The effects of inadequate school building infrastructure on student attendance," *Soc. Sci. Q.*, vol. 85, no. 5 SPEC. ISS., pp. 1112–1128, 2004, doi: 10.1111/j.0038-4941.2004.00266.x.
[8] R. N. PARNWELL, "Influence of School Infrastructure on Academic Performance in Public Primary Schools in Ruiri Location-Meru County,Kenya," 2015.
[9] S. W. Lee and G. Mamerow, "Understanding the role cumulative exposure to highly qualified science teachers plays in students' educational pathways," *J. Res. Sci. Teach.*, vol. 56, no. 10, pp. 1362–1383, 2019, doi: 10.1002/tea.21558.
[10] A. O. Hopland and O. H. Nyhus, "Does student satisfaction with school facilities affect exam results?: An empirical investigation," *Facilities*, vol. 33, no. 13–14, pp. 760–774, 2015, doi: 10.1108/F-09-2014-0076.
[11] S. Rahardjo, "The Effect of Competence , Leadership and Work Environment Towards Motivation and its Impact on the Performance of Teacher of Elementary School in Surakarta City , Central JAVA , Indonesia," *Int. J. Adv. Res. Manag. Soc. Sci.*, vol. 3, no. 6, pp. 59–74, 2014, [Online]. Available: http://www.garph.co.uk/IJARMSS/June2014/7.pdf.
[12] A. Alam and M. Ahmad, "The impact of instructional leadership, professional communities and extra responsibilities for teachers on student achievement," *Int. J. Educ. Manag.*, vol. 31, no. 3, pp. 383–395, 2017, doi: 10.1108/IJEM-09-2015-0126.
[13] T. Bold *et al.*, "Enrollment without learning: Teacher effort, knowledge, and skill in primary schools in Africa," *J. Econ. Perspect.*, vol. 31, no. 4, pp. 185–204, 2017, doi: 10.1257/jep.31.4.185.
[14] F. Ben Yahia, H. Essid, and S. Rebai, "Do dropout and environmental factors matter? A directional distance function assessment of tunisian education efficiency," *Int. J. Educ. Dev.*, vol. 60, no. November 2017, pp. 120–127, 2018, doi: 10.1016/j.ijedudev.2017.11.004.
[15] I. Berk and V. Ediger, "A historical assessment of Turkey's natural gas import vulnerability," *Energy*, vol. 145, no. 2018, pp. 540–547, 2018, doi: 10.1016/j.energy.2018.01.022.
[16] V. Ş. Ediger and I. Berk, "Crude oil import policy of Turkey : Historical analysis of determinants and implications since 1968," *Energy Policy*, vol. 39, no. 4, pp. 2132–2142, 2011, doi: 10.1016/j.enpol.2011.01.058.

Appendix

A.1. Rural Female Model

Table A1: Summary Statistics of indicators.

	INS	ENROL	TCH	TPQ	TAQ	RP	BL	EL	DW	WR	BW	BC	STR	TIR
Mean	0.44	0.46	0.22	0.26	0.27	0.28	0.44	0.43	0.44	0.45	0.49	0.4	0.4	0.37
Maximum	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Minimum	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Std.Dev.	0.38	0.36	0.31	0.32	0.29	0.29	0.38	0.36	0.3	0.31	0.34	0.28	0.32	0.26
skewness	0.241	0.038	1.411	1.326	1.259	1.119	0.237	0.285	0.221	0.248	0.174	0.448	0.44	0.982
Kurtosis	-1.694	-1.745	0.474	0.155	0.338	0.279	-1.678	-1.642	-1.367	-1.47	-1.62	-0.662	-1.367	0.199
# of observation	17	17	17	17	17	17	17	17	17	17	17	17	17	17

Table A2: Correlation matrix (Q) of normalized indicators.

	INS	ENROL	TCH	TPQ	TAQ	RP	BL	EL	DW	WR	BW	BC	STR	TIR
INS	1													
ENROL	.958 ***	1												
TCH	.901 ***	.954 ***	1											
TPQ	.936 ***	.966 ***	.975 ***	1										
TAQ	.315	.387	.466 *	.417 *	1									
RP	-.211	-.191	-.248	-.248	-.159	1								
BL	.996 ***	.958 ***	.914 ***	.938 ***	.316	-.227	1							
EL	.974 ***	.956 ***	.939 ***	.939 ***	.333	-.256	.982 ***	1						
DW	.988 ***	.966 ***	.933 ***	.958 ***	.344	-.225	.989 ***	.979 ***	1					
WR	.98 ***	.969 ***	.946 ***	.946 ***	.341	-.223	.987 ***	.99 ***	.982 ***	1				
BW	.985 ***	.964 ***	.929 ***	.929 ***	.313	-.194	.99 ***	.985 ***	.982 ***	.993 ***	1			
BC	.489 **	.509 **	.521 **	.508 **	.771 ***	-.344	.491 **	.507 **	.501 **	.511 **	.487 **	1		
STR	.627 ***	.587 **	.494 **	.516 **	-.331	-.122	.627 ***	.614 ***	.599 **	.598 **	.629 ***	-.118	1	
TIR	-.021	.106	.221	.169	.764 ***	-.074	-.014	.026	.046	.039	.001	.471 *	-.67 ***	1

Note: * = statistical significance at 10% level of significance,
**= statistical significance at 5% level of significance,
*** = statistical significance at 1% level of significance

Table A3: Eigenvalues evaluated from matrix Q.

	λ_1	λ_2	λ_3	λ_4
Eigenvalues	9.557173	2.61093	0.968059	0.418439
Variability(%)	68.26552	18.6495	6.914708	2.988847
Cumulative Variance Explained (%)	68.26552	86.91502	93.82973	96.81858

Kaiser criterion (1960): Propose to keep eigenvalue greater than one.

Table A4: Rotated component matrix.

	C_1	C_2
INS	0.31727	0.06713
ENROL	0.31711	0.01102
TCH	0.31078	-0.05611
TPQ	0.31333	-0.0269
TAQ	0.13303	-0.52111
RP	-0.08767	0.09481
BL	0.31868	0.06459
EL	0.31901	0.0452
DW	0.31962	0.03826
WR	0.32024	0.03883
BW	0.31844	0.06363
BC	0.18218	-0.39565
STR	0.18755	0.46585
TIR	0.029	-0.56936

Note: Two components extracted

B.1. Rural Male Model

Table B1: Summary Statistics of normalized indicators

	INS	ENROL	TEACHER	TPQ	TAQ	RP	BL	EL	DW	WR	BW	BC	STR	TIR
Mean	0.62	0.43	0.26	0.26	0.3	0.28	0.62	0.56	0.53	0.54	0.58	0.48	0.51	0.25
Maximum	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Minimum	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Std.Dev.	0.33	0.34	0.28	0.3	0.27	0.29	0.33	0.35	0.32	0.33	0.36	0.24	0.33	0.24
skewness	-0.233	0.119	1.471	1.378	1.259	1.472	-0.218	0.01	-0.094	-0.076	-0.129	0.138	-0.239	1.462
Kurtosis	-1.607	-1.529	0.886	0.472	0.65	1.307	-1.634	-1.738	-1.71	-1.713	-1.821	-0.009	-1.488	2.42
# of observation	17	17	17	17	17	17	17	17	17	17	17	17	17	17

Table B2: Correlation matrix (Q) of normalized indicators.

	INS	ENROL	TEACHER	TPQ	TAQ	RP	BL	EL	DW	WR	BW	BC	STR	TIR
INS	1													
ENROL	.883 ***	1												
TEACHER	.764 ***	.779 ***	1											
TPQ	.901 ***	.884 ***	.899 ***	1										
TAQ	.432 *	.387	.802 ***	.6 **	1									
RP	-.184	-.247	-.224	-.238	-.246	1								
BL	.997 ***	.876 ***	.765 ***	.895 ***	.432 *	-.184	1							
EL	.977 ***	.908 ***	.813 ***	.948 ***	.458 *	-.162	.971 ***	1						
DW	.965 ***	.901 ***	.804 ***	.946 ***	.479 *	-.302	.96 ***	.967 ***	1					
WR	.98 ***	.906 ***	.804 ***	.948 ***	.479 *	-.211	.975 ***	.987 ***	.988 ***	1				
BW	.979 ***	.908 ***	.805 ***	.932 ***	.451 *	-.224	.973 ***	.988 ***	.97 ***	.976 ***	1			
BC	.651 ***	.444 *	.687 ***	.655 ***	.746 ***	-.306	.651 ***	.62 ***	.634 ***	.629 ***	.603 **	1		
STR	.696 ***	.807 ***	.512 **	.646 ***	-.02	.034	.688 ***	.703 ***	.667 ***	.681 ***	.711 ***	.134	1	
TIR	-.799 ***	-.65 ***	-.517 **	-.658 ***	-.056	-.077	-.795 ***	-.755 ***	-.745 ***	-.75 ***	-.768 ***	-.381	-.807 ***	1

Note: * = statistical significance at 10% level of significance, **= statistical significance at 5% level of significance, *** = statistical significance at 1% level of significance

Table B4: Rotated component matrix.

	C_1	C_2
INS	0.30623	0.0628
ENROL	0.28821	0.09941
TCH	0.27144	-0.2292
TPQ	0.30171	-0.07011
TAQ	0.16715	-0.55677
RP	-0.0733	0.36464
BL	0.30494	0.05941
EL	0.30814	0.05199
DW	0.30705	0.00247
WR	0.30861	0.02828
BW	0.30779	0.05152
BC	0.21183	-0.40599
STR	0.22406	0.43082
TIR	-0.24055	-0.35447

Note: Two components extracted

Table B3: Eigenvalues evaluated from matrix Q.

	λ_1	λ_2	λ_3	λ_4
Eigenvalues	10.17819	1.815738	0.887501	0.509448
Variability(%)	72.70136	12.96955	6.339291	3.638915
Cumulative Variance Explained (%)	72.70136	85.67091	92.01021	95.64912

Kaiser criterion (1960): Propose to keep eigenvalue greater than one.