UNIVERSITY OF CAPE COAST

LATENT FACTORS THAT AFFECT APPLICANTS' CHOICE OF THE UNIVERSITY OF CAPE COAST

IVY EYIAH INKUM

2012

UNIVERSITY OF CAPE COAST

LATENT FACTORS THAT AFFECT APPLICANTS' CHOICE OF THE UNIVERSITY OF CAPE COAST

BY

IVY EYIAH INKUM

Thesis submitted to the Department of Mathematics and Statistics of the School of Physical Sciences, University of Cape Coast, in partial fulfillment of the requirements for the award of Master of Philosophy Degree in Statistics.

MAY 2012

DECLARATION

Candidate’s Declaration

I hereby declare that this thesis is the results of my own original research and that no part of it has been presented for another degree in this university or elsewhere.

Candidate’s Signature………………………. Date…………………

Name: Ivy Eyiah Inkum

Supervisors’ Declaration

We hereby declare that the preparation and presentation of this thesis were supervised in accordance with the guidelines on supervision of thesis laid down by the University of Cape Coast.

Principal Supervisor’s Signature…………………. Date……………….

Name: Dr. N. K. Howard

Co-Supervisor’s Signature……………………….... Date……………….

Name: Prof. B. K. Gordor

ABSTRACT

Choice of university education in Ghana is a competitive exercise for Senior High School students. Choosing it requires consideration of various factors.

The motivation for the study is to measure high school students' perception of the University of Cape Coast. To this end, the study focuses on various factors. The objective of the study therefore is to determine the latent factors that are considered by the SHS students in their decision. In order to achieve this objective, 20 Senior High Schools were selected for the study. A questionnaire was used to that effect. The hypotheses related to the thesis lie in the choice and significance of the factors the students used to answer the 28 indicators. The main analytical technique used in the study is factor analysis.

After rotation, the main latent factor influencing applicants' choice of the University of Cape Coast is academic resources. Four other factors in decreasing order of influence are quality assurance, cost of study, influences of others and discipline. It is worth mentioning that almost all the factors are dependent on some demographic characteristic such as age and sex of the respondents.

ACKNOWLEDGEMENTS

First of all, I am very grateful to my principal supervisor, Dr. N. K. Howard for his enormous contribution and corrections which aided in the completion of this thesis. Also, I really appreciate my co-supervisor Prof. B. K. Gordor for his constructive criticisms.

I am most thankful to the Head of Department of Mathematics and Statistics, Dr. E. K. Essel for his assistance. To all the headmasters and headmistresses who cooperated with me during my data collection, I say thank you.

I cannot forget Mr. B. K. Nkansah and Mr. F. Eyiah-Bediako of the Department of Mathematics and Statistics for showing deep concern in this study. I wish to thank all the staff of Mathematics and Statistics Department for giving me the opportunity to enroll for this programme.

Above all I wish to express my sincere gratitude to my parent for supporting me financially and spiritually. I also thank my colleagues for being there for me at all times.

DEDICATION

To my family

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

BACKGROUND TO THE STUDY

Each year a huge number of high school students complete school with anticipation of what may be the most significant decision of their young lives. With over 60 universities, polytechnics and training colleges in Ghana, the decision of where to submit applications has become a daunting task for both students and parents. Researchers have examined university choice process with a variety of approaches in an attempt to identify factors that influence the decisions of senior high school students to attend a university. According to Kim (2004), every student has his or her own preferences about universities based on institutional type, prestige, or even a student's 'intuitive feelings' about how his or her personality fits into a certain university. Consequently, the results of studies on the choice of university are of particular interest to students and their parents.

Every human being at a point in time is presented with a problem of choosing between a set of options in order to achieve a purpose in life. To choose means to decide which of a number of different things or people is best or most appropriate. The decision making process is complex and subject to multiple influences that not only interact with each other but also change over time. It is mostly said that the choices you make today will have an influence on your future. Hence a person's inability to choose correctly can affect him or her negatively in the future. This is especially true with major life decisions such as choosing a university to attend.

A number of deliberations are made in making a choice and it is upon such deliberations that one takes a decision to choose. Certain decisions take either a shorter or longer time based on how simple or complex an issue is. For instance deciding on the type of food to eat or the dress to wear in a particular day may take some few minutes whereas deciding on the type of institution to attend may take a longer time.

In the case of university education, the first thing to think about is whether or not to attend a university. Students who have completed the senior high school are faced with this challenge. Most of them do not know the type of institution to attend. Choices include universities, training colleges, nursing training, polytechnics, technical schools, and so on. To decide on one of these institutions results from careful considerations such as finance, the grades obtain previously and many other factors. At this information gathering stage, advice of friends, counselors, and parents has great influence.

The next level of decision making (assuming the choice is to obtain a university education) is the type of university to attend. The student must decide either to attend a public university or a private university based on certain factors. The literature on university choice is vast and investigates many factors that students consider when choosing to enroll at a particular university. Some of these factors may include the programmes offered by a particular university, the location of the university, constructive environment for learning, easy access to information, parental influence, the university's reputation and cost of tuition.

STATEMENT OF THE PROBLEM

Within the last few decades, the competition among students who would like to attend universities has intensified. These students are not only concerned about the university to enroll but they are particularly interested in institutions with higher achievements and prestige. Moreover, with a multitude of students vying for the best universities, it is a greater challenge for some students than others to attract the most desirable institutions of their choice (Geiger, 2002). Hence, it is very essential to find out which factors actually influence the choice of the University of Cape Coast by prospective applicants.

Furthermore, many institutions of higher education in Ghana are striving for greater levels of status and prestige. Besides, universities such as University of Ghana, Kwame Nkrumah University of Science and Technology (KNUST) and other private universities have a long history of prestige for quality, therefore attracting the most qualified students. Studies that have investigated factors influencing the choice of university by high-achieving students repeatedly cite academic reputation as one of the factors of choice (Chapman & Jackson, 1987; Goenner & Snaith, 2004; Manski & Wise, 1983). For that reason, it is very necessary to know these factors so as to help the authorities of the University of Cape Coast to know which areas need to be improved to attain greater level of status and prestige. It will also help both students and parents to make informed decision when choosing a university.

REVIEW OF EDUCATION IN GHANA

Every country has its own educational system and Ghana is no different. The country of Ghana offers its citizens easy access to a quality education.

“Ghana is endowed with a good education system”, a statement made by BBC News Monitoring Department. Over the centuries education has had different goals, from spreading the Gospel to creating an elite group to run the colony. The British laid a solid foundation for the formal education system in Ghana, however only a small group had access to it. The Nkrumah Government saw education as a major instrument for national development and introduced the policy of education for all. After Ghana gained its independence in 1957, the education system has undergone a series of reforms. Especially the reforms in the 1980’s geared the education system away from purely academic to more in tune with the nations manpower needs.

The present structure of education, which starts at the age of six years, consists of six years of Primary education, three years of Junior High School, three years of Senior High School and four years University courses. Naturally students who successfully pass the Senior Secondary School Certificate Examination can also follow courses at a Polytechnic, Teachers Training College or other tertiary institutions. The first nine years form the basic education and are free and compulsory. Primary and Junior High Schools education is tuition-free and will be mandatory when enough teachers and facilities are available to accommodate all the students. At primary school, the students will learn the basics of education such as reading, writing, science and math whilst at junior high school, they will learn a variety of vocational and educational classes that will help them decide what to choose in the senior high school courses.

Pupils may enter senior high or technical/vocational schools for a three-year course after basic school, which prepare them for university education.

Students usually study a combination of three (in some cases, four) elective subjects and a number of core subjects. For example, a science student could study Additional Mathematics, Chemistry, Biology and Physics as his/her elective subjects. A business student might study Economics, Accounting and Managements as his/her elective subjects. In addition to the elective subjects, there are core subjects, which are those studied by all students in addition to their electives. The core subjects include Mathematics, English and Science. At the end of the three year senior high course, students are required to sit for the West African Senior Secondary Certificate Examinations (WASSCE). Students who obtain aggregate 18 or better (six is best) can enter the university. Usually, the score is determined by aggregating the student's grades in his elective subjects. The aggregate score is then added to the aggregate score of his best core subjects, with scores in English and Mathematics considered first. Education in Ghana is mainly in English. This allows Ghana to bring in teachers from all over the world in order to meet the rising demand for teachers in the area. Entrance to universities or tertiary institutions is by examination following completion of senior high school.

Ghana has developed a road map known as Vision 2020. The Vision 2020 document contains an education policy with the objectives to ensure all citizens regardless of gender or social status, are functionally literate and productive at the minimum. It further states that in order to achieve Vision 2020, the education system must embrace science and technology since we are in a technological era and countries that fail to recognize this will not be able to escape the clutches of poverty. As a results Ghana devotes over 40% of its country’s budget to the educational system.

BRIEF ACCOUNT OF THE UNIVERSITY OF CAPE COAST

The University of Cape Coast was inaugurated in December 1962 as a university college. It was given a full independent university status on October 1st 1971. The university is one of the rare sea-front universities in the world. It was established to train graduate teachers for Second Cycle Institutions; Teacher Training Colleges; and Technical Institutions, a mission that the two universities existing then (University of Ghana and Kwame Nkrumah University of Science and Technology) were unequipped to fulfill. With careful planning and execution, the university has grown from a fledgling University College to a giant institution of excellence and choice in Africa and the world.

The institution continues to expand its existing highly qualified Faculty and Administrative staff, by offering a conducive environment that motivates them to position the University to respond effectively to the development needs of a changing world. Due to expansion and globalization, University of Cape Coast currently runs other programmes aside the Education degrees. These programmes include educational music and theatre studies, planning and administration, business administration, commerce, actuarial science, tourism, optometry, labor studies, governance and democracy, computer science, agriculture and information technology among other programmes. The University of Cape Coast Medical School started in 2007. The Faculty of Law which will train students to become professional Lawyers is in progress. This is to allow flexibility and choice in its course offerings and thus cater for specific needs of students, while still focusing on its initial mission.

The above mentioned programmes are run by eight Faculties and Schools; The School of Business, Faculty of Arts, School of Agriculture, Faculty of Education, Faculty of Social Sciences, School of Medical Sciences, School of Biological Sciences and School of Physical Sciences. The University of Cape Coast also has a Centre for Continuous Education (CCE) purposely for distance students. Applicants can get admission into the university through the West African Senior Secondary Certificate Examination (WASSCE), Matured Student Examination, HND certificates and others.

OBJECTIVES OF THE STUDY

The main objective of the study is to determine the latent factors that influence applicants' choice of UCC.

To achieve this objective, there is the need:

1. To compile the ratings of some indicators that influence prospective applicants' choice of UCC)
2. To identify the correlations that exist between these factors which influence prospective applicants' choice of UCC.
3. To determine whether the extracted factors are dependent on some demographic characteristics of the respondents or not.

RESEARCH QUESTION

The research seeks to address the following questions:

1. What are the main factors that may affect prospective applicants' choice of the University of Cape Coast?
2. How do students rate the indicators that influence prospective applicants' choice of UCC?

3. What are the correlations that exist between these factors that influence applicants choice of UCC?
4. Is the extent of influence of a particular factor dependent on the demographic characteristics of the respondents?

SIGNIFICANCE OF THE STUDY

This study has significance for several areas. First, the results of this study will be highly beneficial to students, their parents and senior high school counselors in providing insights into the university selection process. Moreover, the research will guide the administrators and policy makers of this university to know the areas that need urgent development.

This study will add to the body of literature related to university choice by exploring the major factors that influence prospective applicants' choice for senior high school level. This will give a broad view of the perception of senior high school students about the University of Cape Coast. This study can be compared with similar studies to know if there are differences in the results obtained.

DATA COLLECTION

In order to address the objectives of the study, two stages of data collection were adopted. The first stage involves data collection from the Students Records & Management Information Section (S.R.M.I.S) of the University of Cape Coast. The second stage involves data collection from SHS through the administration of questionnaires to the selected senior high schools and data analysis.

First Stage Data Collection

Data on the distribution of successful applicants by schools were collected from S.R.M.I.S of the University of Cape Coast over a period of five years (i.e. from 2006 to 2010). This data was collected because the study population for this research was Secondary Schools that had their candidates getting admission into UCC. Applicants from over 350 secondary schools are admitted into the University of Cape Coast each year within this five year period.

Schools with over 30 candidates getting admission into the University of Cape Coast were purposely selected from these 350 schools. These schools were selected because the researcher cannot have enough time and money to visit over 350 secondary schools scattered all over the country. Only secondary schools were considered because they contribute the majority of candidates who are admitted into the university.

A summary of the schools selected for the study and their respective percentages in the university over the period of five years is presented in Table 1.

Table 1: Schools with 30 or More Students Admitted into UCC

Schools	Academic Year									
	2006/2007		2007/2008		2008/2009		2009/2010		2010/2011	
	No.	%	No.	%	No.	%	No.	%	No.	%
St. Augustine, Cape Coast *	98	2.41	115	2.83	100	2.51	85	2.19	90	2.12
Ghana National, Cape Coast	82	2.02	77	1.89	55	1.38	73	1.88	67	1.58
Mfantsiman Girls, Saltpond*	80	1.97	50	1.23	50	1.25	51	1.31	90	2.12
Pope John's, Koforidua*	79	1.94	83	2.04	83	2.08	49	1.26	60	1.41
Aggrey Memorial, Cape Coast	76	1.87	68	1.67	56	1.40	63	1.62	75	1.77
Prempeh College, Kumasi*	65	1.60	57	1.40	71	1.78	46	1.19	46	1.11
Fijai, Takoradi	62	1.53	48	1.18	58	1.45	51	1.31	65	1.53
Mfantsipim, Cape Coast*	53	1.30	79	1.94	47	1.18	44	1.13	30	0.71
Opoku Ware, Kumasi*	53	1.30	61	1.50	74	1.85	56	1.44	41	0.96
Yaa Asantewaa Girls, Kumasi*	51	1.26	44	1.08	48	1.20	39	1.00	62	1.46
Accra Academy, Accra*	50	1.23	58	1.43	46	1.15	51	1.31	55	1.29

Table 1 (continued)

Schools	Academic Year									
	2006/2007		2007/2008		2008/2009		2009/2010		2010/2011	
	No.	**%**	**No.**	**%**	**No.**	**%**	**No.**	**%**	**No**	**%**
Adisadel College, Cape Coast*	48	1.18	61	1.50	75	1.88	77	1.98	81	1.91
Anglican, Kumasi	47	1.16	55	1.35	43	1.08	30	0.77	43	1.01
Archbishop Poter Girls, Takoradi*	47	1.16	36	0.88	48	1.20	48	1.24	61	1.44
T I Ahmadiya, Kumasi	46	1.13	35	0.88	31	0.78	32	0.82	44	1.04
GSTS, Takoradi *	45	1.11	38	0.93	42	1.05	48	1.24	37	0.87
Okuapeman, Akropong	44	1.08	34	0.84	33	0.83	50	1.29	49	1.15
St John's, Takoradi*	36	0.89	32	0.79	41	1.03	39	1.00	56	1.32
Ghanass, Koforidua	34	0.84	41	1.01	37	0.93	41	1.06	46	1.08
University Practice, Cape Coast	31	0.76	41	1.01	37	0.93	47	1.21	48	1.13
Total percentage		**27.7**		**27.4**		**26.9**		**26.3**		**27.0**

Schools with asterisks (*) are single sex schools

Out of over 350 schools from which students get admission into the university, only 20 schools consistently had a minimum of 30 students being admitted. The total percentage for these 20 schools ranges between 26 and 28 which are over one fourth of the total student population. Six out of these 20 schools (30%) are in Cape Coast, five (25%) are in Kumasi, four (20%) are in Takoradi, two (10%) are in Koforidua, only one (5%) is in Accra, one (5%) is in Akropong and one (5%) is in Saltpond. Throughout the five years St. Augustine's College had the highest percentage of student who got admission into the University of Cape Coast. Twelve schools sampled for this research are single sex schools. Thus, the schools with asterisk are the single sex schools. Nine out of the twelve single sex schools are boys' school whereas only three are girls' school.

Second Stage Data Collection

This stage involves the discussion of the sampling procedure used in the study and the designing of the questionnaire. Also, demographic characteristics of the respondents are examined in this section.

Sampling of Respondents

The study population was made up of all Senior High Schools from which students get admission into the University of Cape Coast. The sampling procedure used for this research is purposive sampling. In this sampling technique, all the final year students in the 20 selected secondary schools are sampled. Within this sample, those who want to attend the University of Cape Coast are purposely chosen. Purposive sampling technique is used because the research aims at students who prefer to attend the University of Cape Coast and the reasons for their choice. Moreover, the final year students are most likely to make their choice of a tertiary institution earlier than those in the lower classes.

The main source of data for this research is primary data and the method used in collecting the data was the use of questionnaires. This method was used because it is less expensive, produces quick result, more convenient and often offers a greater assurance of anonymity compared to interview. A pilot survey involving the students and Teaching Assistants of the University of Cape Coast was conducted to help in the identification of the indicator variables. After the pilot survey using a number of indicators, 28 indicators were considered exhaustive since other indicator variables tend to be similar. After designing the questionnaire, it was administered to the final year students in the 20 selected senior high schools who want to attend the University of Cape Coast.

The questionnaire was divided into two sections that sought to gather different sets of information. The first section consisted of four issues on demographic characteristics of the respondents. The second section asked the respondents to indicate the degree of agreement or disagreement with the following statements using a five-point Likert Scale [strongly disagree (1), disagree (2), undecided (3), agree (4), strongly agree (5)]:

V_1: The only university that offers my choice of programme.

V_2: It is located in Central region where there are many tourist attractions.

V_3: It is base on my counselor’s recommendation.

V_4: The university is close to my place of residence (proximity).

V_5: It will be an opportunity to meet old friends.

V_6: I prefer the grading system of this university to other universities.

V_7: It is cheaper to get accommodation in this university.

V_8: The university offer quality programmes.

V_9: Students who complete this university are competent.

V_{10}: Most of my friends and family members attended this university.

V_{11}: The physical structure of the university is very attractive

V_{12}: I like the university because it is located by the sea.

V_{13}: Both students and lecturers are punctual at lectures.

V_{14}: The workers of the university are friendly.

V_{15}: I like the social life of the people in this university

V_{16}: A lot of facilities are available for research in this university

V_{17}: It was recommended to me by friends/relatives.

V_{18}: The cost of living is relatively cheaper in this university.

V_{19}: The administration monitors performance of students very well.

V_{20}: Students comport themselves well during lecture hours.

V_{21}: The administration regulates activities of the staff very well.

V_{22}: Academic user fee is moderate.

V_{23}: The university is recognized worldwide.

V_{24}: There are many facilities for athletics.

V_{25}: A lot of financial sponsorship is available to the student

V_{26}: I like the way students study in this university.

V_{27}: The university has a competent staff.

V_{28}: The university has conducive learning environment.

Any respondent who chooses option 1 or 2 from the five-point Likert Scale for any item considered not to value that item when choosing the University of Cape Coast. Conversely, if a respondent selects option 4 or 5, it implies the item is of much importance in the choice of University of Cape Coast. Whenever option 3 is indicated against any item, it means that the respondent does not mind if the university has that characteristic or not.

A total of 1150 questionnaires were distributed throughout the 20 selected secondary schools. Eight of these questionnaires were discarded and 1142 were valid for the analysis. According to Tabachnick and Fidell (1996), correlations tend to be less reliable when estimated from small samples. Therefore it is important that the sample size be large enough that correlations are reliably estimated. The required sample size also depends on the magnitude of population correlations and number of factors. If there are strong reliable correlations and a few distinct factors, a small sample size is adequate. Comrey and Lee (1992) give us a guide sample sizes of 50 as very poor, 100 as poor, 200 as fair, 300 as good, 500 as very good and 1,000 as excellent. As a general rule of thumb, it is comforting to have at least 300 cases for factor analysis. For most of the analysis in this study, the Statistical Package for the Social Sciences (SPSS, VERSION 16.0) was used.

Demographic Characteristics of Respondents

Out of the 1142 respondents used for this study, 697 representing 61% were males and 442 representing 38.7% were females. Three of the respondents representing 0.3% did not indicate their gender. None of the respondents were below 16 years of age. Almost all the respondents are aged between 16 to 19 years. Thus, 1045 (91.5%) are aged between 16 to 19 years whiles only 96 (8.4%) are aged 20 years and above. One person representing 0.1% did not indicate the age. In addition, Ghana National College had the highest respondents of 89 (7.8%) who would like to attend the University of Cape Coast. This was closely followed by Accra Academy with a total of 77 (6.7%) respondents who would like to attend UCC. Besides, Ghana Secondary School, Koforidua and Ghana Secondary Technical School, Takoradi had the lowest number of respondents who would like to attend UCC. Respondents from these two schools are 37 (3.2%) and 38 (3.3%) respectively.

Furthermore, the distribution of respondents in the various regions indicates that most of the respondents who are living in Greater Accra, Ashanti, Western, Central and Eastern Regions respectively would like to attend UCC. Only three respondents (0.3%) from Upper West region expressed interest in attending UCC. Frequency distribution tables for sex, age, schools and region of birth are presented in Appendix II (a-d).

The 20 senior high schools would be classified as first class schools and second class schools. According to a survey results conducted by Serve Africa.info (2012), the first class schools within this 20 selected schools include St Augustine, Adisadel, Mfantsipim, Accra Academy, Pope John, Opoku Ware, Anglican, T. I. Ahmadiya and Prempeh College. The remaining 11 senior high schools are considered as second class schools.

The ten regions are also grouped into three zones. Greater Accra, Central, Eastern and Volta Regions constitute the first zone. The second zone includes Ashanti, Brong Ahafo and Western Regions. The three Northern Regions as well as the foreigners constitute the third zone. The groupings would help in subsequent statistical analysis.

OUTLINE OF THE STUDY

The thesis consists of six chapters. The first chapter consists of the background to study, statement of the problem, review of education in Ghana, brief account of the University of Cape Coast, objectives of the study, research question, significance of the study and data collection used in the study. Demographic characteristics of the respondents are also discussed in this chapter. The last section is the outline of the thesis which gives a brief overview of the contents of the study.

The second chapter reviews relevant literature on the topic. It talks about what other people have been able to do in this area of study and how they carried out their study. Particular attention is given to the literature review on the correlation matrix. This is because of the low correlations between the indicator variables when using the Likert scale.

Chapter Three dwells on the review of statistical methods such as factor analysis (FA), factor extraction, factor rotation and many others.

The fourth and fifth chapters entail the preliminary and further analysis, respectively, of the research. Detail discussions, summary, conclusions and recommendations are captured in the sixth chapter.

CHAPTER TWO

LITERATURE REVIEW

GENERAL REVIEW OF RELATED STUDIES

Students generally believe that university education has a positive impact on their future success. In addition, students also recognize that degrees from certain institutions are more valuable than those from other universities. Clearly, students and parents in the twenty-first century continue to put significant effort into selecting the right institutions. Since the early twentieth century, some research studies have been conducted in an effort to understand the various factors which are most important to students and their families when making the choice of which university or institution to attend. Some of the related studies on factors affecting students' choice of an institution to attend are reviewed below.

Schoenherr (2009) conducted a study on the factors that influence the College of first choice for high achieving students. In this study, quantitative research design incorporating secondary analysis of data was used. Multiple regression models were conducted to examine the predictive ability of the independent variables, while controlling for other variables in the model, in relation to the choice of college for high achieving students. The three-stage choice model developed by Hossler and Gallagher (1987) was the basis for this study. The model outlines three stages of the college choice process:

1. Predisposition: Students' decisions/aspirations to enroll in post secondary education.
2. Search: the process of considering types of institution to apply.
3. Choice: the selection of an institution to attend.

Results from this study indicated that for high achieving students the second stage in the model had more influence than the first stage in predicting the outcome of college choice. Also, this study found the availability of financial aid to be the most important factor in predicting whether students will attend a higher-tiered or lower-tiered university. Students who consider the availability of financial aid to be very important tend to attend lower-tiered university. The importance of financial aid accounted for over five times the variability of the importance of college cost. The findings of Schoenherr's research are very informative. However, the multivariate analysis used in his research is different from what is employed in this study. Thus, whilst Schoenherr (2009) used multivariate analysis involving multiple regression models in his research, this study focuses on factor analysis.

The findings of a study carried out by Wang (2009) on the factors that influence international students' choice of universities revealed that students' university choice varied according to different reasons. A questionnaire involving 26 indicator variables were administered to a total of 600 international students in Canada. Six most influential factors in decreasing order were identified: university reputation-related factors, economic-related factors, service/facility-related factors, country-related factors, advice and degree recognition.

Harvard economists Avery and Hoxby (2004) wrote an article, entitled "Do and Should Financial Aid Packages Affect Students' College Choice?" using data from the College Admissions Project. Researchers gathered data on 3,240 students from 396 United State High School and utilized the conditional logit technique. As cited by Cheng (2006), Avery and Hoxby focused on how the college choice behaviours of high-ability students are affected by their financial aid packages. Avery and Hoxby concluded that high-ability students are nearly indifferent to a college's distance from their home, to whether it is in-state and to whether it is public. However, they are sensitive to tuition, room and board in the expected direction. They also prefer to attend the most selective colleges in the set to which they are admitted. They are attracted by grants, loans and work-study commitment.

After reviewing previous research, Paulsen (1990) concluded based on characteristics of student background and ability alone, that individuals are more likely to attend college when family income is higher. Furthermore, he found that when a student's family income is greater, he or she is more likely to apply to, or attend, a more highly selective institution, a high-cost institution, an institution located a greater distance from home, a private (rather than a public) institution, and a four-year institution. Another study by Paulsen revealed that college aspirations among blacks were related directly to academic achievement in school, the influence of others (parents, teachers, and friends) and self-esteem. Achievement and significant others' encouragement, but not self-esteem, were important in determining educational aspiration for white.

According to Kern (2000), the biggest influence on college choice appears to be parents, college reputation, friends and the student's own initiative. Kern describes the results of a study about the college choice process for minority high school students considering postsecondary education. He pointed out that financial aid was an important consideration for many students.

Several researchers (Hayden, 2000; Falsey & Haynes, 1984; Russell, 1980) have examined the relationships between student interaction with other college bound students and their college participation. According to Hayden (2000), opinions of friends and former students weigh heavily on the minds of African American college applicants when deciding between colleges. These studies and others expound upon the knowledge that the more a high school student interacts with other students with college plans, the more likely they are to consider going to college.

Canale and Dunlap (1996) surveyed 543 high school seniors and juniors in order to determine the relative importance of certain college traits in their choice of a prospective institution. They found that teacher attributes, areas of study offered, costs and academic reputation were ranked the highest in terms of importance among the list of college characteristics investigated.

With regard to problems of survey instrument, a major drawback of the Likert Scale is that it takes longer time to complete than other itemized rating scales such as the semantic differential and Staple scale (Naresh, 2004). This is because respondents have to read each statement. The reliability of the Likert Scale tends to increase with the number of items used. However as the number of items increase so does the time taken to complete the questions and this may demotivate the respondent. According to the University of Salford (2006), there is no hard and fast rule to determine the final number of items in a scale and this will reflect the nature and complexity of the question. Considering the two psychometric properties, i.e. reliability and validity, the best number of options for Likert Scale is between four and seven (Lozano, Garcia-Cueto and Muniz, 2008). More than seven options will give better psychometric property but one has to make sure it will not exceed the discriminative capacity of the respondents. Problems associated with the use of survey instruments and ratings scales are well documented and researched (Miller and Mitamura, 2003; Rucinski, 1993). Among the most useful research in this area is a series of articles by Schwarz and some of his colleagues that yielded relevant insights for the present research (Schwarz, Grayson and Knauper, 1998; Schwarz and Hippler, 1995).

LITERATURE REVIEW ON CORRELATION MATRIX

Studies conducted on the correlation matrix signified that extremely low correlations are obtained when using the Likert Scale in factor analysis. A study of empirical identification of the major facets of conscientiousness by MacCann, Duckworth and Roberts (2009) clearly indicates that the use of Likert scale usually produces low correlations between the indicator variables. In their study, students rated 117 items on a five point scale. Parallel analysis of these items indicated that nine factors be extracted. MacCann et al. observed that the relationships of conscientiousness facets to academic outcomes were small, ranging from 0.12 to 0.23. Factor loadings greater than or equal to 0.3 were considered as salient in the study.

Ampimah (2007) conducted a study on factor analysis of applicants' choice of Cape Coast Polytechnic. A total of 317 questionnaires involving 14 items were administered to the Cape Coast Polytechnic students. The highest correlation of 0.379 was obtained between the variables "Cost of living being relatively cheaper" and "Accommodation being cheaper". A cut-off value of 0.142 was used in grouping the variables. Generally, correlations between all the 14 variables were particularly low.

Similarly, a study carried out by Nkansah, Gordor and Horward (2007) on factor analysis of bath soap consumerism also recorded low correlations. The study involved 1,000 questionnaires administered to the University of Cape Coast students. A correlation coefficient of 0.217 was used as a cut-off due to the low correlations between the variables. The highest positive correlation of 0.649 was obtained between the variables "Get rid of skin disorders" and "Prevent skin infections". Nkansah et al. considered a loading of 0.5 as a cut-off value in their study.

The results of a confirmatory factor analysis performed to cross validate the factor structure of the educators' attitudes towards educational research scale by Ozturk (2011) indicated that some of the statistically significant correlations were not high enough to signify meaningful relationship. The original scale was developed by the same author and revised based on the results of an exploratory factor analysis. The scale had 29 Likert-type items intended to measure eight dimensions of the variable. A correlation coefficient of 0.13 was statistically significant.

For some of the papers reviewed in this study, only the factor loading matrixes were presented. Nothing significant was said about the correlation matrixes. For instance, Kallio (1995) considered a factor loading of 0.4 as high in her study on the factors influencing the college choice decisions of graduate students. Students were asked to rate 31 institutional characteristics based on the degree of importance each factor played in their final enrollment decision. A principal axis factor analysis with a varimax rotation was carried out on the 31 indicators but the correlation matrix was not presented. Kallio obtained a four factor solution. These are academic, work, spouse and social factors.

Furthermore, a loading of 0.45 has been successfully used by Andriotis, Agiomirgianakis and Mihiotis, (2008) as a cut-off value in their study on measuring tourist satisfaction. Tourist were asked to indicates their satisfaction to a 38-item, 7-point Likert scale. The 38 statements were group using principal components with varimax rotation. The correlation matrix which shows the relationship between the indicators was not presented. The criterion of eigenvalue-greater-than-one was used to determine the number of factors. Nine factors accounted for the variability in the data.

This study investigated into the correlation matrix when using the Likert Scale in factor analysis. It was observed that, low correlations were mostly obtained because the researchers were dealing with categorical variables. Nevertheless, this study also uses categorical variables.

In summary, there has been some research into the factors that influence the choice of universities and colleges by students. Most of these researchers use methods such as multiple regression analysis, conditional logit technique and others in their analysis. The multivariate methods involving factor analysis used in this study to identify the latent factors influencing the choice of the University of Cape Coast, demonstrate the uniqueness of this work. Similar research carried out by Inkoom (2010) used data from students in the University of Cape Coast whereas this study uses data from prospective applicants from 20 selected senior high schools in Ghana. The selection of the sample for this study also adds to the uniqueness of this research work.

CHAPTER THREE

REVIEW OF METHODS

This chapter describes the techniques used to help solve the research questions. The chapter addresses factor analysis using principal component technique and other multivariate methods due to their usefulness in this research work.

FACTOR ANALYSIS

Factor analysis is a collection of methods used to examine how underlying factors influence the responses on a number of measured variables (DeCoster, 1998). Thus, it is a general name denoting a class of procedure primarily used for data reduction and summarization. Since its inception over a century ago, factor analysis has become one of the most widely used multivariate statistical procedures in applied research endeavors across a multitude of domains (e.g. psychology, education, sociology, management, public health). The fundamental intent of factor analysis is to determine the number and nature of latent variables or factors that account for the variation and covariation among a set of observed measures, commonly referred to as indicators. Specifically, a factor is an unobservable variable that influences more than one observed measure and that accounts for the correlations among these observed measures. Factor analysis has numerous applications. For instance:

1. It can be used in market segmentation for identifying the underlying variables on which to group the customers.
2. In product research, factor analysis can be employed to determine the brand attributes that influence consumer choice.
3. In advertising studies, factor analysis can be used to understand the media consumption habits of the target market.
4. In pricing studies, it can be used to identify the characteristics of price-sensitive consumers. For example, these consumers might be methodical, economy minded and home centered.

In factor analysis, an entire set of interdependent relationship is examined. No distinction is made between the dependent variable and the independent variables. For this reason, factor analysis is referred to as an interdependence technique. Factor analysis is helpful in achieving the following:

1. To identify underlying dimensions, or factors, that explains the correlations among a set of variables.
2. To identify a new, smaller set of uncorrelated variables to replace the original set of correlated variables in subsequent multivariate analysis (regression or discriminant analysis).

There are various methods for estimating loadings and communalities of factor analysis. The approach used to derive the weights or factor score coefficients differentiates the various methods of factor analysis. Some of the approaches for estimating the common factors include principal component, common factor analysis, maximum likelihood and generalized least squares.

The method employed in this study for estimating the common factors is the principal component method.

MODEL DEFINITION AND ASSUMPTIONS

Let the variables $v_1, v_2, \cdots, v_p$ with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ be a linear combination of the underlying common factors $f_1, f_2, \cdots, f_m$ with accompanying error term $\boldsymbol{\varepsilon}$ to account for that part of the variable that is not unique (not in common with the other variables). For each variable $v_1, v_2, \cdots, v_p$, the model is as follows:

$$
\begin{aligned}
v_1 - \mu_1 &= \lambda_{11} f_1 + \lambda_{12} f_2 + \cdots + \lambda_{1m} f_m + \varepsilon_1 \\
v_2 - \mu_2 &= \lambda_{21} f_1 + \lambda_{22} f_2 + \cdots + \lambda_{2m} f_m + \varepsilon_2 \\
&\vdots \\
v_p - \mu_p &= \lambda_{p1} f_1 + \lambda_{p2} f_2 + \cdots + \lambda_{pm} f_m + \varepsilon_p
\end{aligned}
\qquad (1)
$$

where $m < p$ and the coefficients λ_{ij} are called loadings. With appropriate assumptions, λ_{ij} indicates the importance of the *j*th factor f_j to the *i*th variable v_i and can be used in the interpretation of f_j.

Considering the common factors, it is assumed that for $j = 1,2,\cdots,m$ $\mathrm{E}(f_j) = 0$, $\mathrm{var}(f_j) = 1$ and $\mathrm{cov}(f_j, f_k) = 0, j \neq k$. The assumption for the residuals ε_i, $i = 1,2,\cdots,p$ are similar, except that each ε_i, must have a different variance. It is assumed that $\mathrm{E}(\varepsilon_i) = 0$, $\mathrm{var}(\varepsilon_i) = \Psi_i$ and $\mathrm{cov}(\varepsilon_i, \varepsilon_k) = 0, i \neq k$. In addition, we assume that $\mathrm{cov}(\varepsilon_i, f_j) = 0$ for all *i* and *j*. Ψ is referred to as specific variance. Also $\mathrm{E}(v_i - \mu_i) = 0$. From the assumptions above, the variance of *v* can be express as

$$\mathrm{var}(v_i) = \lambda_{i1}^2 + \lambda_{i2}^2 + \cdots + \lambda_{im}^2 + \boldsymbol{\Psi}_i \qquad (2)$$

According to Rencher (2002), the emphasis in factor analysis is on modeling the covariances or correlations among the *v*'s. Model (3.1) can be written in matrix notation as

$$\mathbf{v} - \boldsymbol{\mu} = \boldsymbol{\Lambda}\mathbf{f} + \boldsymbol{\varepsilon} \qquad (3)$$

where $\mathbf{v} = (\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_\mathbf{p})'$, $\boldsymbol{\mu} = (\mu_1, \mu_2, \cdots, \mu_p)'$, $\mathbf{f} = (f_1, f_2, \cdots, f_m)'$, $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_p)'$ and

$$\boldsymbol{\Lambda} = \begin{pmatrix} \lambda_{11} & \lambda_{12} & \cdots & \lambda_{1m} \\ \lambda_{21} & \lambda_{22} & \cdots & \lambda_{2m} \\ \vdots & \vdots & & \vdots \\ \lambda_{p1} & \lambda_{p2} & \cdots & \lambda_{pm} \end{pmatrix} \qquad (4)$$

The assumptions listed between (1) and (2) can be expressed concisely using vector and matrix notation:

$$\mathrm{E}(\mathbf{f}) = \mathbf{0} \qquad \mathrm{cov}(\mathbf{f}) = \mathbf{I} \qquad (5)$$

$$\mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0}, \quad \mathrm{cov}(\boldsymbol{\varepsilon}) = \boldsymbol{\Psi} = \begin{pmatrix} \Psi_1 & 0 & \cdots & 0 \\ 0 & \Psi_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \Psi_P \end{pmatrix}, \qquad (6)$$

$$\mathrm{cov}(\mathbf{f}, \boldsymbol{\varepsilon}) = \mathbf{O}$$

The notation $\mathrm{cov}(\mathbf{f}, \boldsymbol{\varepsilon}) = \mathbf{O}$ indicates a rectangular matrix containing the covariances of the *f*'s with ε's. The covariance matrix $\boldsymbol{\Sigma}$ can be expressed in terms of $\boldsymbol{\Lambda}$ and $\boldsymbol{\Psi}$ using the model (3) and the assumptions in (5) and (6).

Since $\boldsymbol{\mu}$ does not affect variances and covariances of $\mathbf{v}$, we have, from (3), $\boldsymbol{\Sigma} = \text{cov}(\mathbf{v}) = \text{cov}(\boldsymbol{\Lambda}\mathbf{f} + \boldsymbol{\varepsilon})$

Since $\boldsymbol{\Lambda}\mathbf{f}$ and $\boldsymbol{\varepsilon}$ are uncorrelated, it implies that, the covariance matrix of their sum is the sum of their covariance matrices:

$$\boldsymbol{\Sigma} = \text{cov}(\boldsymbol{\Lambda}\mathbf{f}) + \text{cov}(\boldsymbol{\varepsilon})$$

$$= \boldsymbol{\Lambda}\,\text{cov}(\mathbf{f})\boldsymbol{\Lambda}' + \boldsymbol{\Psi}$$

$$= \boldsymbol{\Lambda}\mathbf{I}\boldsymbol{\Lambda}' + \boldsymbol{\Psi}$$

$$= \boldsymbol{\Lambda}\boldsymbol{\Lambda}' + \boldsymbol{\Psi} \qquad (7)$$

In (2), the variance of v_i was partitioned into a component due to common factors, called the **communality** and a component unique to v_i, called the **specific variance**:

$$\sigma_{ii} = \text{var}(v_i) = \left(\lambda_{i1}^2 + \lambda_{i2}^2 + \cdots + \lambda_{im}^2\right) + \Psi_i$$

$$= h_i^2 + \Psi_i$$

where

$$h_i^2 = \lambda_{i1}^2 + \lambda_{i2}^2 + \cdots + \lambda_{im}^2 \qquad (8)$$

is the communality of variables with the factors extracted and ψ_i is the specific variance. The communality h_i^2 is also referred to as common variance, and the specific variance Ψ_i has been called specificity, unique variance or residual variance.

PRINCIPAL COMPONENT METHOD

Once it has been determined that factor analysis is an appropriate technique for analyzing a data, an appropriate method must be selected. In principal component factoring, which is the approach used in this study, the total variance in the data is considered. The diagonal of the correlation matrix consist of unities and full variance is brought into the factor matrix. Principal component method of factoring is recommended when the primary concern is to determine the minimum number of factors that will account for maximum variance in the data for use in subsequent multivariate analysis.

From the random sample $v_1, v_2, \cdots, v_n$, the sample covariance matrix $\mathbf{S}$ is obtained and an attempt is made to find an estimator $\hat{\mathbf{\Lambda}}$ that will approximate the fundamental expression in (7) with $\mathbf{S}$ in place of $\mathbf{\Sigma}$:

$$\mathbf{S} \cong \hat{\mathbf{\Lambda}}\hat{\mathbf{\Lambda}}' + \hat{\mathbf{\Psi}} \tag{9}$$

In principal component method, $\hat{\mathbf{\Psi}}$ is neglected and $\mathbf{S}$ is factored into $\mathbf{S} = \hat{\mathbf{\Lambda}}\hat{\mathbf{\Lambda}}'$. In order to factor **S,** spectral decomposition is used in

$$\mathbf{S} = \mathbf{CDC}' \tag{10}$$

where $\mathbf{C}$ is an orthogonal matrix constructed with normalized eigenvectors $(\mathbf{c}_i'\mathbf{c}_i = 1)$ of $\mathbf{S}$ as columns and $\mathbf{D}$ is a diagonal matrix with the eigenvalues $\theta_1, \theta_2, \cdots, \theta_p$ of $\mathbf{S}$ on the diagonal:

$$\mathbf{D} = \begin{pmatrix} \theta_1 & 0 & \cdots & 0 \\ 0 & \theta_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \theta_p \end{pmatrix} \tag{11}$$

To finish factoring $\mathbf{CDC}'$ in (10) into the form $\hat{\mathbf{\Lambda}}\hat{\mathbf{\Lambda}}'$, we can factor **D** into $\mathbf{D} = \mathbf{D}^{1/2}\mathbf{D}^{1/2}$, since the eigenvalues θ_i of the positive semidefinite matrix **S** are all positive or zero.

$$\mathbf{D}^{1/2} = \begin{pmatrix} \sqrt{\theta_1} & 0 & \cdots & 0 \\ 0 & \sqrt{\theta_2} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \sqrt{\theta_p} \end{pmatrix}$$

With this factoring of **D**, expression (10) becomes

$$\mathbf{S} = \mathbf{CDC}' = \mathbf{CD}^{1/2}\mathbf{D}^{1/2}\mathbf{C}'$$

$$= (\mathbf{CD}^{1/2})(\mathbf{CD}^{1/2})' \qquad (12)$$

This is of the form $\mathbf{S} = \hat{\mathbf{\Lambda}}\hat{\mathbf{\Lambda}}'$, but we do not define $\hat{\mathbf{\Lambda}}$ to be $\mathbf{CD}^{1/2}$ because $\mathbf{CD}^{1/2}$ is $p \times p$, and we are seeking a $\hat{\mathbf{\Lambda}}$ that is $p \times m$ with $m < p$. Defining $\mathbf{D}_1 = \mathrm{diag}(\theta_1, \theta_2, \cdots, \theta_m)$ with the m largest eigenvalues $\theta_1 > \theta_2 > \cdots > \theta_m$ and $\mathbf{C}_1 = (\mathbf{c}_1, \mathbf{c}_2, \cdots, \mathbf{c}_m)$ containing the corresponding eigenvectors. We then estimate $\mathbf{\Lambda}$ by the first m columns of $\mathbf{CD}^{1/2}$,

$$\hat{\mathbf{\Lambda}} = \mathbf{C}_1\mathbf{D}_1^{1/2} = \left(\sqrt{\theta_1}\mathbf{c}_1, \sqrt{\theta_2}\mathbf{c}_2, \cdots, \sqrt{\theta_m}\mathbf{c}_m\right) \qquad (13)$$

where $\hat{\mathbf{\Lambda}}$ is $p \times m$, $\mathbf{C}_1$ is $p \times m$, $\mathbf{D}_1^{1/2}$ is $m \times m$, θ_i are the eigenvalues and $\mathbf{c}_i$ are the eigenvectors.

The *i*th diagonal element of $\hat{\mathbf{\Lambda}}\hat{\mathbf{\Lambda}}'$ is the sum of squares of the *i*th row of $\hat{\mathbf{\Lambda}}$, or $\hat{\boldsymbol{\lambda}}_i'\hat{\boldsymbol{\lambda}}_i = \sum_{j=1}^{m} \hat{\lambda}_{ij}^2$. Hence to complete the approximation of $\mathbf{S}$ in (9), we define

$$\Psi_I = s_{ii} - \sum_{j=1}^{m} \hat{\lambda}_{ij}^2 \qquad (14)$$

and write

$$\mathbf{S} \cong \hat{\mathbf{\Lambda}}\hat{\mathbf{\Lambda}}' + \hat{\mathbf{\Psi}} \qquad (15)$$

where $\hat{\mathbf{\Psi}} = \mathrm{diag}(\Psi_1, \Psi_2, \cdots, \Psi_p)$. Thus in (15) the variances on the diagonal of **S** are modeled exactly, but the off-diagonal covariances are only approximate.

From (8):

$$\hat{h}_i^2 = \sum_{j=1}^{m} \hat{\lambda}_{ij}^2 \tag{16}$$

which is the sum of squares of the *i*th row of $\hat{\mathbf{\Lambda}}$. The sum of squares of the *j*th column of $\hat{\mathbf{\Lambda}}$ is the *j*th eigenvalue of $\mathbf{S}$:

$$\sum_{i=1}^{p} \hat{\lambda}_{ij}^2 = \sum_{i=1}^{p} \left(\sqrt{\theta_j}\, c_{ij}\right)^2$$

$$= \theta_j \sum_{i=1}^{p} c_{ij}^2$$

$$= \theta_j \tag{17}$$

since the normalized eigenvectors (column of **C**) have length one.

By expressions (14) and (16), the variance of the *i*th variable is partitioned into a part due to the factors and a part due uniquely to the variables:

$$s_{ii} = \hat{h}_i^2 + \hat{\Psi}_i$$

$$= \hat{\lambda}_{i1}^2 + \hat{\lambda}_{i2}^2 + \cdots + \hat{\lambda}_{im}^2 + \hat{\Psi}_i \tag{18}$$

Thus the *j*th factor contributes $\hat{\lambda}_{ij}^2$ to s_{ii}. The contribution of the *j*th factor to the total sample variance, $\mathrm{tr}(\mathbf{S}) = s_{11} + s_{22} + \cdots + s_{pp}$, is therefore,

$$\text{Variance due to } j\text{th factor} = \sum_{i=1}^{p} \hat{\lambda}_{ij}^2 = \hat{\lambda}_{1j}^2 + \hat{\lambda}_{2j}^2 + \cdots + \hat{\lambda}_{pj}^2, \tag{19}$$

which is the sum of squares of loadings in the *j*th column of $\hat{\mathbf{\Lambda}}$. By (17), this is equal to the *j*th eigenvalue, θ_j. The proportion of total sample variance due to the *j*th factor is therefore,

$$\frac{\sum_{i=1}^{p}\hat{\lambda}_{ij}^{2}}{\text{tr}(\mathbf{S})} = \frac{\theta_j}{\text{tr}(\mathbf{S})} \tag{20}$$

If the variables are not commensurate, we can use standardized variables and work with the correlation matrix **R**. The eigenvalues and eigenvectors of **R** are then used in place of **S** in (13) to obtain estimates of the loadings. Since the emphasis in factor analysis is on reproducing the covariances or correlations rather than the variances, use of the correlation matrix **R** is more appropriate. In applications, **R** often gives better results than **S**. If we are factoring **R**, the proportion corresponding to (20) is

$$\frac{\sum_{i=1}^{p}\hat{\lambda}_{ij}^{2}}{\text{tr}(\mathbf{R})} = \frac{\theta_j}{p} \tag{21}$$

where p is the number of variables.

We can assess the fit of the factor analysis model by comparing the left and right sides of (15). The error matrix $\mathbf{E} = \mathbf{S} - \left(\hat{\mathbf{\Lambda}}\hat{\mathbf{\Lambda}}' + \hat{\mathbf{\Psi}}\right)$ has zeros on the diagonal but nonzero off-diagonal elements. If the eigenvalues are small, the residuals in the error matrix are small and the fit is good.

APPROPRIATENESS OF THE FACTOR ANALYSIS MODEL

For factor analysis to be appropriate, the variables must be correlated. Also, to decide whether or not the data are appropriate for factor analysis, a number of measures are used for this purpose.

As suggested by Sharma (1996), one can subjectively examine the correlation matrix. High correlations among the variables indicate that the variables can be grouped into homogeneous sets of variables such that each set of variables measure the same underlying factors. Low correlations among the variables signify that the variables do not have much in common or are a group of heterogeneous variables. In this sense, one could view factor analysis as a technique that tries to identify groups of variables such that variables in each group are indicators of a common trait or factor. This suggests that the correlation matrix is appropriate for factoring. However, visual examination of a correlation matrix for a large number of variables is almost impossible.

Formal statistics available for testing the appropriateness of the factor model includes **Bartlett's test of sphericity.** The statistic can be used to test the null hypothesis that the variables are uncorrelated in the population. In other words, the population correlation matrix is an identity matrix. In an identity matrix, all the diagonal terms are 1, and all the off-diagonal terms are 0. The test statistic for sphericity is based on a chi-square transformation of the determinant of the correlation matrix. A large value of the test statistic will favor the rejection of the null hypothesis. If the null hypothesis cannot be rejected, then the appropriateness of factor analysis should be questioned. The test statistic is given by;

$$\chi^2 = -2\left\{1 - \frac{1}{6pn}\left(2p^2 + p + 2\right)\right\} \text{In}\,\Lambda \tag{22}$$

where

$$\Lambda = \left\{\frac{\text{Geometric mean } \theta_i}{\text{Arithmetic mean } \theta_i}\right\}^{\frac{n}{2}} = \left\{\frac{\prod_{i=1}^{p}\theta_i}{\left(\frac{1}{p}\sum_{i=1}^{p}\theta_i\right)^p}\right\}^{\frac{n}{2}} \tag{23}$$

where $\theta_i\left(i = 1,2,\cdots,p\right)$ are the eigenvalues of the component factors, p is the number of indicator variables and *n* is the sample size. The test statistic has a chi-square distribution with degrees of freedom equal to $\frac{1}{2}(p-1)(p+2)$. At an α–level of significance, a large value of χ^2 than the corresponding table value of $\chi^2_{\alpha,\frac{1}{2}(p-1)(p+2),}$ leads to the rejection of the null hypothesis. From equation (22), it can be observed that the statistic is greatly dependent on the sample size (*n*). If *n* is very large, the value of $\text{In}\,\Lambda$ causes the statistic to be large. This makes results of the Bartlett's test of sphericity highly dependent on the sample size.

Another useful statistic is the **Kaiser-Meyer-Olkin (KMO)** measure of sampling adequacy. This measure is a popular diagnostic measure which provides a means to assess the extent to which the indicators of a construct belong together. Thus, the index compares the magnitude of the observed correlation coefficients to the magnitude of the partial correlation coefficients. Small values of the KMO statistic indicate that the correlation between pairs of variables cannot be explained by other variables and that factor analysis may not be appropriate. Table 2 gives a guideline for the KMO measure (Sharma, 1996).

Table 2: A Guide for Interpreting KMO Measure

KMO Measure	Recommendation
≥ 0.90	Marvelous
0.80+	Meritorious
0.70+	Middling
0.60+	Mediocre
0.50+	Miserable
Below 0.50	Unacceptable

To have satisfactory results, it is expected that the overall KMO measure should exceed 0.8 although a measure of above 0.6 is acceptable (Rencher, 2002). An equation which could be used to measure the sampling adequacy (MSA) is given by:

$$\text{MSA} = \frac{\sum_{i \neq j} r_{ij}^2}{\sum_{i \neq j} r_{ij}^2 + \sum_{i \neq j} q_{ij}^2} \qquad (24)$$

where r_{ij}^2 is the square of an element from **R** and q_{ij}^2 is the square of an element from $\mathbf{Q} = \mathbf{D}\mathbf{R}^{-1}\mathbf{D}$, with $\mathbf{D} = \left[(\text{diag } \mathbf{R}^{-1})^{1/2} \right]^{-1}$. As $\mathbf{R}^{-1}$ approaches a diagonal matrix, MSA approaches one.

In summary, there are many data sets to which factor analysis should not be applied. If the scree plot does not have a pronounced bend or the eigenvalues do not show a large gap around one, then **R** is likely to be unsuitable for factoring. In addition, the communality estimates after factoring should be fairly large.

CHOOSING THE NUMBER OF FACTORS

There are so many criteria for choosing the number of factors. Four of these criteria will be considered in this study.

First, we choose m to be equal to the number of factors necessary for the variance accounted for to achieve a predetermined percentage of the total variance $\text{tr}(\mathbf{S})$ or $\text{tr}(\mathbf{R})$. This method applies particularly to the principal component technique of factoring. By equation (20), the proportion of total sample variance due to the jth factor from $\mathbf{S}$ is $\frac{\sum_{i=1}^{p}\hat{\lambda}_{ij}^2}{\text{tr}(\mathbf{S})}$. The corresponding proportion from $\mathbf{R}$ is $\frac{\sum_{i=1}^{p}\hat{\lambda}_{ij}^2}{p}$ as in equation (21). The contribution of all m factors to $\text{tr}(\mathbf{S})$ or p is therefore $\sum_{i=1}^{p}\sum_{j=1}^{m}\hat{\lambda}_{ij}^2$ which is the sum of squares of all elements of $\hat{\mathbf{\Lambda}}$. This sum of squares is also equal to the sum of the first m eigenvalues or to the sum of all p communalities:

$$\sum_{i=1}^{p}\sum_{j=1}^{m}\hat{\lambda}_{ij}^2 = \sum_{i=1}^{p}\hat{h}_i^2 = \sum_{j=1}^{m}\theta_j \qquad (25)$$

Thus we choose m sufficiently large so that the sum of the communalities or the sum of the eigenvalues (variance accounted for) constitutes a relatively large portion of $\text{tr}(\mathbf{S})$ or p.

The second method considered in this study is the determination of *m* based on the eigenvalues. In this approach, only factors with eigenvalues greater than 1.0 are retained. Factors with variance less than 1.0 are no better than a single variable, because, due to standardization, each variable has a variance of 1.0. The eigenvalue represents the total variance explained by each factor. The eigenvalue-greater-than-one rule is the most popular heuristics and is the default in many statistical software packages. It has been observed by Naresh (2004) that if the number of variables is less than 20, this approach will results in a conservative number of factors.

Thirdly, the use of the scree test based on the plot of the eigenvalues of **S** or **R** is the next method considered. The scree test was named after the geological term scree, referring to the debris at the bottom of a rocky cliff Rencher (2002). In a scree plot, the eigenvalues are plotted against the number of factors in order of extraction. The shape of the plot is used to determine the number of factors. If the graph drops sharply, followed by a straight line with much smaller slope, then we choose *m* equal to the number of components before the straight line begins. Usually, the graph has a distinct break between the steep slope of factors, with large eigenvalues and a gradual trailing off associated with the rest of the factors. This gradual trailing off is referred to as the scree. Zwick and Velicer (1986) found out in their simulation studies that the scree plot is one of the best-performing rules.

Lastly, we consider the test of the hypothesis that *m* is the correct number of factors. Thus, we wish to test

$$H_0 : \mathbf{\Sigma} = \mathbf{\Lambda\Lambda}' + \mathbf{\Psi} \qquad \text{against} \qquad H_1 : \mathbf{\Sigma} \neq \mathbf{\Lambda\Lambda}' + \mathbf{\Psi}$$

where $\mathbf{\Lambda}$ is $p \times m$. The test statistic, a function of the likelihood ratio, is given by:

$$\chi^2 = \left(\mathrm{n} - \frac{2\mathrm{p} + 4\mathrm{m} + 11}{6} \right) \mathrm{In} \left(\frac{\left| \hat{\mathbf{\Lambda}}\hat{\mathbf{\Lambda}}' + \hat{\mathbf{\Psi}} \right|}{|\mathbf{S}|} \right) \tag{26}$$

which is approximately χ_d^2 when H_0 is true, where $d = \frac{1}{2}\left[(p-m)^2 - p - m\right]$ and $\hat{\mathbf{\Lambda}}$ and $\hat{\mathbf{\Psi}}$ are the maximum likelihood estimates. Rejection of H_0 implies that *m* is too small and more factors are needed.

The choice of the number of factors, *m* for many data sets will not be obvious. When a data set is appropriate for factoring, the first three methods will almost always give the same value of *m*, and there will be little question as to what this value should be.

FACTOR ROTATION

Rotation is one of the important aspects of factor analysis. The main objective of rotation is to transform the factor matrix into a simpler one that can be meaningfully interpreted. In rotating the factors, we would like each factor to have nonzero, or significant, loadings or coefficients for only some of the variables. Also, the variables should have nonzero or significant loadings with only few factors, if possible with only one. If several factors have high loadings with the same variable, it is difficult to interpret them. Rotation does not affect the communalities and the percentage of total variance explained.

However, the percentage of variance accounted for by each factor does change. The estimated loading matrix $\hat{\boldsymbol{\Lambda}}$ can likewise be rotated to obtain

$$\hat{\boldsymbol{\Lambda}}^{*} = \hat{\boldsymbol{\Lambda}}\mathbf{T} \qquad (27)$$

where $\mathbf{T}$ is orthogonal. Since $\mathbf{TT}' = \mathbf{I}$, the rotated loadings provide the same estimates of the covariance as before:

$$\mathbf{S} \cong \hat{\boldsymbol{\Lambda}}^{*}\hat{\boldsymbol{\Lambda}}^{*\prime} + \hat{\boldsymbol{\Psi}} = \hat{\boldsymbol{\Lambda}}\mathbf{TT}'\hat{\boldsymbol{\Lambda}}' + \hat{\boldsymbol{\Psi}} = \hat{\boldsymbol{\Lambda}}\hat{\boldsymbol{\Lambda}}' + \hat{\boldsymbol{\Psi}} \qquad (28)$$

If a rotation in which every point is close to an axis can be achieved, then each variable loads highly on the factor corresponding to the axis and has small loadings on the remaining factors. Consequently, mere observations are made on which variables are associated with each factor and the factor is named accordingly.

To achieve the objective of rotation, two types of rotation can be performed: namely orthogonal and oblique. In orthogonal rotation, the factors are constrained to be uncorrelated whilst in oblique rotation the factors are allowed to intercorrelate. A rotation is said to be orthogonal if the axes are maintained at right angles. Orthogonal rotations preserve communalities. This is because the rows of $\hat{\boldsymbol{\Lambda}}$ are rotated and the distance to the origin is unchanged. The rotation in equation (28) involving an orthogonal matrix is an orthogonal rotation. Field (2005) recommended the use of orthogonal rotation if the factors are theoretically independent (unrelated to each other). The most popular types of orthogonal rotations are varimax and quartimax rotations. The oblique rotations are not used in this study, hence they are not discussed.

Varimax Rotation

The varimax rotation seeks rotated loadings that minimize the variance of the squared loadings in each column of $\hat{\Lambda}^*$. Thus, varimax rotation minimizes the number of variables with high loadings on a factor, thereby enhancing the interpretability of the factors. The major objective of varimax rotation is to have a factor structure in which each variable loads highly on one and only one factor. That is, a given variable should have a high loading on one factor and near zero loadings on other factors. This is achieved by maximizing the variance of the squared loadings across variables, subjects to the constraint that the communality of each variable is unchanged. That is, for any factor,

$$C_j = \frac{\left(\sum_{i=1}^{p} \lambda_{ij}^2 - \lambda_j^2\right)^2}{p}$$

$$= \frac{p\sum_{i=1}^{p} \lambda_{ij}^4 - \left(\sum_{i=1}^{p} \lambda_{ij}^2\right)^2}{p^2} \qquad (29)$$

where

C_j is the variance of communality of the variables within factor *j* and

$\lambda_{.j}^2$ is the average squared loading for factor *j*.

The total variance for all the factors is given by:

$$C^* = \sum_{j=1}^{m} C_j$$

$$= \sum_{j=1}^{m} \left[\frac{p\sum_{i=1}^{p} \lambda_{ij}^{4} - \left(\sum_{i=1}^{p} \lambda_{ij}^{2} \right)^{2}}{p^{2}} \right]$$

$$= \frac{\sum_{j=1}^{m}\sum_{i=1}^{p} \lambda_{ij}^{4}}{p} - \frac{\sum_{j=1}^{m}\left(\sum_{i=1}^{p} \lambda_{ij}^{2} \right)^{2}}{p^{2}} \qquad (30)$$

Since the number of variables remains the same, maximizing equation (30) is the same as maximizing,

$$pC^{*} = \sum_{j=1}^{m}\sum_{i=1}^{p} \lambda_{ij}^{4} - \frac{\sum_{j=1}^{m}\left(\sum_{i=1}^{p} \lambda_{ij}^{2} \right)^{2}}{p} \qquad (31)$$

The varimax procedure cannot guarantee that all variables will highly load on only one factor. This type of rotation is available in virtually all factor analysis software programmes. The output typically includes the rotated loading matrix $\hat{\mathbf{\Lambda}}^{*}$, the variance accounted for (sum of squares of each column of $\hat{\mathbf{\Lambda}}^{*}$), the communalities (sum of squares of each row of $\hat{\mathbf{\Lambda}}^{*}$) and the orthogonal matrix $\mathbf{T}$ used to obtain $\hat{\mathbf{\Lambda}}^{*} = \hat{\mathbf{\Lambda}}\mathbf{T}$.

Quartimax Rotation

The major objective of this type of rotation is to obtain a pattern of loadings such that:

1. All the variables have a fairly high loading on one factor
2. Each variable should have a high loading on one other factor and near zero loadings on the remaining factors.

This objective is realized by maximizing the variance of the loadings across factors, subject to the constraint that the communality of each variable is unchanged. Thus, suppose for any given variable *i*, we define

$$Q_i = \frac{\sum_{j=1}^{m}\left(\lambda_{ij}^2 - \lambda_{i.}^2\right)^2}{m} \qquad (32)$$

where Q_i is the variance of the communality (i.e. square of the loadings) of variable *i*, λ_{ij}^2 is the squared loading of the *i*th variable on the *j*th factor, $\lambda_{i.}^2$ is the average squared loading of the *i*th variable and *m* is the number of factors.

By expanding and substituting $\lambda_{i.}^2 = \frac{\sum_{j=1}^{m}\lambda_{ij}^2}{m}$ into equation (32), it can be rewritten as:

$$Q_i = \frac{m\sum_{j=1}^{m}\lambda_{ij}^4 - \left(\sum_{j=1}^{m}\lambda_{ij}^2\right)^2}{m^2} \qquad (33)$$

The total variance of all the variables is given by:

$$Q^* = \sum_{i=1}^{p} Q_i = \sum_{i=1}^{p}\left[\frac{m\sum_{j=1}^{m}\lambda_{ij}^4 - \left(\sum_{j=1}^{m}\lambda_{ij}^2\right)^2}{m^2}\right] \qquad (34)$$

Once the initial factor solution has been obtained, the number of factors, *m*, remains constant. Furthermore, the second term in equation (34), $\sum_{j=1}^{m}\lambda_{ij}^2$, is the communality of the variable and hence, it will also be constant. Therefore, maximization of equation (32) reduces to maximizing the equation:

$$Q^* = \sum_{i=1}^{p}\sum_{j=1}^{m}\lambda_{ij}^4 \qquad (35)$$

Quartimax rotation is the most appropriate when the presence of a general factor is suspected. The quartimax rotation gives an interpretation of the factor structure similar to that of varimax rotation. The only difference is the fact that varimax rotation suppresses the general factor and should not be used when the presence of a general factor is suspected. The communality estimates of the variables are not affected. In most cases, prior to performing rotation, the loadings of each variable are normalized by dividing the loading of each variable by the total communality of the respective variable.

COMPUTATION AND INTERPRETATION OF FACTOR SCORES

Factor scores are used for various purposes such as to serve as proxies for latent variables and to determine a participant's relative standing on the latent dimension. Conceptually, a factor score is the score that would have been observed for a person if it had been possible to measure the latent factor directly. Factor scores can be calculated if necessary. If the goal of factor analysis is to reduce the original set of variables to a smaller set of composite variables (factors) for use in subsequent multivariate analysis, it is useful to compute factor scores for each respondent. The factor score for the *i*th factor may be estimated as follows:

$$\mathrm{F}_i = \mathrm{w}_{i1}\mathrm{V}_1 + \mathrm{w}_{i2}\mathrm{V}_2 + \mathrm{w}_{i3}\mathrm{V}_3 + \cdots + \mathrm{w}_{ip}\mathrm{V}_p$$

or $$\mathrm{F}_j = \sum_{j=1}^{p}\mathrm{w}_{ij}\mathrm{V}_i \qquad (36)$$

F_i = estimate of *i*th factor,

w_i = weight or factor score coefficient,

V_i = *i*th indicator variable

p = the number of variables.

The data is about opinions of some Senior High School students on their choice of the University of Cape Coast. Any observed value is 1, 2, 3, 4 or 5. Clearly, the value of this expression in Equation (36) would be influenced by the observed values of the variables, V_i. The variability in the values of these variables would then be reflected in the factor scores. This can influence the interpretation of the scores. To eliminate the effect of the variation in the individual variables on interpretation, we first standardized the data. Thus, the factor score f_i corresponding to F_i is given by

$$f_i = \sum_{j-1}^{p} w_{ij} \frac{V_j - \mu_j}{s_j} \tag{37}$$

The standardized process determines the magnitude and the sign of the *j*th term in equation (37). Assuming that all the weights, w_{ij}, are positive, we consider three typical scenarios: f_i could be a high positive value, a high negative value or close to zero.

A high positive score is obtained if the values of the items $(V_j, j = 1,2,\ldots, p)$ are higher than the average values of the respective items. Thus, in relation to the data used for this study, a high positive score, f_i, indicates that the respondent indicated scores on all the items that are consistently much higher than the average score of the various items. Such a respondent is strongly influenced by that factor, since he/she strongly agreed or agreed to most of the attributes.

A high negative score is obtained if the values of most of the items $(V_j, j = 1,2,\dots, p)$ are lower than the average values of the respective items. In reference to the data used for this study, a high negative score, f_i, indicates that the respondent indicated scores on all the items that are consistently much lower than the average score of the respective items. Such a respondent is least influenced by that factor, since he/she strongly disagreed or disagreed to most of the attributes.

A very small score (close to zero) is obtained if the values of most of the items are just about the same as the average values of the respective items. Thus, in relation to the data used for this study, a small value of f_i indicates that the respondent consistently obtained scores that are about the same value as the mean score in each of the items. Such a respondent is moderately influenced by that factor, since he/she did not express extreme opinions on most of the attributes.

The above discussion is basically on the effects of standardization of the data on the sign and size of the factor score assuming that the weights are all positive. However, the magnitude of the score is also determined to a large extent by the size of the weight, w_{ij}, on the variables V_j.

The factor score coefficients, used to combine the standardized variables are obtained from the factor score coefficient matrix. The factor scores can be found on most computer programmes. Due to factor indeterminacy problem, a number of loadings are possible, each resulting in a separate set of factor scores. In other words, the factor scores are not unique.

For this reason many researchers hesitate to use the factor scores in further analysis.

The factor scores would be classified as strongly influenced, moderately influenced and least influenced. Respondents with scores above the 75^{th} percentile would be classified as strongly influenced whiles those with scores below the 25^{th} percentile are considered as least influenced. Respondents with scores ranging from 25^{th} to 75^{th} percentile would be classified as moderately influenced. These ranges are used because there are enough observations within each range and the sample size is reasonably large.

CHAPTER FOUR

PRELIMINARY ANALYSIS

In this chapter, exploratory methods are used to investigate fundamental characteristics of the variables used in the study. Exploratory methods such as bar charts, correlation analysis, analysis of the eigenvalues and analysis based on scree plots are discussed. The bar charts for the various indicator variables are grouped based on the pattern of responses. These groups would then be compared with the groups that would be obtained in the correlation analysis. The eigenvalue analysis and the scree plot analysis of the indicators would also give an indication of how many factors to extract.

FREQUENCY DISTRIBUTION OF RESPONSES

The focus in this section is to use bar charts to find out if the indicator variables can be grouped based on the pattern of responses. The anticipation is that the result obtained in this section would emphasize the outcome of subsequent sections. The bar charts provide a quick summary of the distribution of the five levels of agreement. These levels are: strongly disagree (1), disagree (2), undecided (3), agree (4) and strongly agree (5). All the twenty eight indicator variables have been presented on separate bar charts to illustrate the pattern of responses of the students. The percentage frequencies are plotted against the levels of agreement on the variables. The frequency distribution tables for these bar charts are shown in Appendix II.

The percentage bar charts indicating the views of respondents on variables V_1 (the only university that offers my choice of programme) and V_2 (It is located in Central Region where there are many tourist attractions) are presented in Figure 1.

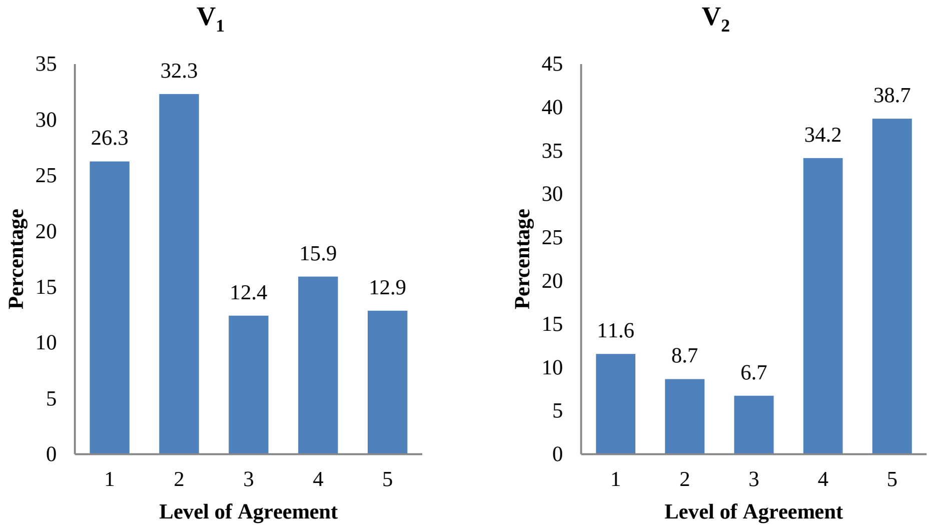

Figure 1: Percentage Distribution of Opinions on V_1 and V_2

V_1: The only university that offers my choice of programme

From the figure, it can be observed that the distribution on V_1 is positively skewed. Thus, this indicator does not generally influence applicants' choice of UCC. A little below one third (32 percent) of the respondents disagree that UCC is the only university that offers their choice of programme. About 12 percent would not mind whether UCC offers their choice of programme or not and 2 people did not give their opinion on this indicator. In all, only 29 percent of the respondents agree (or strongly agree) that the University of Cape Coast is the only university that offers their choice of programme. Most (59 percent) of the respondents disagree (or strongly disagree) on this issue. This means that most of the students do not consider this factor to be important in the choice UCC.

V_2: It is located in Central Region where there are many tourist attractions

It can be seen from Figure 1 that quite a large number (about 39 percent) of the respondent strongly agree that their choice of the University of Cape Coast is based on its location in Central Region where there are many tourist attractions. Only a few of the respondents (7 percent) could not identify their position on the importance of this issue. As a whole, 79 out of every 100 respondents either agree or strongly agree on this attributes. About one fifth of the respondents (20 percent) disagrees or strongly disagrees to this attribute. The figure reveals that the pattern of distribution is negatively skewed. Generally, most of the respondents would like to choose the University of Cape Coast because of where it is located.

The bar chats demonstrating the opinions of respondents for the indicator variables V_3, V_4, V_5, V_6, V_7 and V_8 are presented in Figure 2. The comments on each graph are presented in subsequent subsections.

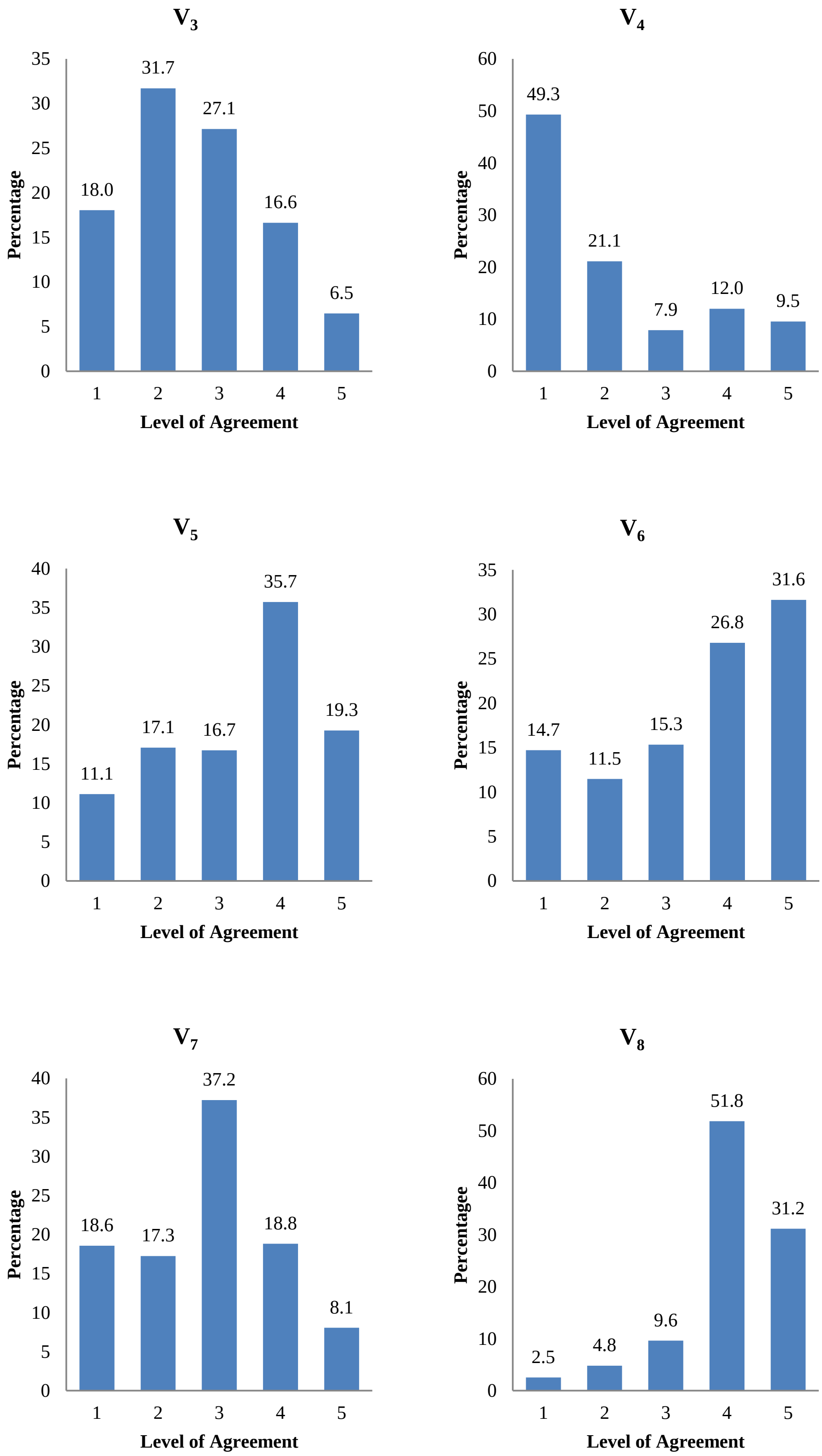

Figure 2: Percentage Distribution of Opinions on V_3, V_4, V_5, V_6, V_7 and V_8

V_3: It is base on my counselor's recommendation

A close examination of Figure 2 reveals that 1 out of every 3 respondent disagree that their choice of UCC was based on their counselor's recommendation. Only 7 out of 100 respondents strongly agree on this subject. The distribution tends to be positively skewed. Approximately half of the respondents disagree (or strongly disagree) that their choice of UCC depends on their counselor's recommendation. A little above one fourth of the respondents (27 percent) were not specific about this statement. Also, 23 out of every 100 respondents claim to base their choice of UCC on their counselor's recommendation.

V_4: The University is close to my place of residence

The pattern of distribution in Figure 2 show that greater part (49 percent) of the respondents might not stay close to the university. Accordingly, 70 percent dispute that the university is close to their place of residence. Very few (8 percent) are not sure and 22 percent stays close to the university. The graph is positively skewed. This shows that most of the respondents would choose the university not because it is close to their place of residence.

V_5: It will be an opportunity to meet old friends

From Figure 2, it can be seen that about 36 out of every 100 respondents agree that their choice of UCC would give them the opportunity to meet old friends. Very few respondents (11 percent) strongly disagree on this issue. A closer look at the figure indicates that the distribution is negatively skewed; hence, this subject attracts many respondents to UCC. Respondents who could not express their views on this subject were below one fifth (17 percent) of the total number. In total, about 28 percent do not believe that they could meet their old friends in UCC; hence they do not base their choice on this reason. Out of every 100 respondent, 55 of them tend to agree or strongly agree on this indicator.

V_6: I prefer the grading system of this university to other universities

Surprisingly, 1 out of every 3 respondents strongly prefers the grading system of UCC to other universities as indicated in Figure 2. A few (11 percent) disagree to this attribute. The graph depicts a negatively skewed distribution which point outs that majority of the respondents (58 percent) prefer the grading system of the University of Cape Coast. A total of 25 percent do not like the grading system of the university. About 15 percent of the students are unable to tell whether they like the grading system or not. Only one person did not respond to this attribute.

V_7: It is cheaper to get accommodation in this university

As shown in Figure 2, the responses on the indicator variable V_7 appears to create doubt in many respondents. More than one third of the respondents (37 percent) could not decide whether accommodation is cheaper in this university or not. A small number of respondents (8 percent) strongly agree on this matter. In effect, about 27 out of every 100 respondents tend to agree or strongly agree on this issue whereas 36 of every 100 respondents disagree or strongly disagree with this issue. Thus, over one third of the respondents maintain that it is not cheaper to get accommodation in UCC. There seems to be large variability on the opinions of the respondents.

V_8: The University offers quality programmes

The distribution on the indicator variable V_8 as shown in Figure 2 indicates that more than half (52 percent) of the respondents agree that the University of Cape Coast offer quality programmes. Very few (3 percent) respondents strongly disagree to this attribute. It is worth mentioning that the distribution is highly skewed negatively. This shows that virtually all (83 percent) the respondents would choose UCC base on the idea that they could get the opportunity to study quality programmes. A little below 10 percent of the respondents could not decide on this factor. Only 7 percent of the students either disagree strongly disagree on this factor. Distribution of opinions on this variable shows evidence of little disparity.

A graphical representation of opinions on the indicator variables V_9, V_{10}, V_{11}, V_{12}, V_{13} and V_{14} are shown in Figure 3. Discussions on each graph are presented in the subsections below.

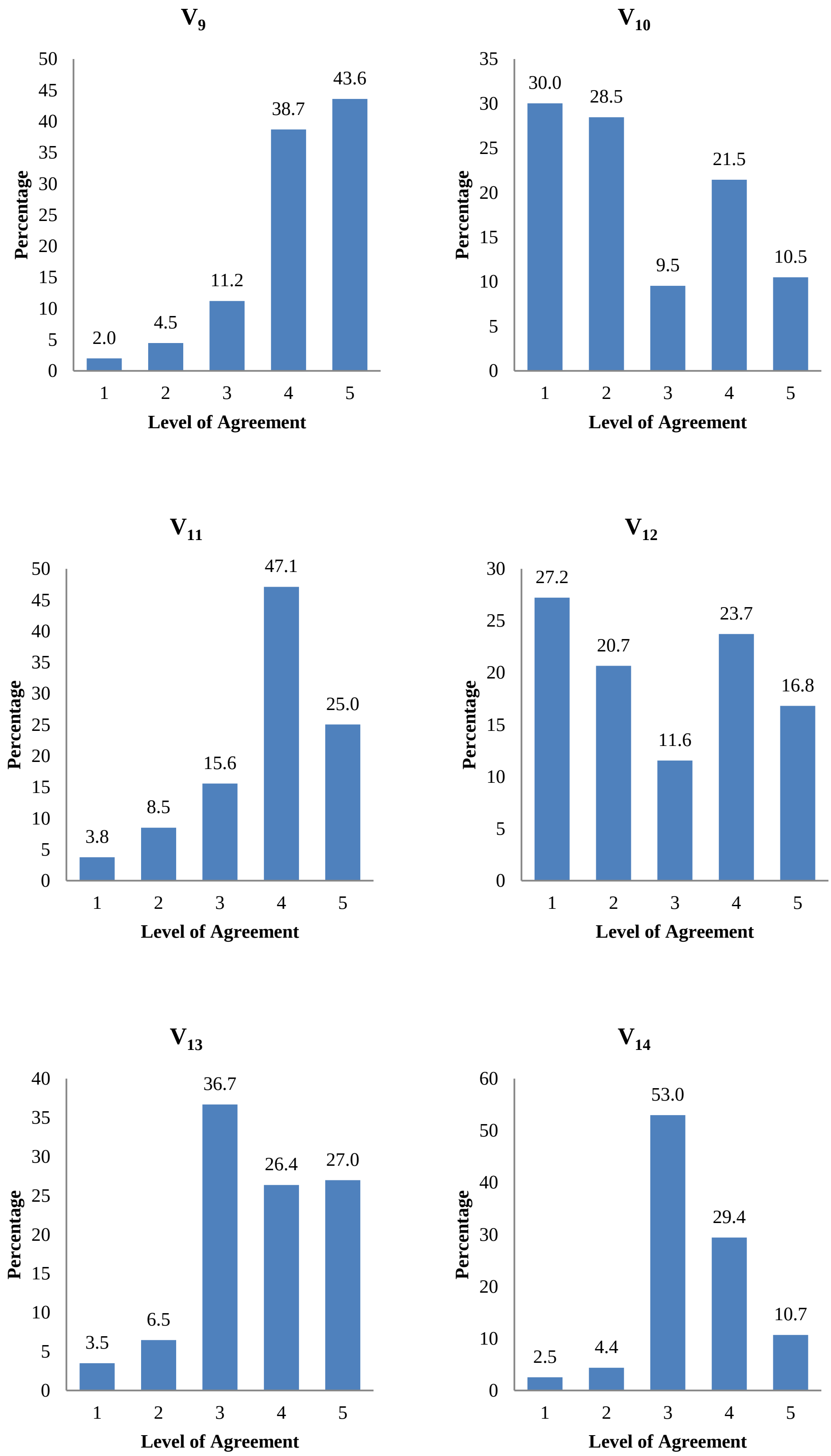

Figure 3: Percentage Distribution of Opinions on V_9, V_{10}, V_{11}, V_{12}, V_{13} and V_{14}

V_9: Students who complete this university are competent

A similar pattern of distribution occurred in the case of the indicator variables V_8 and V_9. Virtually all the respondents (82 percent) either agree or strongly agree that students who complete UCC are competent. Very few respondents (2 percent) strongly disagree that UCC graduates are competent. The distribution shows a high level of negative skewness which means that this item is of much importance to majority of the respondent. Approximately 11 of every 100 respondents could not decide on this attribute.

V_{10}: Most of my friends and family members attended this university

The percentage distribution of opinions on the indicator V_{10} as shown in Figure 3 tends to be positively skewed. Consequently, the figure point out that, more than half of the respondents (58 percent) disagree that their friends and family members attended UCC. Approximately, 1 of every 10 respondent is not sure on this issue. Precisely 32 percent accepted to base their choice on the reason that most of their friends and family members attended UCC. The distribution shows some level of variation in opinions on this indicator variable.

V_{11}: The physical structure of the university is very attractive

Figure 3 on the indicator V_{11} shows that a small number (approximately 4 percent) of respondents strongly disagree with this attributes. A lot more (47 percent) tend to agree on this attributes. The nature of the graph is negatively distributed. The implication is that the majority (72 percent) of respondents are in agreement that the physical structure of UCC is very attractive. As a result, these students would choose to attend UCC because of its attractive physical structure. About 16 of every 100 respondents could not decide whether or not they are attracted to the university because of its physical structure.

V_{12}: I like the university because it is located by the sea

According to Figure 3 on variable V_{12}, there is no clear bias for this indicator variable. About 40 out of every 100 respondents like the University of Cape Coast because it is located by the sea. Almost 12 percent are unable to indicate their views on this attribute. Most of the respondents (48 percent) would not choose the University of Cape Coast because it is located by the sea. The pattern of the distribution shows very large variability in opinion on the variable.

V_{13}: Both students and lecturers are punctual at lectures

It can be observed from Figure 3, that the pattern of distribution for the indicator variable V_{13} is negatively skewed. Only a few (about 4 percent) respondents strongly disagree on this issue. It is not surprising that more than one third of the respondents (37 percent) are unable to determine whether students and lecturers are punctual at lectures or not. Besides, more than half of the respondents (53 percent) believe that both students and lecturers are punctual at lectures. The pattern of the distribution reveals some level of variability in opinions on this indicator variable.

V_{14}: The workers of the university are friendly

Unpredictably, over half of the number (53 percent) could not tell whether the workers of the university are friendly or not as indicated in Figure 3 on variable V_{14}. As much as 40 percent also believe that the workers of the university are friendly. A small number (about 7 percent) of the students disagree with this assertion. The shape of the graph is negatively skewed. This implies respondents generally agree that UCC workers are friendly.

A quick summary of the distribution of the students' level of agreement on the indicator variables V_{15} V_{16} V_{17} V_{18} V_{19} and V_{20} are provided in Figure 4. The successive subsections talk about the pattern of distribution for each indicator variable demonstrated by the graph.

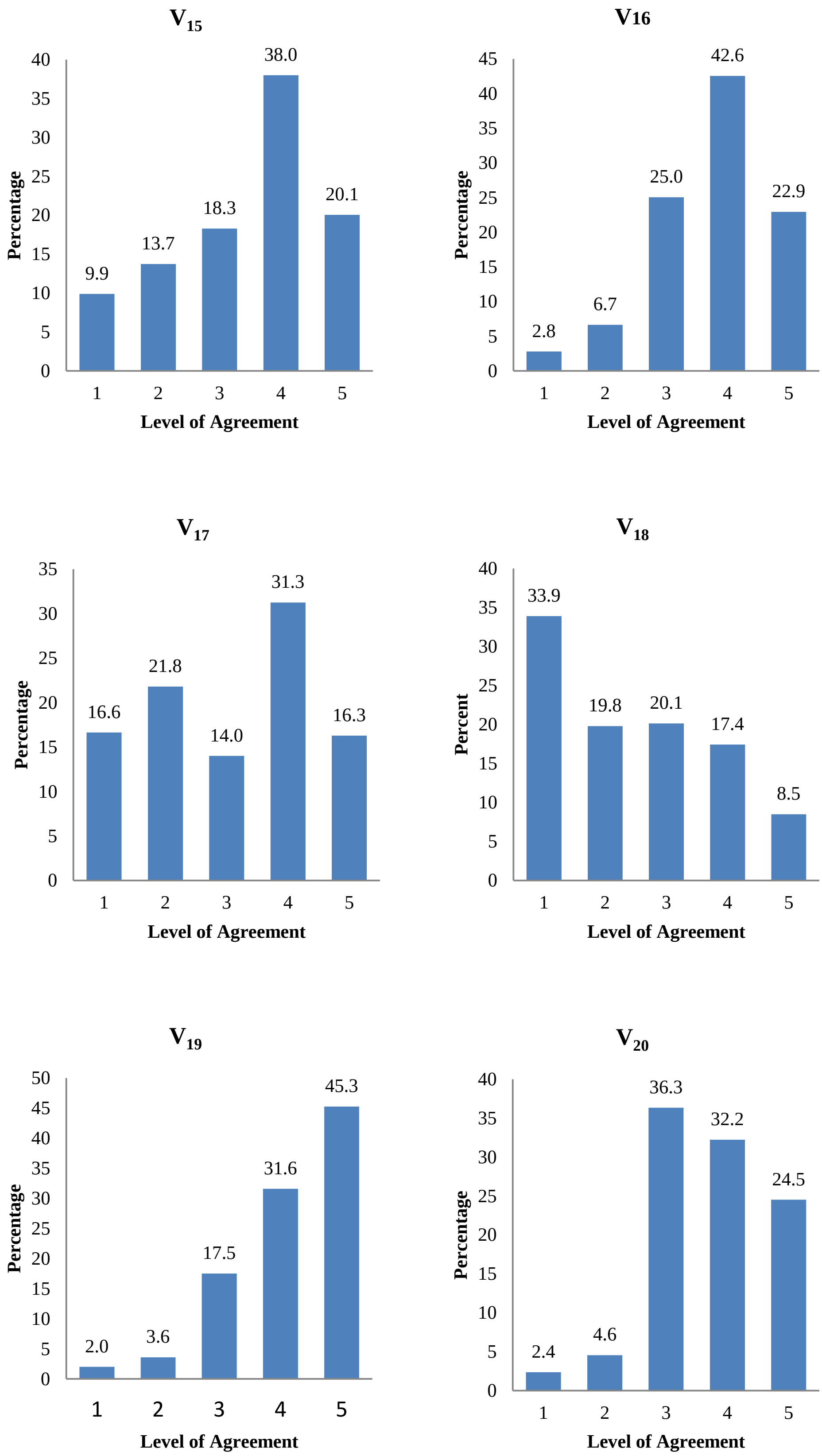

Figure 4: Percentage Distribution of Opinions on V_{15}, V_{16}, V_{17}, V_{18}, V_{19} and V_{20}

V_{15}: I like the social life of the people in this university

The distribution of opinions on variable V_{15} demonstrated in Figure 4 indicates that over one third of the respondents like the social life of the people in UCC. The distribution displayed in this figure tends to be negatively skewed. Accordingly, majority (58 percent) of the students like the social life of the people in UCC. Approximately, 1 of every 10 respondents strongly disagrees with this statement, suggesting that their choice of UCC does not depend on the social life of the people in the university. It is also observed that less than one fifth (18 percent) were unable to express their view on this attributes. The pattern of distribution indicated in Figure 4 on variable V_{15} is almost like that of Figure 2 on variable V_5.

V_{16}: A lot of facilities are available for research in this university

It can be observed from Figure 4 that very few (about 3 percent) respondents strongly disagree to the attributes in V_{16}. A lot more (43 percent) agree that a lot of facilities are available for research in UCC. Exactly one fourth of the respondents are not sure of choosing the University of Cape Cost based on its resources for research work. Only 3 out of every 100 respondents strongly disagree on this issue. This figure demonstrates a negatively skewed distribution. It shows that the availability of facilities for research is of much importance to the respondents when choosing the University of Cape Coast. Distribution of opinions on this indicator variable is similar to that in Figure 3 on the indicator variable V_{11}.

V_{17}: It was recommended to me by friends/relatives

With regard to the variable V_{17} in Figure 4, there is no clear bias on the opinions of respondents for this indicator variable. Close to half (47 percent) of the respondents tend to agree or strongly that UCC was recommended to them by their friends/relatives. Thus, the thought of coming to UCC was based on recommendations from friends and relatives. About 14 percent are not sure that UCC was recommended to them by their friends/relatives. The remaining students (38 percent) disagree (or strongly disagree) that UCC was recommended to them by their friends/ relatives. Almost the same percentage of respondents agrees with the attributes in V_{17} and V_5.

V_{18}: The cost of living is relatively cheaper in Cape Coast

It can be observed from Figure 4 that a small number (9 percent) of respondents strongly agree that the cost of living is relatively cheaper in Cape Coast. A lot more (54 percent) either disagree or strongly disagree with this statement. The shape of the graph is positively skewed. This means that most of the respondents perceive the cost of living in Cape Coast to be relatively expensive. Nearly, one fourth of the respondents are not sure. In all, 26 percent are certain that the cost of living is relatively cheaper. Exactly 3 respondents could not give their opinion on this variable. Generally, there is an indication that applicants are not driven by the cost of living on UCC campus.

V_{19}: The administration monitors performance of student very well

The trend of distribution of opinions on this indicator variable as displayed in V_{19} is negatively skewed. The implication is that majority (77 percent) of the respondents believe that the administration monitors performance of student very well. Only a few (2 percent) strongly disagree with this assertion. A little below one fifth of the respondents (17 percent) could not decide on this statement. Altogether, 6 of every 100 respondents either disagree or strongly disagree that the administration monitors performance of student very well. A similar pattern of distribution as indicated in Figure 3 on variable V_9 occurred in V_{19}.

V_{20}: Students comport themselves well during lecture hours

Surprisingly, 36 of every 100 respondents are not sure whether this attributes is important or not. The trend of distribution in Figure 4 on variable V_{20} is negatively skewed. This means that quite a large number (57 percent) of the respondents believe that student comport themselves well at lectures. A few (7 percent) objected to this attribute. The distribution of the views of respondents in this figure is similar to that of Figure 3 on the indicator variable V_{13}.

On the attributes of V_{21}, V_{22} V_{23} V_{24} V_{25} and V_{26} the distribution of opinions are displayed in Figure 5. The trend of distribution for each indicator variable is discussed subsequently in the subsections below.

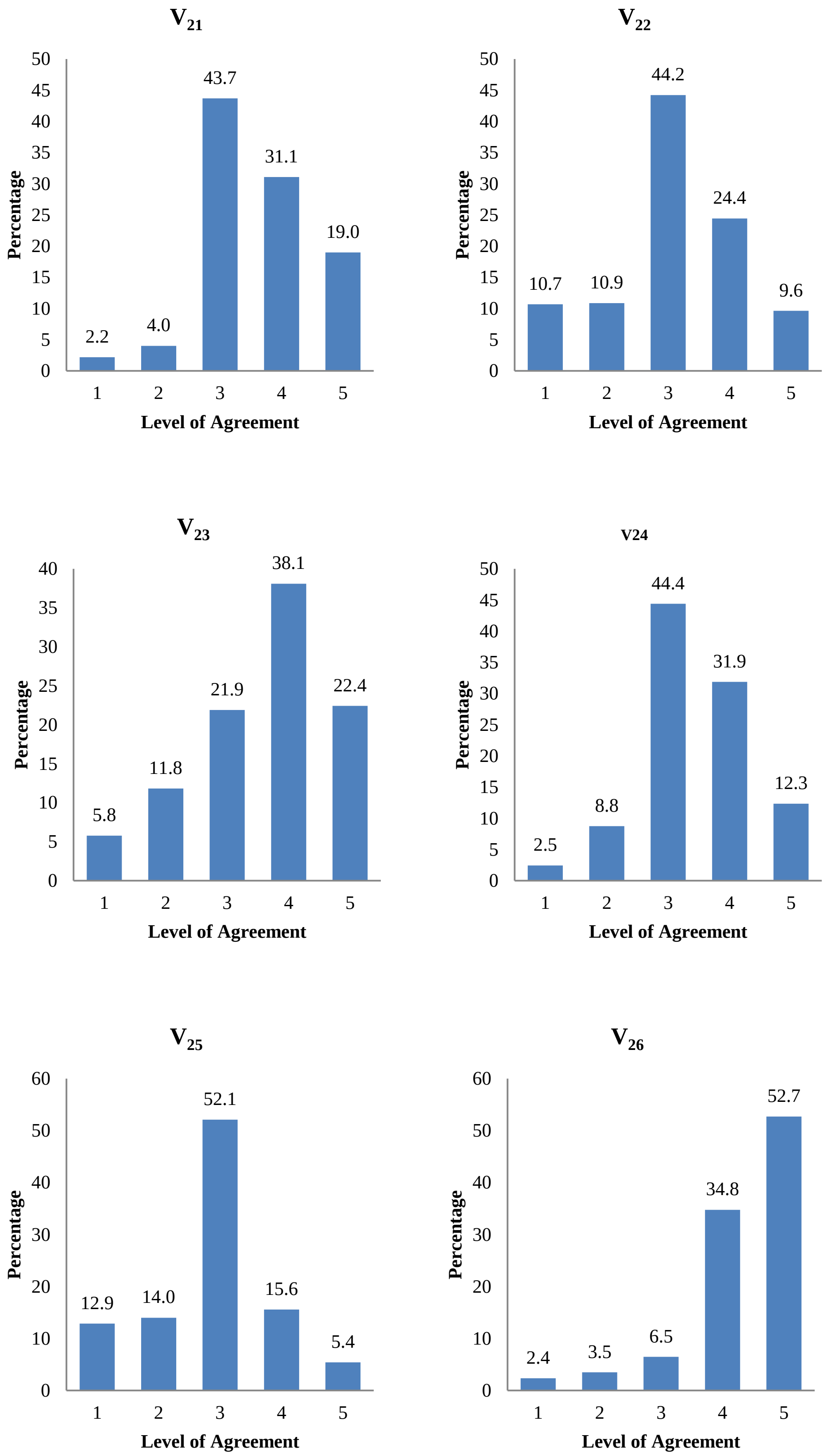

Figure 5: Percentage Distribution of Opinions on V_{21}, V_{22}, V_{23}, V_{24}, V_{25} and V_{26}

V_{21}: The administration regulates activities of the staff very well

A similar situation in the case of Figure 4 on variable V_{20} is shown in Figure 5 on variable V_{21}. The percentage of respondents who could not decide on the indicator variables V_{20} and V_{21} is high. For instance, 44 of every 100 respondents are unable to decide whether the administration regulates activities of the staff very well in UCC or not. Precisely 50 of every 100 respondents tend to agree or strongly agree with the statement. Very few respondents (6 percent) either disagree or strongly disagree on this issue.

V_{22}: Academic user fee is moderate

Figure 5 on variable V_{22} shows that almost half (44 percent) of the respondents are unable to indicate whether the academic user fee is moderate or not. The figure portrays a negatively skewed distribution. Accordingly around 34 percent of the students recognize that the academic user fee is moderate hence they would choose UCC because of this reason. Approximately 22 out of every 100 students argue that the academic user fee is not moderate. Consequently, they are not encouraged by this reason in their choice of UCC. There seems to be some level of variation in opinion on this indicator variable. The trend of distribution in V_{22} is almost like that of V_7.

V_{23}: The University is recognized worldwide

The nature of the graph in Figure 5 on variable V_{23} is a negatively skewed distribution. As a result, majority (60 percent) of the respondents either agrees or strongly agrees that UCC is recognized worldwide. Thus, more than half of the students consider worldwide recognition of the university as an important factor in their choice of UCC. Only a few (6 percent) strongly disagree to this attribute. Approximately 22 percent of the students are not sure whether they will choose the university based on its worldwide recognition or not. As a whole, 18 of every 100 respondents tend to disagree (or strongly disagree) that their choice of UCC depends on the worldwide recognition of the university.

V_{24}: There are many facilities for athletics

A look at V_{24} in Figure 5 reveals that a small number (3 percent) of respondents strongly disagree that there are many facilities for athletics in UCC. Consequently, they would not consider this factor when choosing the University of Cape Coast. Amazingly, almost half of the students (45 percent) were not decisive on this matter. The distribution of opinions on this indicator variable is negatively skewed. Accordingly, a little above 44 percent of the respondents either agree or strongly agree that they would choose the University of Cape Coast based on the facilities available for athletics. Some level of disparity exists in the views of the respondents concerning this factor.

V_{25}: A lot of financial sponsorship is available to students

The graph in Figure 5 on variable V_{25} shows that an unexpected number (52 percent) of respondents are not certain whether there is a lot of financial sponsorship in UCC or not. A little above one fifth of the students (21 percent) are motivated by the availability of financial sponsorship in the university. Close to 27 percent do not agree that financial sponsorship is available for students in UCC. Therefore they do not depend on this factor when choosing the University of Cape Coast. The graph displays a negatively skewed distribution. Distribution of opinions on this indicator variable is related to the responses on V_{22}.

V_{26}: I like the way students study in this university

An overwhelming number (88 percent) of respondents are attracted to the way students study in UCC as indicated in Figure 5 on variable V_{26}. However, it can be observed from the figure that just a few (2 percent) respondents strongly disagree to this attributes. Thus, these few respondents are not attracted to the way students study in the university. Nearly 7 of every 100 respondents are not certain about this attributes. The distribution is negatively skewed. The implication is that most of the prospective applicants who would choose the University of Cape Coast place particular emphasis on the way students study at UCC.

The distribution of opinions on the indicators V_{27} (the university has a competent staff) and V_{28} (the university has conducive learning environment) are displayed in Figure 6. The necessary comments are given in the subsections below.

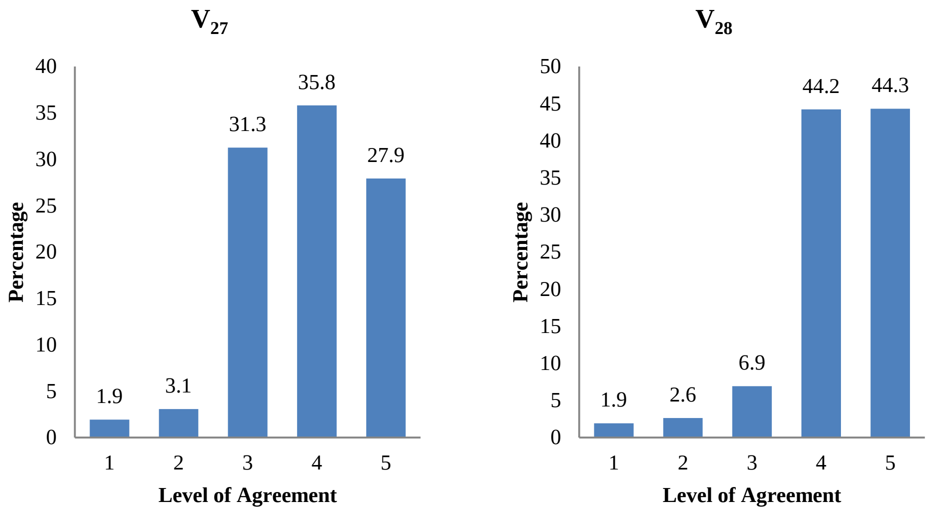

Figure 6: Percentage Distribution of Opinions on V_{27} and V_{28}

V_{27}: The University has a competent staff

According to the graph on the indicator variable V_{27}, very few (2 percent) respondents strongly disagree that UCC has a competent staff. Surprisingly 1 of every 3 respondents is not sure. In total, more than half (64 percent) of the respondents are in agreement that UCC has a competent staff. The shape of the diagram is negatively skewed. Therefore, greater part of the respondents would choose to attend UCC base on this idea. The trend of distribution in V_{27} is similar to that of V_{21} in Figure 5.

V_{28}: The University has conducive learning environment

In relation to variable V_{28}, virtually all the respondents (89 percent) tend to agree or strongly agree that they would choose the University of Cape Coast because of it's conducive learning environments. On the other hand, very few (2 percent) respondents strongly disagree to this issue. Merely 6 of every 100 students could not indicate whether they found this factor important or not. The figure is negatively skewed indicating that most of the respondents tend to appreciate the conducive learning environments in UCC.

Comparison of the Indicator Variables

A close examination and comparison of the bar charts for the indicator variables reveal that most of the charts for the variables are negatively skewed with the exception of V_1, V_3, V_4, V_{10}, V_{12}, and V_{18} which were positively skewed. The implication of this negative skewness is that most of the factors considered to influence applicants' choice of the University of Cape Coast are generally attractive. The extent of the negative skewness is more pronounced in the case of V_{26} (I like the way students study in this university) and V_{28} (the university has conducive learning environment) but less obvious in V_7 (It is cheaper to get accommodation in this university). The smaller amount of negative skewness on the indicator variable V_7 could mean that the importance of this factor in the choice of UCC is debatable.

A comparison of the bar charts for the 28 indicator variables reveals some similarities in the pattern of distribution of opinions. There are observed similarities in the pattern of distribution on the indicator variables V_{13} (Both students and lecturers are punctual at lectures), V_{20} (Students comport themselves well during lecture hours), V_{21} (The administration regulates activities of the staff very well) and V_{27} (The university has a competent staff). The two indicator variables V_9 (Students who complete this university are competent) and V_{19} (The administration monitors performance of students very well) also have similar pattern of distribution. All the six indicators (V_9,

V_{13}, V_{19}, V_{20}, V_{21} and V_{27}) had very small number of respondent who strongly disagree to these attributes.

Also, the indicator variables V_5 (It will be an opportunity to meet old friends), V_{15} (I like the social life of the people in this university) and V_{17} (It was recommended to me by friends/relatives) are related in the distribution of opinions. In each case, majority of the respondents agreed with these indicators.

Furthermore, the similarities in the pattern of distribution on V_7 (It is cheaper to get accommodation in this university), V_{22} (Academic user fee is moderate) and V_{25} (A lot of financial sponsorship is available to the student) are recognized. Thus, the percentage of respondents who are undecided on these indicators is surprisingly high.

The trend of distribution on the indicators V_{11} (The physical structure of the university is very attractive) and V_{16} (A lot of facilities are available for research in this university) were also similar. Most of the respondents agreed on these issues and very few respondents disagreed with these indicators.

The indicators with similar pattern of distribution identified in this section would be compared with the groups that would be obtained from the correlation analysis.

CORRELATION ANALYSIS OF THE INDICATOR VARIABLES

For the purpose of this study and based on the literature reviewed, a correlation coefficient of 0.2 will be used as a cut-off. This means that, a correlation coefficient greater than or equal to 0.2 will be considered high while that which is less than 0.2 will be taken as low. The use of the cut-off value will help to identify the variables that belong to a group. Thus, with a cut-off value of 0.2, we are able to clearly identify homogeneous set of variables from the correlation matrix. The 0.2 cut-off value is used because, most of the research conducted on factor analysis using the Likert Scale as the survey instrument had low correlations. As a result, Auger and Devinney (2007) considered a correlation greater than or equal to 0.2 as high in their study on the title, “Do What Consumers Say Matter? The Misalignment of Preferences with Unconstrained Ethical Intensions”. Also, Ampimah, (2007) used a cut-off value as low as 0.142 and had the highest correlation of 0.379. Besides, Nkansah et al. (2007) used a cut-off of 0.217 in their study on factor analysis of bath soup consumerism. Other references on the cut-off value used by some researchers are discussed in Chapter Two.

The correlations among the twenty eight indicator variables are shown in Table 3 and the meaning of the variables $V_1, V_2, V_3 \ldots, V_{28}$ are given.

Table 3: Correlation Matrix of the Indicator Variables

	V_1	V_2	V_3	V_4	V_5	V_6	V_7	V_8	V_9	V_{10}	V_{11}	V_{12}	V_{13}	V_{14}	V_{15}	V_{16}	V_{17}	V_{18}	V_{19}	V_{20}	V_{21}	V_{22}	V_{23}	V_{24}	V_{25}	V_{26}	V_{27}
V_2	.030																										
V_3	.169	.091																									
V_4	.108	.052	.160																								
V_5	.119	.155	.155	.076																							
V_6	.116	.064	.078	.051	.084																						
V_7	.010	.094	.107	.106	.055	.102																					
V_8	**.248**	.056	.122	.052	.137	.159	.035																				
V_9	.172	.093	.112	.063	.103	.111	.020	**.211**																			
V_{10}	.134	.107	**.316**	**.207**	**.263**	.103	.024	.131	.117																		
V_{11}	.079	.175	.181	.054	.138	.133	.089	**.243**	.161	.128																	
V_{12}	.019	**.402**	.111	.133	.117	.054	.090	.050	.057	.105	.062																
V_{13}	.057	.099	.117	-.031	.051	.145	.104	.172	**.246**	.066	.124	.094															
V_{14}	.074	.084	.150	.102	.147	.133	.123	.151	.085	.157	.199	.100	.142														
V_{15}	.124	.158	.198	.083	**.291**	**.246**	.121	.124	.131	**.234**	.198	.071	.123	.144													
V_{16}	.116	.059	.121	.066	.172	.102	.096	**.293**	**.205**	.117	**.367**	.035	.155	.199	.197												
V_{17}	**.240**	.137	**.436**	.173	**.244**	.143	.037	.162	**.226**	**.386**	.128	.124	.078	.090	**.216**	.135											
V_{18}	.026	.077	.059	.197	.038	.044	**.409**	-.043	.001	.002	.065	.076	.024	.060	.076	.028	.046										
V_{19}	.109	.112	.079	.021	.068	.179	.030	.196	**.229**	.116	.153	.004	**.303**	.062	.115	.172	.107	.027									
V_{20}	.064	.152	.062	-.033	.074	.099	.086	.179	**.207**	.081	.160	.128	**.619**	.177	.119	.125	.116	.034	**.323**								
V_{21}	.065	.110	.086	.082	.069	.149	.118	**.230**	**.255**	.007	.150	.044	**.255**	.173	.153	.188	.060	.080	**.430**	**.250**							
V_{22}	.020	.098	.146	.122	.150	.164	**.388**	.079	.080	.049	.094	.079	.127	.134	.135	.117	.045	**.245**	.066	.145	.132						
V_{23}	**.203**	.099	.110	.142	**.208**	.171	.061	**.360**	.145	.165	**.230**	.073	.193	**.202**	.199	**.266**	.184	-.018	.156	.163	.140	.075					
V_{24}	.101	.074	.111	.109	.164	.096	.103	.189	.109	.144	**.292**	.026	.120	**.254**	.188	**.376**	.105	.039	.077	.120	.085	.156	**.212**				
V_{25}	.112	.063	.103	.122	.138	.135	**.208**	.136	.106	.100	.153	.088	.104	.102	.167	.134	.132	**.276**	.113	.090	.124	**.217**	.139	.172			
V_{26}	.136	.097	.082	.031	.089	.163	.035	**.263**	**.394**	.087	**.207**	.027	.191	.177	.198	**.213**	.124	.058	**.247**	**.235**	.176	.081	.139	.137	.101		
V_{27}	.159	.135	.135	.032	.072	.113	.080	.197	**.254**	.095	**.301**	.078	**.217**	**.309**	.113	**.294**	.111	.014	**.204**	**.224**	**.304**	.119	.158	.196	.116	**.233**	
V_{28}	.063	.082	.029	.014	.097	.147	.029	**.280**	**.242**	.052	**.287**	.005	.177	.171	.131	**.381**	.080	-.004	**.242**	**.221**	.171	.122	**.248**	**.207**	.114	**.291**	**.298**

KMOMSA = 0.811 Value of Bartlett's Test of Sphericity = 5776 df = 378 Sig = 0.000 Determinant = 0.006

From Table 3, the highest positive correlation of 0.619 was recorded between the variables V_{13} (Both students and lecturers are punctual at lectures) and V_{20} (Students comport themselves well during lecture hours). This high correlation implies that respondents who consider the punctuality of both students and lecturers when choosing the University of Cape Coast are also interested in how students comport themselves during lecture hours. In other words, these two variables have much in common. The lowest correlation of 0.001 was obtained between V_9 (Students who complete UCC are competent) and V_{18} (The cost of living is relatively cheaper in this university). This low correlation implies that respondents who consider the competent level of UCC graduates are not necessarily interested in the cost of living in UCC. A correlation of -0.004 between V_{28} (the university has conducive learning environment) and V_{18} (The cost of living is relatively cheaper in this university) was the lowest negative correlation obtained. All the negative correlations obtained were low. These low correlations imply that there may be no significant pairwise relationship between the variables. Although correlations among all the pairs of the indicator variables were not very high, there were some significant pairs which were interesting.

Considering the correlation matrix in Table 3, the variables can be grouped based on their pairwise correlations. As discussed earlier in Chapter Two, the cut-off value of 0.2 is not out of place since a cut-off value less than 0.2 has been used successfully by other researchers. With a cut-off value of 0.2 in mind, a careful examination of the correlation matrix reveals six groups of indicators. To identify these groupings, there is the need to consider each indicator and identify the indicators with high correlations. Indicators with high correlations between them are grouped. The next step is to ensure that indicators in each group correlates highly with each other. Each indicator must belong to only one group. Based on this idea, the following groups were identified:

Group 1 (Academic Issues)

V_9 Students who complete this university are competent

V_{13} Both students and lecturers are punctual at lectures

V_{19} The administration monitors performance of students very well

V_{20} Students comport themselves well during lecture hours

V_{21} The administration regulates activities of the staff very well

V_{27} The university has a competent staff

Group 2 (Availability of Resources)

V_{11} The physical structure of the university is very attractive

V_{16} A lot of facilities are available for research in this university

V_{24} There are many facilities for athletics

V_{28} The university has conducive learning environment

Group 3 (Influence of others)

V_5 It will be an opportunity to meet old friends.

V_{10} Most of my friends and family members attended this university.

V_{15} I like the social life of the people in this university

V_{17} It was recommended to me by friends/relatives.

Group 4 (Cost of Study)

V_7 It is cheaper to get accommodation in this university.

V_{18} The cost of living is relatively cheaper in this university.

V_{22} Academic user fee is moderate.

V_{25} A lot of financial sponsorship is available to the student

Group 5 (Quality Academic Programme)

V_1 The only university that offers my choice of programme.

V_8 The university offer quality programmes.

V_{23} The university is recognized worldwide.

Group 6 (Geographic Location)

V_2 It is located in Central region where there are many tourist attractions.

V_{12} I like the university because it is located by the sea.

Of the six groups, Group 1 contains the highest number of indicator variables. This attribute has implications for identifying latent factors. This will be discussed further in the next section. The groupings indicate that, prospective applicants lay emphasis on academic issues before choosing the University of Cape Coast.

The second, third and fourth groups contain the next highest number of indicator variables. These groups may be labeled as availability of resources, influence of others and cost of study respectively. The indicator variables in the fifth group largely talk about quality academic programmes.

Group six contains the lowest number of indicator variables. As much as possible, a group containing just two indicator variables must be avoided unless the correlation between the two variables is very high. Generally, the correlations between the variables were low, therefore a correlation of 0.402 can be considered very high. Besides, the correlation between these two variables, V_2 (It is located in Central region where there are many tourist attractions) and V_{12} (I like the university because it is located by the sea) was one of the highest correlation obtained. Since these two variables do not have relatively high correlation with the other variables, the two variables may form a group.

It is not surprising that these groups have been identified from the correlation matrix. This is because almost all the indicators in each group seem to have similar pattern of distribution. For instance, the pattern of distribution on V_{13}, V_{20}, V_{21} and V_{27} in Group 1 is presented in Figure 7 for easy comparison.

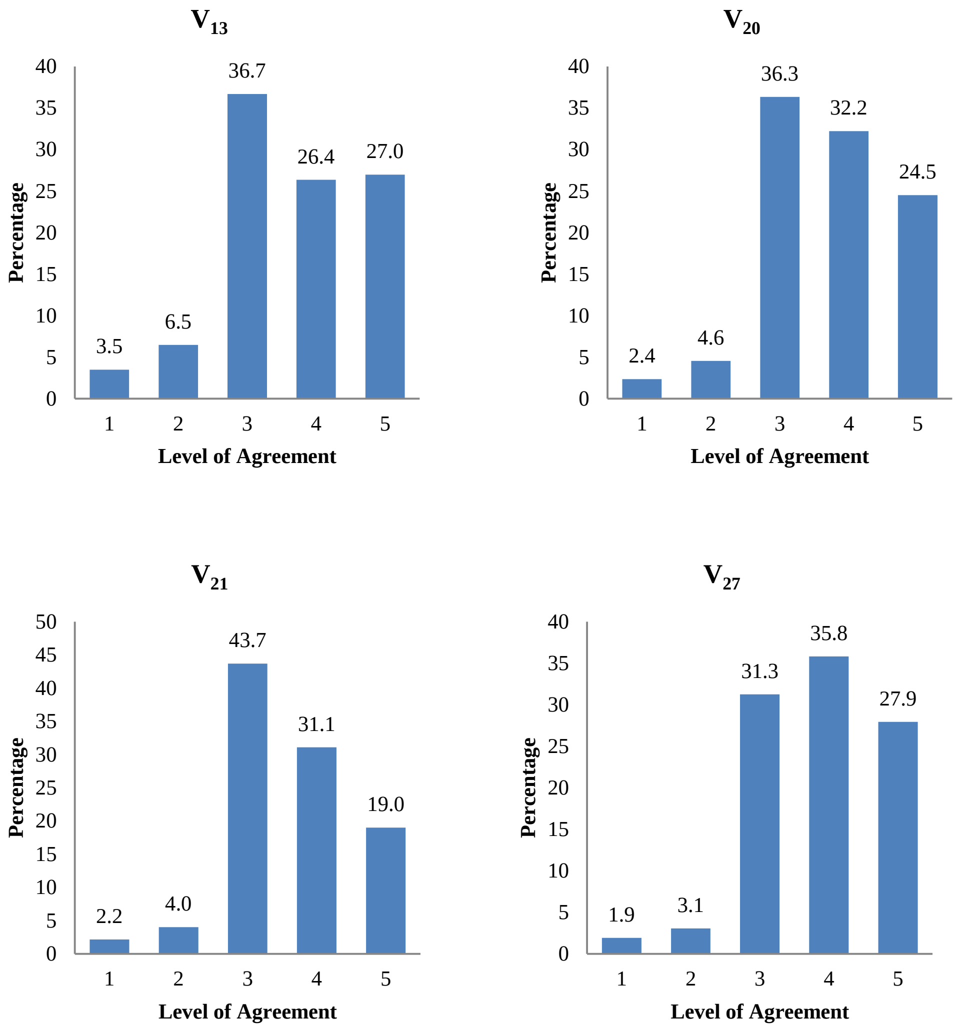

Figure 7: Comparing the Distribution of Opinions on V_{13}, V_{20}, V_{21} and V_{27}

It can be observed from Figure 7 that the percentages of respondents who strongly disagree on the four indicators are very few. On the other hand, a lot more of the respondents tend to either agree or strongly agree on these indicator variables.

The determinant listed at the bottom of the correlation matrix is 0.006 which is greater than the necessary value of 0.00001. Therefore, multicollinearity is not a problem for these data. To sum up, all the variables correlate fairly well and none of the correlation coefficients are particularly large; hence there is no need to consider eliminating any variable at this stage.

INTERPRETING KMO MEASURE FOR THE COMPONENTS

The KMO statistic varies between 0 and 1. An overall KMO measure of 0.811 for this data as shown at the bottom of Table 3 suggests that the correlation matrix is appropriate for factor analysis. Since this value is close to one, it indicates that patterns of correlations are relatively compact and so factor analysis should yield distinct and reliable factors. Summary of the recommendations given to any KMO measure are shown in Table 2. The value for these data falls into the range of being meritorious which means that factor analysis is appropriate for these data.

A higher number of indicators per factor contribute immensely to the high value of the KMO measure. It can be observed from the correlation analysis that six variables come together as a group to serve as a pointer to a factor identified as Academic Issues. The large number of indicators in group 1 contributes enormously to the substantial reason why this data is appropriate for factoring. In addition, three of the groups have four variables each, one group has three variables, and one group has only two variables. It is also recognized that relatively low correlations within group decrease the value of the KMO measure. The groups with very few indicators may also decrease the value of the KMO measure as in the case of two-indicator group identified.

It is important to note that high correlations among indicator variables do not contribute much to a high value of the KMO statistic. This might be the reason why factor analysis is appropriate for these data although the correlations among the indicator variables are relatively low.

BARTLETT'S TEST OF SPERICITY FOR THE COMPONENTS

The test statistic used for Bartlett's test of sphericity is given in Chapter Three, specifically, Equations (22) and (23). The sample size n = 1143, the number of indicator variables p = 28 and the eigenvalues of the component θ_i are given in Table 4. As indicated at the bottom of Table 3, the value of Bartlett's test of sphericity obtained for the data was 5776 and the significance level was 0.000. This value is the statistic for testing the null hypothesis that the population correlation matrix is an identity matrix. At any level of significance, the high value of the test statistic leads to the rejection of the null hypothesis which means that the population correlation matrix is not an identity matrix. This shows that Bartlett's test of sphericity is highly significant for these data.

Since the test is significant, it indicates that there are some relationships between the variables. In other words, there are correlations among the variables and hence factor analysis is appropriate for these data.

The total sample size used for this research is large. Hence it is not surprising that a huge value was obtained for the Bartlett's test of sphericity since the test is widely known for its sensitivity to sample size (Dillon and Goldstein, 1984)

INTERPRETING THE EIGENVALUES OF THE COMPONENTS

The eigenvalues associated with each factor represents the variance explained by that particular linear component. Furthermore, eigenvalues are the variances of the components extracted. It provides a lower bound for the number of factors present in the data. The eigenvalues and the corresponding percentage variances associated with each of the 28 possible factors before extraction and after extraction are listed in Table 4.

Table 4: Eigenvalues and Total Variances Explained

Component	Eigenvalues	Percentage of Variance	Cumulative Percentage
1	**4.876**	17.413	17.413
2	**2.025**	7.233	24.647
3	**1.772**	6.329	30.975
4	**1.513**	5.403	36.378
5	**1.299**	4.638	41.016
6	**1.108**	3.959	44.974
7	**1.082**	3.865	48.840
8	**1.028**	3.673	52.513
9	0.971	3.469	55.981
10	0.930	3.323	59.304
11	0.891	3.180	62.485
12	0.861	3.075	65.560
13	0.836	2.985	68.546
14	0.768	2.742	71.287
15	0.758	2.707	73.994
16	0.714	2.549	76.544
17	0.694	2.478	79.021
18	0.655	2.338	81.360
19	0.642	2.293	83.653
20	0.628	2.244	85.896
21	0.617	2.204	88.100
22	0.557	1.990	90.091
23	0.525	1.876	91.967
24	0.500	1.786	93.753
25	0.493	1.759	95.512
26	0.486	1.735	97.247
27	0.433	1.545	98.792
28	0.338	1.208	100.000

Before extraction, 28 linear components (factors) were identified within the data set. From Table 4, Factor 1 explains 17.413 percent of the total variance. The first eight components have eigenvalues greater than one. All the factors with eigenvalues greater than one are extracted based on the rule of eigenvalue-greater-than-one. This implies that eight components appear appropriate to estimate the correlation matrix. These eight components explain about 53 percent of the total variation in the data. If the first eight components are selected to approximate the observed matrix, then it indicates that information sacrificed is about 47 percent.

THE SCREE PLOT OF THE COMPONENTS

Often, when using the eigenvalue-greater-than-one rule, too many components are extracted, so it is important to look at the scree plot. If there is a change (or elbow) in the shape of the plots, then we choose the number of extracted factors to be equal to the number of components where the elbow appears to fall. The plot of the eigenvalues against the corresponding components is shown in Figure 8. The eigenvalues and the corresponding components used in the scree plots are given in Table 4 above.

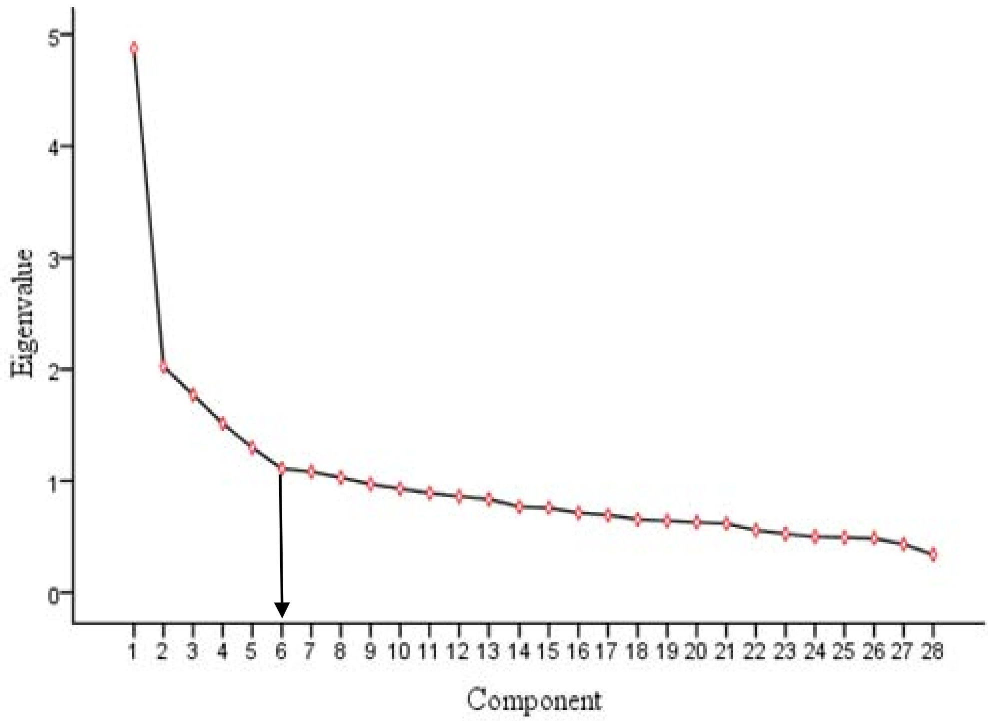

Figure 8: Scree Plot of the Components

The plot tends to show steep declines initially and then levels off after the sixth component. That is, the elbow is at the sixth component. The implication is that practically six factors may account for the variation in the data.

SUMMARY OF PRELIMINARY ANALYSIS OF THE VARIABLES

Various exploratory techniques have been used to analyze the data in this chapter. First, the percentage distributions of opinions on all the indicator variables were plotted on bar charts. These charts were examined to know if the indicator variables can be grouped based on the pattern of responses. It was observed that there were some similarities in the pattern of distribution on the indicator variables. In addition, these charts show patterns that tend to support the groupings based on the correlation matrix.

From the correlation matrix in Table 3, it was observed that the 28 variables involved in the study can be grouped into six sets of indicators. The group with the highest number of indicators talked about Academic Issues and the group with the lowest number of indicators talked about Geographical Location of UCC.

The two formal statistics used in testing the appropriateness of the factor analysis model revealed that the data is appropriate for factoring although the correlations among the variables were generally low. The overall KMO measure of 0.811 suggested that, patterns of correlations are relatively compact and so factor analysis should yield distinct and reliable factors. The large statistic value (5776) of the Bartlett's test of sphericity also support that the data is appropriate for factoring.

The eigenvalue analysis also revealed that the first eight components have eigenvalues greater than one. Therefore based on the eigenvalue-greater-than-one rule, eight factors could be extracted. Analysis based on the scree plots also showed that six factors could account for the variation in the data. The results from frequency distribution, correlation analysis, scree plot and eigenvalue analyses suggest that between six to eight factors can be used to explain the correlations among the 28 indicator variables.

CHAPTER FIVE

FURTHER ANALYSIS

The preliminary analysis reported in Chapter Four gave enough evidence that the data would be appropriate for factoring. As stated in Chapter One, the main objective of this research is to determine the latent factors that influence applicants' choice of the University of Cape Coast. As a result, this chapter focuses on finding the exact number of underlying factors that influence applicants' choice of UCC.

EXTRACTION OF THE UNDERLYING FACTORS

Factor extraction involves determining the smallest number of factors that can be used to best represent the interrelations among a set of variables. Determining the number of factors to use has been an issue since the beginning of factor analysis. Ultimately it is a judgment call. Most of the commonly used techniques of extraction have been reviewed in Chapter Three and applied in Chapter Four. A formal approach to confirm or otherwise this number of factors is discussed in subsequent sections.

Identifying the Number of Factors

The degrees of freedom of the statistic for testing the hypothesis that m is the correct number of factors was given in Chapter Three as

$$d = \frac{1}{2}\left[(p-m)^2 - p - m\right]$$

where p = 28 and m is the number of factors to extracted. Since the degrees of freedom must be positive, the possible values of m can be obtained by solving the inequality

$$(p-m)^2 - p - m > 0$$

Substituting the value of p into the quadratic inequality gives

$$(28-m)^2 - 28 - m > 0,$$

it implies $$m^2 - 57 + 756 > 0$$

$$m < 21 \text{ or } m > 36$$

The number of factors m, cannot be greater than the number of indicator variables p. This indicates that $m > 36$ is not appropriate since there are only 28 indicator variables. As a result, the number of latent factors to extract should be less than 21. If m is zero, it implies that all the p indicator variables are uncorrelated. It would be recalled that the Bartlett's test of sphericity indicated that the population correlation matrix is not an identity. Therefore, the lower limit of m cannot be zero. Thus, all the p indicators are correlated with each other. If the lower limit of m is one, then there is only one underlying factor which is possible. It follows that the range of values of m is $1 \le m \le 20$.

Furthermore, to determine the exact number of factors m, the results of the eigenvalue-greater-than-one rule and the scree plot discussed in Chapter Four would be used. The eigenvalue analysis indicated that eight factors could be identified whereas the scree plot in Figure 1 gave the indication that six factors seem appropriate. This could imply that the number of factors m may be in the interval $6 \leq m \leq 8$. This interval will now serve as a guide to the most suitable choice of *m*. Ultimately, the choice of *m* would be guided by the interpretability of the factors and parsimony consideration. Table 5 shows the loadings of each indicator on each factor.

Table 5: Unrotated 8-Factor Loading Matrix

Indicators	Factor							
	1	2	3	4	5	6	7	8
V_1	0.324	0.094	-0.287	0.025	0.330	0.207	0.067	0.329
V_2	0.305	0.171	0.096	0.341	**-0.571**	0.254	0.247	-0.044
V_3	0.380	0.389	-0.235	0.197	0.128	-0.050	-0.386	-0.143
V_4	0.222	**0.419**	-0.006	-0.019	0.139	0.315	-0.165	0.281
V_5	0.373	0.306	-0.206	0.012	-0.129	-0.265	0.260	-0.108
V_6	0.364	0.031	0.046	0.015	0.189	-0.194	**0.425**	-0.064
V_7	0.262	0.350	**0.595**	-0.121	0.098	-0.044	-0.049	0.005
V_8	**0.508**	-0.176	-0.203	-0.147	0.106	0.090	0.171	0.365
V_9	**0.474**	-0.226	-0.086	0.164	0.204	0.334	0.011	-0.255
V_{10}	0.371	0.377	-0.383	0.188	0.044	-0.159	-0.164	-0.089
V_{11}	**0.513**	-0.046	-0.061	-0.297	-0.252	0.066	-0.063	-0.145
V_{12}	0.222	0.268	0.100	0.376	**-0.550**	0.277	0.131	0.184
V_{13}	**0.468**	-0.340	0.238	0.387	0.002	-0.356	-0.176	0.174
V_{14}	**0.432**	0.038	0.033	-0.176	-0.222	-0.111	-0.329	0.048
V_{15}	**0.455**	0.228	-0.095	0.007	0.011	-0.285	0.339	-0.308
V_{16}	**0.541**	-0.123	-0.102	**-0.430**	-0.142	0.023	-0.063	-0.047
V_{17}	**0.425**	0.368	-0.388	0.280	0.179	0.024	-0.131	-0.098
V_{18}	0.171	0.397	**0.564**	-0.086	0.193	0.193	-0.013	-0.016
V_{19}	**0.457**	-0.342	0.100	0.267	0.226	0.032	0.078	-0.029
V_{20}	**0.480**	-0.341	0.239	**0.405**	-0.079	-0.309	-0.164	0.150
V_{21}	**0.462**	-0.268	0.234	0.149	0.159	0.152	-0.020	-0.003
V_{22}	0.338	0.271	**0.481**	-0.112	0.077	-0.157	0.005	-0.058
V_{23}	**0.495**	-0.011	-0.207	-0.146	-0.026	-0.121	0.197	**0.496**
V_{24}	**0.447**	0.057	-0.054	**-0.425**	-0.197	-0.173	-0.152	0.031
V_{25}	0.365	0.256	0.265	-0.137	0.185	0.036	0.154	0.084
V_{26}	**0.485**	-0.254	-0.027	0.015	0.136	0.248	0.132	-0.319
V_{27}	**0.519**	-0.224	0.036	-0.098	-0.136	0.212	-0.305	-0.110
V_{28}	**0.497**	-0.323	-0.022	-0.275	-0.088	0.088	0.083	-0.090

In this study, a loading of 0.4 is considered high enough to associate a factor with a variable. This value is fairly good because the literature review indicates that some researchers such as Kallio (1995) successfully used factor loadings greater than or equal to 0.4. Also, Brown (2006) stated that in applied research, factor loadings greater than or equal to 0.30 or 0.40 are often interpreted as salient. However, widely accepted guidelines do not exist and the criteria for salient and nonsalient loadings often depend on the empirical context. The factor loading matrix above shows the correlation between each variable and each factor. For example, V_1 (The only university that offers my choice of programme) has a 0.324 correlation with Factor 1 and a 0.094 correlation with Factor 2.

From Table 5, it can be observed that most of the indicator variables load highly on Factor 1. Thus, Factor 1 is related most closely to V_{16} (A lot of facilities are available for research in this university) followed by V_{27} (The university has a competent staff), V_{11} (The physical structure of the university is very attractive) and V_8 (The university offer quality programmes). Consequently, exactly 16 indicator variables load highly (above 0.4) on the first factor, three indicators each load highly on the third and fourth factors, two indicator variables load highly on the fifth factor and only one indicator each load highly on the second, seventh and eighth factors. Surprisingly, all the indicators load very low on the sixth factor. Since the non-zero loadings on Factor 6 are not high enough to help identify its indicators, 6-factor model may not be appropriate.

Since the first factor loads highly on as many as sixteen indicators but the second factor loads highly on just one factor, it suggests that there is one important underlying factor influencing applicants' choice of UCC. This factor largely reflects academic resources. In view of the fact that several indicators loaded highly on the first factor, it is an indication that the number of factors are not many. Strictly speaking, the unrotated factor loadings support a single factor solution. However, a single factor obviously will not be consistent with the rules of extraction that have already being examined. It will also not reflect adequately the actual factors that influence applicants' choice of UCC. It is therefore necessary to examine a rotation of the loadings.

ROTATION OF THE EIGHT FACTORS EXTRACTED

Having determined the appropriate number of factors, the extracted factors are rotated to foster their interpretability. In instances where two or more factors are involved, rotation is possible because of the indeterminate nature of the common factor model. Thus, for any given multiple-factor model, there exist an infinite number of equally good fitting solutions, each represented by a different factor loading matrix (Brown, 2006).

For this data, eight factors are involved hence, rotation is possible. Before rotation, most variables loaded highly onto the first factor and the remaining factors did not give a good look. However, the rotation of the factor structure is intended to clarify this issue considerably to obtain a simple structure. By a simple structure, it means we obtain the most readily interpretable solutions in which:

1. Each factor is defined by a subset of indicators that load highly on the factor.

2. Each indicator (ideally) has a high loading on one factor (often referred to as a primary loading) and has a trivial or close to zero loading on the remaining factors (referred to as secondary loading).

The eight factors extracted constitute the original unrotated factor solution. There was no definite representation for some of the factors extracted. Therefore, by rotating the factor solution, it is expected that there would be a definite representation for all the factors which will make interpretation considerably easier. The following sections examine both varimax and quartimax rotations of the original factor solution. It has been noted earlier in this study that a loading of 0.4 is relatively high to associate a factor with a variable.

Varimax Rotation of the 8-Factor Model

The varimax rotation is the most widely used type of rotation. As noted in Chapter Three, the main objective is to make some of these loadings as large as possible and the rest as small as possible in absolute value. Using the matrix algebra in Equation (27), the factor loadings in Table 5 is multiplied by the transformation matrix in Table 6 to produce the rotated factor loading matrix in Table 7. In this data set, the varimax transformation matrix is as follows:

Table 6: Varimax Transformation Matrix

Factors	1	2	3	4	5	6	7	8
1	0.539	0.444	0.258	0.317	0.343	0.328	0.300	0.171
2	-0.130	-0.424	0.498	0.550	-0.409	0.011	0.172	0.236
3	-0.083	0.026	0.761	-0.456	0.302	-0.296	-0.136	0.094
4	-0.627	0.175	-0.171	0.337	0.494	-0.122	-0.015	0.418
5	-0.436	0.326	0.264	0.189	-0.028	0.279	0.011	-0.721
6	-0.067	0.534	0.076	-0.050	-0.504	0.182	-0.522	0.378
7	-0.304	0.125	-0.019	-0.461	-0.235	0.270	0.703	0.238
8	-0.078	-0.430	0.041	-0.158	0.263	0.778	-0.309	0.113

These values convey how much the axes were rotated to foster simple structure. Specifically, the values on the diagonal are cosines and the values on the off-diagonal are sines. It must be noted that although this transformation fosters the interpretability of the factor solution, it does not alter the communality of the indicators. The factor loadings of the new factor solution after varimax rotation of the 8-factor model are shown in Table 7.

Table 7: Varimax Rotated Factor Loading Matrix

Indicators	Factors							
	1	2	3	4	5	6	7	8
V_1	-0.033	0.187	0.023	0.263	-0.044	**0.593**	-0.007	-0.045
V_2	0.081	0.124	0.040	0.035	0.051	-0.048	0.151	**0.805**
V_3	0.126	0.060	0.111	**0.736**	0.073	-0.014	0.011	-0.001
V_4	0.025	-0.010	0.340	0.340	-0.156	0.352	-0.226	0.140
V_5	0.174	-0.071	0.026	0.271	-0.005	0.087	**0.545**	0.164
V_6	-0.015	0.191	0.167	-0.024	0.109	0.186	**0.530**	-0.036
V_7	0.097	-0.037	**0.739**	0.007	0.099	-0.061	0.047	0.034
V_8	0.273	0.216	-0.040	-0.014	0.120	**0.619**	0.113	0.005
V_9	0.095	**0.689**	-0.015	0.178	0.067	0.081	0.027	0.042
V_{10}	0.113	-0.024	-0.041	**0.669**	0.045	0.086	0.221	0.054
V_{11}	**0.610**	0.201	0.047	0.066	-0.020	0.034	0.122	0.122
V_{12}	0.008	-0.041	0.083	0.092	0.075	0.076	-0.021	**0.820**
V_{13}	0.097	0.140	0.050	0.055	**0.829**	0.056	0.051	0.025
V_{14}	**0.536**	-0.048	0.127	0.191	0.208	0.020	-0.057	0.058
V_{15}	0.154	0.130	0.120	0.224	0.020	-0.018	**0.671**	0.056
V_{16}	**0.669**	0.193	0.037	0.014	-0.013	0.170	0.118	-0.035
V_{17}	0.006	0.169	-0.003	**0.717**	0.005	0.167	0.167	0.078
V_{18}	-0.044	0.078	**0.751**	0.022	-0.080	-0.023	-0.059	0.069
V_{19}	-0.007	**0.511**	0.038	0.011	**0.411**	0.152	0.110	-0.017
V_{20}	0.123	0.159	0.031	0.045	**0.813**	0.028	0.045	0.110
V_{21}	0.098	**0.483**	0.192	-0.025	0.335	0.125	-0.033	0.037
V_{22}	0.157	-0.005	**0.613**	0.029	0.156	-0.066	0.188	0.000
V_{23}	0.297	-0.068	-0.009	0.028	0.185	**0.651**	0.225	0.077
V_{24}	**0.645**	-0.091	0.114	0.091	0.040	0.112	0.129	-0.049
V_{25}	0.091	0.099	**0.499**	0.038	-0.005	0.226	0.185	0.017
V_{26}	0.196	**0.655**	0.015	0.034	0.026	0.033	0.168	0.018
V_{27}	**0.514**	0.387	0.047	0.114	0.172	0.002	-0.179	0.092
V_{28}	**0.498**	0.376	-0.025	-0.147	0.075	0.144	0.140	-0.002

From Table 7, it can be observed that the indicator variables V_{11} (The physical structure of the university is very attractive), V_{14} (The workers of the university are friendly), V_{16} (A lot of facilities are available for research in this university), V_{24} (There are many facilities for athletics), V_{27} (The university has a competent staff) and V_{28} (The university has conducive learning environment) have uniquely high loadings on the first factor. These variables were part of the set of indicators for the first factor in the unrotated factor solution.

It can further be observed from the table that the second factor has relatively high loadings on V_9 (Students who complete this university are competent), V_{19} (The administration monitors performance of student very well), V_{21} (The administration regulates activities of the staff very well) and V_{26} (I like the way students study in this university). In the unrotated factor loading matrix, all these indicators loaded highly on the first factor.

Also, the third factor has relatively high loadings on four indicator variables. These variables are V_7 (It is cheaper to get accommodation in this university), V_{18} (The cost of living is relatively cheaper in Cape Coast), V_{22} (Academic user fee is moderate) and V_{25} (A lot of financial sponsorship is available to students). The first three indicators V_7, V_{18} and V_{22} were identified in the unrotated factor loadings but the fourth indicator V_{25} was not identified. Thus, rotation of the latent factors has helped to identify a fourth variable which will contribute to the interpretation of this factor.

Indicator variables that load highly on the fourth factor are V_3 (It is base on my counselor's recommendation), V_{10} (Most of my friends and family members attended this university) and V_{17} (It was recommended to me by friends/ relatives). None of these indicators were identified in the unrotated factor loading matrix except V_{17} which was identified as part of the indicators in the first factor.

Three indicator variables have high loadings on the fifth factor. The indicator V_{19} (The administration monitors performance of student very well) was identified as part of the second factor and it has a relatively high loading on the second factor than the fifth factor. Hence, only two variables have very high loadings on the fifth factor. These two variables are V_{13} (Both students and lecturers are punctual for lectures) and V_{20} (Students comport themselves well during lecture hours). In the unrotated matrix, these two indicators were also identified as part of the indicators in the first factor.

Further observation of the factor matrix reveals that the indicator variables V_1 (The only university that offers my choice of programme), V_8 (The university offers quality programmes) and V_{23} (The university is recognized worldwide) have relatively high loadings on the sixth factor. The indicator variables V_8 and V_{23} were associated with the first factor in the case of the unrotated factor loadings.

Close examination of Table 7 also indicates that there are three variables that load relatively high on the seventh factor. These variables are V_5 (It will be an opportunity to meet old friends), V_6 (I prefer the grading system of this university to other universities) and V_{15} (I like the social life of the people in this university). Only V_6 had a high loading on this factor in the unrotated factor loadings. Also, from the unrotated factor matrix, V_{15} had a high loading on the first factor.

Finally, indicator variables V_2 (It is located in Central region where there are many tourist attractions) and V_{12} (I like the university because it is located by the sea) have very high loadings on the eighth factor. The unrotated factor loadings indicated that these two variables had relatively high loadings on the fifth factor. Clearly, rotation of the factor matrix has helped to produce a factor solution with the best simple structure. It can be observed from Table 7 that variable V_{19} loaded highly on the second and fifth factor. Consequently, there is the need for quartimax rotation in order to verify exactly where to place this variable.

Quartimax Rotation of the 8-Factor Model

The quartimax criterion on the other hand, seeks to maximize the variance of the squared loadings for each variable and tends to produce factors with high loadings for all variables. Using Equation (27), the original factor loading matrix in Table 5 is multiplied by the transformation matrix in Table 8 to obtain the quartimax rotated factor loadings in Table 9. The transformation matrix in the case of the quartimax rotation is as shown below:

Table 8: Quartimax Transformation Matrix

Factors	1	2	3	4	5	6	7	8
1	0.573	0.454	0.259	0.317	0.318	0.303	0.160	0.280
2	-0.132	-0.432	0.502	0.554	-0.393	0.018	0.232	0.170
3	-0.088	0.035	0.761	-0.459	0.299	-0.293	0.094	-0.136
4	-0.614	0.206	-0.163	0.343	0.498	-0.108	0.421	-0.005
5	-0.425	0.339	0.263	0.187	-0.037	0.282	-0.720	0.014
6	-0.065	0.516	0.071	-0.052	-0.523	0.176	0.379	-0.524
7	-0.282	0.133	-0.013	-0.449	-0.237	0.276	0.243	0.714
8	-0.068	-0.406	0.044	-0.156	0.278	0.790	0.114	-0.300

The factor loading matrix for the quartimax rotation is displayed in Table 9.

Table 9: Quartimax Rotated Factor Loading Matrix

Indicators	Component							
	1	2	3	4	5	6	7	8
V_1	-0.012	0.196	0.024	0.265	-0.052	**0.589**	-0.048	-0.009
V_2	0.094	0.126	0.045	0.041	0.046	-0.052	**0.804**	0.148
V_3	0.134	0.061	0.112	**0.734**	0.069	-0.022	-0.008	0.000
V_4	0.033	-0.011	0.339	0.339	-0.158	0.350	0.136	-0.231
V_5	0.191	-0.069	0.031	0.276	-0.005	0.081	0.160	**0.540**
V_6	0.008	0.203	0.170	-0.019	0.100	0.181	-0.037	**0.528**
V_7	0.101	-0.034	**0.739**	0.005	0.095	-0.065	0.029	0.041
V_8	0.299	0.226	-0.039	-0.012	0.107	**0.608**	0.001	0.107
V_9	0.114	**0.691**	-0.017	0.177	0.040	0.065	0.038	0.017
V_{10}	0.125	-0.021	-0.038	**0.671**	0.045	0.080	0.048	0.213
V_{11}	**0.619**	0.189	0.044	0.064	-0.036	0.013	0.115	0.106
V_{12}	0.019	-0.035	0.088	0.097	0.077	0.077	**0.819**	-0.022
V_{13}	0.115	0.170	0.054	0.054	**0.822**	0.051	0.022	0.047
V_{14}	**0.539**	-0.052	0.125	0.187	0.202	0.006	0.050	-0.070
V_{15}	0.173	0.133	0.123	0.229	0.013	-0.028	0.052	**0.663**
V_{16}	**0.679**	0.181	0.034	0.011	-0.029	0.147	-0.042	0.102
V_{17}	0.024	0.174	0.001	**0.720**	0.000	0.160	0.072	0.159
V_{18}	-0.041	0.078	**0.750**	0.021	-0.086	-0.025	0.066	-0.063
V_{19}	0.016	**0.529**	0.039	0.011	0.391	0.143	-0.019	0.106
V_{20}	0.141	0.188	0.035	0.044	**0.804**	0.022	0.107	0.041
V_{21}	0.117	**0.495**	0.191	-0.027	0.314	0.114	0.034	-0.039
V_{22}	0.165	-0.001	**0.614**	0.028	0.151	-0.072	-0.005	0.180
V_{23}	0.322	-0.054	-0.007	0.031	0.182	**0.643**	0.073	0.221
V_{24}	**0.650**	-0.100	0.112	0.088	0.034	0.094	-0.057	0.115
V_{25}	0.106	0.104	**0.500**	0.039	-0.013	0.220	0.013	0.179
V_{26}	0.215	**0.653**	0.013	0.034	-0.001	0.015	0.014	0.158
V_{27}	**0.522**	0.381	0.044	0.108	0.151	-0.019	0.085	-0.195
V_{28}	**0.513**	0.371	-0.027	-0.150	0.053	0.124	-0.007	0.128

The loadings on this matrix are similar to those for the varimax rotation in Table 7 except that the indicators constituting the seventh and eighth factors have been interchanged. Also, there are three indicators comprising the fifth factor when the varimax rotation was applied whiles only two indicators represents the fifth factor in the case of the quartimax rotation. Thus, the indicator variable V_{19} has been identified to have a high loading on the second factor and not on the fifth factor in the quartimax rotation. This identification in the case of quartimax rotation will make it easier in the interpretation of these factors.

It is essential to note that transformation of the original factor loadings does not alter the communality of the indicators. In a solution entailing more than one variable, communalities in an orthogonal analysis are calculated by taking the sum of squared loadings for a given indicator across all factors. For instance, before rotation, the proportion of variance explained in V_1 is equal to $0.324^2 + 0.094^2 + (-0.287)^2 + \cdots + 0.207^2 + 0.067^2 + 0.329^2 = 0.461$. After rotation, the proportion of variance explained in V_1 is equal to $(-0.033)^2 + 0.187^2 + 0.023^2 + \cdots + 0.593^2 + (-0.007)^2 + (-0.045)^2 = 0.461$. All the indicators will have the same communality for both unrotated and rotated factor solution. This confirms that rotation does not alter the fit of the factor solution. The communalities for the indicator variables are presented in Appendix III.

After rotation, the first five factors seem to have a unique interpretation. An attempt to interpret the sixth and seventh factors may be a repetition of a factor that has already been identified. Consequently, a five factor model is adopted for this data. The interpretation of the extracted factors and the goodness-of-fit of the five-factor model are presented in the next two sections.

INTERPRETATION OF FACTORS

From the rotated factor loadings, the first factor consists of V_{11}, V_{14}, V_{16}, V_{24}, V_{27} and V_{28}. These indicator variables largely reflect availability of academic resources. Therefore, Factor 1 can be labeled as Academic Resources.

The indicator variables constituting the second factor are V_9, V_{19}, V_{21}, and V_{26}. Each of these variables is concern with provision of competent graduates at the University of Cape Coast. Thus, the second factor represents Quality Assurance.

Each of the indicators constituting the third factor (V_7, V_{18}, V_{22} and V_{25}) is related to the cost of study at the University of Cape Coast. Hence these variables reveal that Cost of Study is also an underlying factor which influences applicants' choice of UCC.

Indicator variables which represent the fourth factor are V_3, V_{10} and V_{17}. These indicators largely suggest that applicants are influenced by friends, relatives and counselors. Consequently, Factor 4 can be represented as Influence of others.

Only two indicator variables (V_{13} and V_{20}) comprise the fifth factor. These variables mainly disclose issues of discipline. The fifth factor may be labeled discipline.

Basically, the sixth and the seventh factor emphasize issues that have already been discussed. As a result, an attempt to interpret these factors may be a repetition of factors identified earlier.

Thus, it can be observed that Academic Resources is the most important factor that influences applicants' choice of UCC. The second most important factor deals with issues relating to Quality Assurance. Other factors that influence selection of UCC by prospective applicants are Cost of Study, Influence of Others and Discipline.

TEST OF ADEQUACY OF THE FIVE-FACTOR MODEL

The likelihood ratio chi-square test statistic used to assess the statistical fit of the model is quite sensitive to sample size (DeCoster, 1998). In large samples, the hypothesis that a particular model fits the data may be rejected even though the discrepancy between the model and the data may be in a practical sense quite small. This problem has led researchers to develop a large number of alternative overall goodness-of-fit indices (such as the Tucker-Lewis index) that are independent of sample size. However, most of these indices are not yet provided as part of the diagnostic output in commonly used software such as SPSS. Nevertheless, there is the need to test the goodness-of-fit of the five-factor model. The hypotheses for testing the adequacy of the model are given as:

$$H_0 : \boldsymbol{\Sigma} = \boldsymbol{\Lambda\Lambda}' + \boldsymbol{\Psi}$$

against $$H_1 : \boldsymbol{\Sigma} \neq \boldsymbol{\Lambda\Lambda}' + \boldsymbol{\Psi}$$

where $\boldsymbol{\Psi}$ is a 28 x 28 diagonal matrix of unique variances of the indicators;

$\boldsymbol{\Lambda}$ is a 28 x 5 matrix of factor loadings;

$\boldsymbol{\Sigma}$ is a 28 x 28 population covariance matrix.

The test statistic, a function of the likelihood ratio was given in Equation (26) as:

$$\left(n - \frac{2p + 4m + 11}{6}\right) \text{In}\left(\frac{\left|\hat{\boldsymbol{\Lambda}}\hat{\boldsymbol{\Lambda}}' + \hat{\boldsymbol{\Psi}}\right|}{|\mathbf{S}|}\right)$$

which is approximately χ^2 with degrees of freedom $d = \frac{1}{2}\left[(p - m)^2 - p - m\right]$

Here $n = 1142$, $p = 28$, $m = 5$, $\hat{\boldsymbol{\Lambda}}$ and $\hat{\boldsymbol{\Psi}}$ are the solutions of the maximum likelihood equations. This implies that

$$\chi^2 = 1127.5 \ \text{In}\left(\frac{\left|\hat{\boldsymbol{\Lambda}}\hat{\boldsymbol{\Lambda}}' + \hat{\boldsymbol{\Psi}}\right|}{|\mathbf{S}|}\right)$$

and the degrees of freedom, $d = 248$. It is difficult to compute $\ln\left(\frac{\hat{\boldsymbol{\Lambda}}\hat{\boldsymbol{\Lambda}}' + \hat{\boldsymbol{\Psi}}}{|\mathbf{S}|}\right)$, hence the method of maximum likelihood ratio will be employed to obtain the value of the chi-square goodness-of-fit statistic. Using the method of maximum likelihood, chi-square goodness-of-fit can easily be generated as $\chi^2 = 876.374$. The corresponding critical value from the chi-square table at 0.05 level of significance is approximately $\chi^2 = 124$. The large value of the test statistic leads to the rejection of the null hypothesis that the population correlation matrix can be approximated by the five-factor model. Thus five factors are not enough to explain the correlation among the indicator variables. This is not surprising since it has been noted by Rencher (2002) that when *n* is large, the goodness-of-fit test often shows more factors to be significant.

The result obtained could also mean that there are other minor factors affecting the choice of UCC by applicants which were not identified by the methods employed in this research. However, it is the aim of the study to find the first few important factors underlying the correlation matrix. Therefore, the result of the goodness-of-fit test is used in the confirmatory sense to point out that with five factors extracted; there may still be other minor factors that could be considered.

It must be pointed out that the relevance of this adequacy test cannot be insisted in this study. This is because it would imply that the number of factors constituting the model must be increased beyond five. However, it has already been observed that an interpretation for more than five factors is not apparent. Besides, interpretations for these minor factors may not be distinct from the first five factors. As mentioned earlier, identifying the number of factors to extract largely depends on the ability to interpret these factors distinctively. Hence, it is better to maintain the five factors as the final factor solution.

CALCULATING FACTOR SCORES

The procedure for calculating the factor score is provided in Chapter Three. The weights or factor score coefficients used to combine the standardized variables are provided in the factor score coefficients matrix in Appendix IV. For example, using the factor score coefficients matrix in Appendix IV, one could compute five factor score for each respondent. The factor scores for this data are obtained using SPSS. In this study, these scores would be used to estimate the level of dependency of the extracted factors on the demographic characteristics of the respondents.

LEVEL OF DEPENDENCY OF EXTRACTED FACTORS ON DEMOGRAPHIC CHARACTERISTICS OF RESPONDENTS

The main focus in this section is to find the extent of influence of the demographic characteristics of the respondents on the extracted factors. The demographic characteristics of respondents considered in this study are sex, age, school and region of birth. The factor scores for the respondents are classified into three ranges as discussed in Chapter Three. Thus, respondents with scores above the 75^{th} percentile are classified as strongly influenced whiles those with scores below the 25^{th} percentile are considered as least influenced. Respondents with scores ranging from 25^{th} to 75^{th} percentile are classified as moderately influenced. Each range suggests the rate at which a respondent is being influenced by a particular factor. The summary of each range and the corresponding interpretations are given in Table 10.

Table 10: Interpretation of Factor Score

Factor Score	Interpretation
-0.70 and below	Least influenced
-0.69 to 0.69	Moderately influenced
0.70 and above	Strongly influenced

Any respondents whose factor score falls within the first range is least influenced and those within the third range are strongly influenced by that factor. The categories under age have been reduced from three to two since there was no respondent within the age of 12 to 15 years. Besides, the schools and the regions have been classified in Chapter One in order to meet the conditions of the contingency table. Distribution of respondents' demographic characteristic and the extent of influence on the five extracted factors are presented on separate contingency tables. The distribution of applicants' level of influence on the first factor is presented in Table 11.

Table 11: Influence on the First Factor by Demographic Characteristics

	Level of influence			
Demographic characteristics	Least influenced	Moderately influenced	Strongly influenced	Total
Sex				
Male	27.7	53.0	19.3	100.0
Female	14.6	56.5	28.9	100.0
Age				
16 to 19yrs	23.0	53.9	23.1	100.0
20yrs and above	19.0	58.9	22.1	100.0
School				
First class sch.	29.0	54.4	16.6	100.0
Second class sch.	17.3	54.3	28.4	100.0
Region				
Zone 1	24.7	51.1	24.2	100.0
Zone 2	19.5	58.2	22.3	100.0
Zone 3	28.9	55.6	15.5	100.0

It can be observed from Table 11 that more than half of the respondents are moderately influenced by the first factor irrespective of the background. Generally, an extracted factor must have some moderate influence on a respondent irrespective of the background.

The number of respondents that are influenced in the extreme will determine whether or not this factor depends on the demographic characteristics of the respondents. For example, the percentage of females who are strongly influenced by the first factor is much higher than those who are least influenced. However, the percentage of males who are strongly influenced by the first factor is much less than those who are least influenced. This suggests that females are strongly influenced by academic resources than their male counterpart. Similarly, respondents in the second class schools are strongly influence by academic resources than those in the first class schools.

The percentage of respondents within the two age groups who are strongly influenced by the first factor is almost the same. As a result, academic resources might be independent of a respondents' age. The distribution of applicants' level of influence on the second factor is presented in Table 12 and the remaining three factors are presented in Appendix V.

Table 12: Influence on Second Factor by Demographic Characteristics

Demographic characteristics	Level of influence			
	Least influenced	Moderately influenced	Strongly influenced	Total
Sex				
Male	22.2	50.9	26.9	100.0
Female	27.1	54.0	18.9	100.0
Age				
16 to 19yrs	24.0	52.5	23.5	100.0
20yrs and above	25.3	48.4	26.3	100.0
School				
First class sch.	24.4	50.3	25.3	100.0
Second class sch.	23.9	53.6	22.5	100.0
Region				
Zone 1	22.0	56.1	21.9	100.0
Zone 2	25.6	49.2	25.2	100.0
Zone 3	31.1	33.3	35.6	100.0

From Table 11 it can be seen that the percentage of males who are strongly influenced by the second factor is much higher than those who are least influenced. However, the percentage of females who are strongly influenced by the second factor is much less than those who are least influenced. This suggests that males are strongly influenced by the second factor (quality assurance) than their female counterpart.

Almost the percentages of respondents within the age groups are influenced in the extreme by the second factor. Accordingly, quality assurance might be independent of a respondents' age.

The result of chi-square test of independence of the five factors on the demographic characteristics of the respondents is presented in Table 13. Generally, if a *p*-value is less than 0.05, then the factor depends on the demographic characteristic of the respondents.

Table 13: Chi-Square Test of Independence for the Extracted Factors on Demographic Characteristics

	Demographic Characteristics							
	Sex		Age		School		Region	
Factor	Chi-sq	*p*-val.	Chi-sq	*p*-val.	Chi-sq	*p*-val.	Chi-sq	*p*-val.
Factor 1	31.935	**0.000**	1.053	0.591	33.956	**0.000**	7.977	0.092
Factor 2	10.283	**0.006**	0.618	0.734	1.549	0.461	12.030	**0.017**
Factor 3	36.485	**0.000**	1.596	0.450	14.845	**0.001**	14.258	**0.007**
Factor 4	4.057	0.132	0.168	0.919	0.296	0.863	7.780	0.100
Factor 5	2.212	0.331	1.179	0.555	21.810	**0.000**	7.907	0.095

p-values less than 0.05 are highlighted

From Table 12, it can be observed that none of the *p*-values relating to age is less than 0.05. This implies that all the five factors are independent on a respondent's age. On the other hand, some of the *p*-values relating to sex, school and region are less than 0.05. For instance, the first factor (Academic resources) significantly depends on sex and school that the respondent is attending. The second factor (quality assurance) is considerably dependent on sex and region. The third factor (cost of study) is found to be dependent on respondents' sex, school and region. Factor four (influence of others) is seen to be independent on respondents' demographic characteristic. Furthermore, the fifth factor (discipline) depends solely on the school that the respondents attend.

CHAPTER SIX

SUMMARY, DISCUSSION AND CONCLUSIONS

In this study, we have been talking about factor analysis of prospective applicants' choice of the University of Cape Coast. The findings of this study are summarized in this chapter. Particular attention is given to the research questions posed for the study and how the results relate to the literature review in Chapter Two. The necessary recommendations and implications for future studies would be given based on the findings of this research.

SUMMARY

The objectives of this study were outlined in Chapter One. The main purpose of this study is to determine the latent factors that influence applicants' choice of the University of Cape Coast. In order to achieve this purpose, two stages of data collection were adopted. The first stage involves the collection of secondary data from the Students Records & Management Information Section of the University of Cape Coast. This data assisted in getting 20 out of over 350 schools from which students get admission into the University of Cape Coast. Details of how the schools were selected are presented in Chapter One. The second stage involves the sampling of respondents from these 20 selected Senior High Schools. Questionnaires involving 28 items were administered to 1142 respondents from the selected schools. The opinions of respondents on these indicators were analyzed.

Of these 1142 respondents, 697 were males, 442 were females and 3 did not indicate their gender. Most of the respondents (1045) are aged 16 to 19 years. Ghana National Senior High School in Cape Coast had the highest number of respondents who would like to attend UCC. Most of the respondents (326) living in Greater Accra Region would like to attend UCC.

The results obtained in the preliminary analysis were fascinating. The bar charts for the frequency distribution of responses revealed that some of the indicators display similar pattern. From the correlation matrix, it was noted that the 28 indicator variables can be grouped into six latent variables. Almost all the indicators in each group seem to have similar pattern of distribution.

The analysis of the indicators based on a scree plot and eigenvalue-greater-than-one rule indicated that between six and eight latent factors can practically explain the correlation among the 28 indicators.

The value of the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy for this data was 0.811. This value indicated that the data is appropriate for factoring although the correlations among the variables were generally low. Also, the large statistic value (5776) of the Bartlett's test of sphericity supported that the data is appropriate for factoring.

Further analysis in Chapter Five was based on the results obtained in Chapter Four. It was noted that the number of factors influencing applicants' choice of UCC could be less than 21. Based on the result of the preliminary analysis, it was suggested that the factors could be between six and eight. It was found that when the factors were unrotated, there was one major factor accounting for the variation in the data. This is because almost all the other factors identified in the unrotated factor loadings had either one or two indicators loading highly on it. The major factor identified in the unrotated factor loading is *academic resources*.

Conversely, the rotated factor loadings indicated that there were eight underlying factors. Base on the interpretations of these factors, it was observed that five factors actually influence applicants' choice of the University of Cape Coast. The most important factor is academic resources followed by quality assurance. The other factors are cost of study, influence of others and discipline. Comparably, academic resources emerged as a common factor from both the rotated and the unrotated factor solutions.

The scores of each respondent were used to find the extent of influence of the extracted factors on the demographic characteristics. After performing a chi-square test of independence, it was found that academic resources is dependent on respondents' sex and school whereas quality assurance is dependents on sex and region. In particular, females and respondents from second class schools are strongly influenced by academic resources. Also, cost of study is dependent on respondents' sex, school and region whiles discipline is solely dependent on school. It is worth mentioning that apart from the fourth factor, all the extracted factors depends on at least one demographic characteristic of the respondents.

DISCUSSION

The relationships between the bar charts for the frequency distributions of responses are systematically discussed. Also, the correlation matrix of the indicator variables is discussed. Similarities between the two results would be clarified. The results of the eigenvalue analysis and the scree plot would also be emphasized. The section would also focus on the rotated and unrotated factor solutions. Finally, the extracted factors which actually influence applicants' choice of UCC would be highlighted.

Comparison of the bar charts for the 28 indicator variables revealed some interesting results. First, almost all the indicator variables on cost of study had similar pattern of distribution. Thus, majority of the respondents were undecided on these indicators. This is not surprising because, most (68 percent) of the students who would like to attend the University of Cape Coast are living in Greater Accra, Ashanti and Western Regions. Hence, it would be difficult for them to decide on the cost of study in UCC. Appendix II(c) shows the frequency distribution of responses by Regions. Secondly, the rest of the groups identified from the pattern of distribution indicate that very few respondents strongly disagree whiles a lot more tend to be in agreement with the indicators in these groups. This confirms that applicants' choice of UCC is indeed based on these factors.

The most important question to ask when conducting factor analysis is whether or not the data is appropriate for factoring. First, one can subjectively examine the correlation matrix. High correlations among the indicator variables show that the variables can be grouped into homogeneous sets such that each set of variables measure the same underlying factor. An observation of the correlation matrix in Table 3 may give the impression that the variables do not have much in common. This is because of the generally low correlations among the indicator variables. However, the low correlation among the indicator variables is consistent with similar studies. As reviewed in the literature, MacCann et al. (2009) observed low correlations ranging from 0.12 to 0.23 in their study. Also, Schwarz and Hippler (1995) identified the problem of low correlations in their study. The ability to identify homogeneous set of indicators was as result of the high KMO value of 0.811 which gave the indication that the data is appropriate for factoring. Furthermore, the Bartlett's test of sphericity suggested that the correlation matrix is not an identity matrix.

As stated earlier in Chapter Four, six groups of indicators were obtained from the correlation matrix. Some of the indicators in these groups had similar pattern. Besides, some of the groups identified in the correlation matrix were similar to the groups identified in the plots of bar charts. As shown in Figure 7, the indicators in Group 1 had similar pattern of distribution. This confirms that although the correlations were generally low, there were homogeneous sets of variables measuring the same underlying factors.

The analysis of the eigenvalues and the scree plots in Chapter Four also indicated the number of factors to extract. In this study, identification of the exact number of factors to extract was basically dependent on the eigenvalue-greater-than-one rule. This is because the scree plot is a diagram which serves as a guard in identifying the factors. However, the eigenvalue analysis is a more analytic way of identifying the exact number of factors.

From the unrotated matrix, almost all the indicators loaded highly on the first factor. Most of these indicators which loaded highly on Factor 1 are issues on academic resources. The implication is that the most important factor which influences applicants' choice of UCC is academic resources. The unrotated factor model was not adopted because the other factors were not obvious. Therefore, both varimax and quartimax rotations were carried out in order to interpret these factors clearly. The results of the two rotations were almost the same. In each case, academic resources tend to be the most important factor influencing applicants' choice of UCC. The reason for approving the rotated factor model is its ability of identifying the other factors.

It must be realized that the final decision on the exact number of factors to extract depends on the plausibility of factor interpretation. Accordingly, the eight factors extracted in the rotated factor model were reduced to five. This was done because some of these factors had similar interpretations. For instance the indicators which represented the second factor are issues on quality assurance. Also, the indicators constituting the sixth factor largely reflect quality assurance. This indicated that the sixth factor cannot be uniquely interpreted. Similar situation occurred in the case of the fourth and seventh factors. Each of these set of indicators shows that applicants are influenced significantly by friends, relatives and counselors. Hence, this factor was labeled influence of others. In brief, five factors give the most comprehensive interpretation of the factor solution.

Relating the results of this study to the literature review, three factors identified in this study are similar to that of Wang (2009). Wang identified service/facility-related factor, economic-related factor and advice as some of the factors influencing international students' choice of universities. These factors are labeled in this study as academic resources, cost of study and influence of others.

Also, Schoenherr (2009) found the availability of financial aid to be the most important factor in predicting whether a student will attend a higher-tiered or a lower-tiered university. In this study, the third most important factor influencing applicants' choice of UCC is the cost of study.

Another study by Paulsen (1990) revealed that one of the most important factors in determining college aspirations among blacks is influence of significant others. This factor is also identified in this study.

From the above, it can be noted that the results obtained from this study is no departure from results of similar studies. Relatively, it is consistent with what some researchers have found.

CONCLUSIONS

This study presents the results of a multivariate analysis (with special focus on factor analysis) of the factors influencing applicants' choice of UCC. The results of the study suggest that senior high school students, when selecting the University of Cape Coast, base their decision on five main factors.

These five factors in order of importance are academic resources available at UCC, quality assurance, cost of study at UCC, influence of others (friends, relatives and counselors) and the level of discipline at the University of Cape Coast. With the exception of influence of others, it was noted that the extracted factors are dependent on at least one of the demographic characteristics of the respondents.

It was observed that the first two factors are most influential in applicants' selection of UCC. This was because of the extent of negative skewness on these indicators.

RECOMMENDATIONS

Wide variability exists in opinions on the indicators of the third and fourth factors. On the whole, cost of study is a controversial issue among the respondents involved in this study. Also, the issue of applicants being influenced by others is debatable. Consequently, it is recommended that issues on cost of study as well as influence of others would be investigated further to know the exact views of applicants. Nevertheless, these two factors contribute to applicants' selection of UCC.

Apart from these five factors which fulfils the main purpose of the study, applicants are usually influenced by some factors that were not identified in the factor solution. Closeness of applicants' residence (proximity) is one such factor worth observing.

Results from this study shows that prospective applicants of UCC rated some factors higher than others. As a result, the university administrators admitting students from Senior High Schools should develop procedures that would emphasize these factors. Possible actions include providing more resources for research and athletics, maintaining a conducive learning environment, employing competent and friendly staff.

With regards to the issues on quality assurance, the researcher recommends that the number of students who are given special awards/prizes for outstanding academic performance should be increased. This would motivate more students to study and help the university to produce competent graduates. Also, quality academic programmes should be introduced in UCC.

Based on the findings of this study, it is recommended that future studies should utilize other itemized rating scales such as the semantic differential and Staple scale as the survey instruments. This could help reduce the problem of low correlations among the indicators. Also, it is suggested that future studies should include more schools as they may help to clarify issues on cost of study and influence of others.

REFERENCES

Ampimah, B. C. (2007). *Factor analysis of applicants' choice of Cape Coast Polytechnic*. Unpublished master's thesis, University of Cape Coast, Cape Coast.

Andriotis, K., Agiomirgianakis, G., & Mihiotis, A. (2008). Measuring tourist satisfaction: A factor-cluster segmentation. *Journal of Vacation Marketing*, *14*(3), 221-235.

Auger, P., & Devinney, T. M. (2007). Do what consumers say matter? The Misalignment of Preferences with Unconstrained Ethical Intentions. *Journal of Business Ethics*, *76*, 361-386.

Avery, C., & Hoxby, C. M. (2004). Do and should financial aid packages affect students' college choices? The economics of where to go, when to go, and how to pay for it. *A National Bureau of Economic Research Conference Report.* University of Chicago Press.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York, NY: The Guilford Press.

Canale, J. R., & Dunlap, L. (1996). The relative importance of various college characteristics to students in influencing their choice of a college. *College Student Journal, 30*(2), 214.

Chapman, R. G., & Jackson, R. (1987). College choices of academically able students: The influence of no-need financial aid and other factors. *Research Monograph 10***.** New York, NY: College Board.

Cheng, B. (2006). Editor's Review. College choices: The economics of where to go, when to go, and how to pay for it. *Harvard Educational Review, 76*(1), 110-119.

DeCoster, J. (1998). *Overview of factor analysis*. Retrieved October 16, 2011 from *http://www.stat-help.com/notes.html*

Dillon, W. R., & Goldstein, M. (1984). *Multivariate Analysis*. New York, NY: Wiley.

Falsey, B., & Haynes, B. (1984). The college cannel: Private and public schools reconsidered. *Sociology of Higher Education 57*, 11-22

Field, A. P. (2005). *Discovering statistics using SPSS* (2nd ed.). London: Sage.

Geiger, R. L. (2002). The competition for high-ability students: Universities in a key market place. *The future of the city of the intellect*. Stanford University Press.

Goenner, C., & Snaith, S. (2004). Assessing the effects of increased admission standards. *College and University Journal, 80*(1), 29-34.

Hayden, M. (2000). Factors that influence the college choice process for African American students. *The Review of Higher Education*. *25*(3), 315-330

Hossler, D., & Gallagher, L. (1987). Studying college choice: A three-phase model and the implication for policy makers. *College and University Journal, 62*(3)**,** 207-221.

Inkoom, G. (2010). *Factor analysis of applicants' choice of the University Cape Coast*. Unpublished master's thesis, University of Cape Coast, Cape Coast.

Kallio, R. E. (1995). Factors influencing the college choice decisions of graduate students. *Research in Higher Education*, *36*(1), 109-124.

Kern, C. W. K. (2000). College choice influence: Urban high school students respond. *Community College Journal of Research and Practice*, *24*(6), 487-494.

Kim, D. (2004). The effect of financial aid on students' college choice: Differences by racial groups. *Research in Higher Education, 45*(1), 47.

Lozano, L. M., Garcia-Cueto, E., & Muniz, J. (2008). *Effects of the number of response categories on the reliability and validity of rating scales*. Retrieved July 22, 2011 from *http://doresearch.wordpress.com*.

MacCann, C., Dunkworth, A. L., & Roberts, R. D. (2009). Empirical identification of the major facets of conscientiousness. *Journal of Learning and Individual Differences*, *19*, 451-458.

Manski, C. F., & Wise, D. A. (1983). Econometric modeling of student behavior. *College Choice in America*. In C. F. Manski & D. A. Wise (Eds.), Cambridge, MA: Harvard University Press.

Miller, A. S., & Mitamura, T. (2003). Are surveys on trust trustworthy? *Social Psychology Quarterly, 66*(1), 62–70.

Naresh. K. M. (2004). *Marketing research: An applied orientation* (4th ed.). New Jersey, Upper Saddle River: Prentice Hall.

Nkansah, B. K., Gordor, B. K., & Howard, N. K. (2007). Factor analysis of bath soap consumerism: A case study of the University of Cape Coast, *Journal of Natural Science, 3*(1), 30-41.

Ozturk, M. A. (2010). A confirmatory factor analysis of educators' attitudes toward educational research. *Educational Sciences, Theory & Practice, 11*(2), 737-747.

Paulsen, M. B. (1990). College choice: Understanding student enrollment behavior. *ASHE-ERIC Higher Education Report, 6,* George Washington University, Washington D. C.

Rencher, A. C. (2002). *Methods of Multivariate Analysis* (2nd ed.). New York, NY: John Wiley and Sons.

Rucinski, D. (1993). Rush to judgment? 'Fast reaction polls in Anita Hill-Clarence Thomas Controversy'. *Public Opinion Quarterly, 57*(4), 575-592.

Russell, C. (1980). Survey of grade 12 students' postsecondary plans and aspirations. *Canadian Department of Education*, Manitoba.

Schoenherr, H. J. (2009). Beyond academic reputation: Factors that that influences the college of first choice for high achieving students.

Department of Adult, Career and Higher Education. University of South Florida.

Schwarz, N., Grayson, C. E., & Knauper, B. (1998). Formal features of rating scales and the interpretation of question meaning. *International Journal of Public Opinion Research*, *10*(2), 177–183.

Schwarz, N., & Hippler, H. J. (1995). Subsequent questions may influence answers to preceding questions in mail surveys. *Public opinion Quarterly*, *59*(1), 93-97.

Serve Africa.info, (2012). Survey results: Top secondary schools in Ghana. Retrieved March 13, 2012, from *http://serveafrica.info/secondary-school-satisfaction-survey.*

Sharma, S. (1996). *Applied Multivariate Techniques*. New York: Wiley.

Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics* (3rd ed.). California State University, Northridge: Harpe Collins College.

University of Salford (2006). Attitude scales. Retrieved July 16, 2011, from *http://www.chssc.salford.ac.uk/healthSci/resmeth/attitude.com.*

Wang, X. (2009). Factors influencing international students' choice of universities. *A Country-Specific Analysis*. Charleston, South Carolina

Zwick, W. R., & Velicerh, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin*, *99*(3), 432-442.

APPENDICES

APPENDIX A

Questionnaire on Latent Factors that affect Applicants' Choice of UCC.

This questionnaire is designed by an M.Phil Statistics student for academic purpose. It is design to help seek information from some selected secondary school students on some of the reasons that are considered in their decision to come to the University of Cape Coast. Your view will help to determine the Latent Factors that influence the choice of a potential University of Cape Coast student. It will also help the university authorities to improve upon the areas that attract a prospective University of Cape Coast student. Any information given on this questionnaire will be held in utmost confidence.

Please complete this questionnaire to the best of your knowledge. Thank you very much.

Part A

Please tick only one option for each question

1. Sex Male [ ]
   Female [ ]
2. Age 12 – 15 years [ ]
   16 – 19 year [ ]
   20 years and above [ ]
3. School……………………………………………………………
4. Region of birth ………………………………………………………

Appendix A (continued)

Please tick one appropriate option for each question.

Reasons	Strongly disagree	Disagree	Undecided	Agree	Strongly agree
1. The only university that offers my choice of programme.					
2. It is located in Central region where there are many tourist attractions.					
3. It is base on my counselor's recommendation.					
4. The university is close to my place of residence (proximity).					
5. It will be an opportunity to meet old friends.					
6. I prefer the grading system of this university to other universities.					
7. It is cheaper to get accommodation in this university.					
8. The university offers quality programmes.					
9. Students who complete this university are competent.					
10. Most of my friends and family members attended this university.					
11. The physical structure of the university is very attractive.					
12. I like the university because it is located by the sea.					
13. Both students and lecturers are punctual for lectures.					
14. The workers of the university are friendly.					
15. I like the social life of the people in this university.					
16. A lot of facilities are available for research in this university.					
17. It was recommended to me by friends/ relatives.					
18. The cost of living is relatively cheaper in Cape Coast.					
19. The administration monitors performance of student very well.					
20. Students comport themselves well during lecture hours.					
21. The administration regulates activities of the staff very well.					
22. Academic user fee is moderate.					
23. The university is recognized worldwide.					
24. There are many facilities for athletics.					
25. A lot of financial sponsorship is available to students.					
26. I like the way students study in this university.					
27. The university has a competent staff.					
28. The university has conducive learning environment.					

APPENDIX B

Frequency Distribution of Responses

(a) Frequency and Percentage Distribution by Sex

Sex	Frequency	Percent	Cumulative
Males	697	61.0	61.0
Females	442	38.7	99.7
Total	1139	99.7	
Missing	3	0.3	100.0
Total	1142	100.0	

(b) Frequency and Percentage Distribution by Age

Age	Frequency	Percent	Cumulative
16-19years	1045	91.5	91.5
20yrs & above	96	8.4	99.9
Total	1141	99.9	
Missing	1	0.1	100.0
Total	1142	100.0	

(c) Frequency and Percentage Distribution by Regions

Regions	Frequency	Percent	Cumulative
Central	141	12.4	12.4
Greater Accra	326	28.5	40.9
Eastern	107	9.4	50.3
Volta	29	2.5	52.8
Ashanti	254	22.2	75.0
Brong Ahafo	38	3.3	78.3
Western	197	17.3	95.6
Northern	18	1.6	97.2
Upper East	13	1.1	98.3
Upper West	3	0.3	98.6
Foreigners	11	1.0	99.6
Total	1137	99.6	
Missing	5	0.4	100.0
Total	1142	100.0	

Appendix B continued

(d) Frequency and Percentage Distribution by Schools

Schools	Frequency	Percent	Cumulative
St Augustine, Cape Coast	64	5.6	5.6
Aggery Memorial, Cape Coast	73	6.4	12.0
Adisadel College, Cape Coast	46	4.0	16.0
Mfantsipim, Cape Coast	46	4.0	20.1
Ghana National Cape Coast	89	7.8	27.8
Mfanstiman Girls, Saltpond	62	5.4	33.3
Accra Academy, Accra	77	6.7	40.0
Okuapeman, Akropong	48	4.2	44.2
Ghana Sec. Sch, Koforidua	37	3.2	47.5
Pope Johns, Koforidua	66	5.8	53.2
Opoku Ware, Kumasi	48	4.2	57.4
Anglican Sec. Sch. Kumasi	50	4.4	61.8
TI Ahmadiya , Kumasi	58	5.1	66.9
Yaa Asantewaa Girls, Kumasi	44	3.9	70.8
Prempeh College, Kumasi	69	6.0	76.8
University Practice, Cape Coast	44	3.9	80.6
Fijai Sec. Sch. Takoradi	61	5.3	86.0
St. Johns Sec. Sch. Takoradi	66	5.8	91.8
GSTS, Takoradi	38	3.3	95.1
Porter Girls, Takoradi	56	4.9	100.0
Total	1142	100.0	

(e) Frequency and Percentage Distribution of Opinions by V_1

Opinion	Frequency	Percent	Cumulative
1	300	26.3	26.3
2	369	32.3	58.6
3	142	12.4	71.0
4	182	15.9	86.9
5	147	12.9	99.8
Total	1140	99.8	
Missing	2	0.2	100.0
Total	1142	100.0	

Appendix B continued

(f) Frequency and Percentage Distribution of Opinions by V_2

Opinion	Frequency	Percent	Cumulative
1	132	11.6	11.6
2	99	8.6	20.2
3	77	6.7	26.9
4	390	34.2	61.1
5	442	38.7	99.8
Total	1140	99.8	
Missing	2	0.2	100.0
Total	1142	100.0	

(g) Frequency and Percentage Distribution of Opinions by V_3

Opinion	Frequency	Percent	Cumulative
1	206	18.0	18.0
2	362	31.7	49.7
3	310	27.2	76.9
4	190	16.6	93.5
5	74	6.5	100.0
Total	1142	100.0	

(h) Frequency and Percentage Distribution of Opinions by V_4

Opinion	Frequency	Percent	Cumulative
1	563	49.3	49.3
2	241	21.1	70.4
3	90	7.9	78.3
4	137	12.0	90.3
5	109	9.5	99.8
Total	1140	99.8	
Missing	2	0.2	100.0
Total	1142	100.0	

Appendix B continued

(i) Frequency and Percentage Distribution of Opinions by V_5

Opinion	Frequency	Percent	Cumulative
1	127	11.1	11.1
2	195	17.1	28.2
3	191	16.7	44.9
4	408	35.7	80.6
5	220	19.3	99.9
Total	1141	99.9	
Missing	1	0.1	100.0
Total	1142	100.0	

(j) Frequency and Percentage Distribution of Opinions by V_6

Opinion	Frequency	Percent	Cumulative
1	168	14.7	14.7
2	131	11.5	26.2
3	175	15.3	41.5
4	306	26.8	68.3
5	361	31.6	99.9
Total	1141	99.9	
Missing	1	0.1	100.0
Total	1142	100.0	

(k) Frequency and Percentage Distribution of Opinions by V_7

Opinion	Frequency	Percent	Cumulative
1	212	18.6	18.6
2	197	17.2	35.8
3	425	37.2	73.0
4	215	18.8	91.8
5	92	8.1	99.9
Total	1141	99.9	
Missing	1	0.1	100.0
Total	1142	100.0	

Appendix B continued

(l) Frequency and Percentage Distribution of Opinions by V_8

Opinion	Frequency	Percent	Cumulative
1	29	2.5	2.5
2	55	4.9	7.4
3	110	9.6	17.0
4	592	51.8	68.8
5	356	31.2	100.0
Total	1142	100.0	

(m) Frequency and Percentage Distribution of Opinions by V_9

Opinion	Frequency	Percent	Cumulative
1	23	2.0	2.0
2	51	4.5	6.5
3	128	11.2	17.7
4	442	38.7	56.4
5	498	43.6	100.0
Total	1142	100.0	

(n) Frequency and Percentage Distribution of Opinions by V_{10}

Opinion	Frequency	Percent	Cumulative
1	343	30.0	30.0
2	325	28.5	58.5
3	109	9.5	68.0
4	245	21.5	89.5
5	120	10.5	100.0
Total	1142	100.0	

(o) Frequency and Percentage Distribution of Opinions by V_{11}

Opinion	Frequency	Percent	Cumulative
1	43	3.8	3.8
2	97	8.5	12.3
3	178	15.6	27.8
4	538	47.1	75.0
5	286	25.0	100.0
Total	1142	100.0	

Appendix B continued

(p) Frequency and Percentage Distribution of Opinions by V_{12}

Opinion	Frequency	Percent	Cumulative
1	311	27.2	27.2
2	236	20.7	47.9
3	132	11.6	59.5
4	271	23.7	83.2
5	192	16.8	100.0
Total	1142	100.0	

(q) Frequency and Percentage Distribution of Opinions by V_{13}

Opinion	Frequency	Percent	Cumulative
1	40	3.5	3.5
2	74	6.5	10.0
3	419	36.7	46.7
4	301	26.3	73.0
5	308	27.0	100.0
Total	1142	100.0	

(r) Frequency and Percentage Distribution of Opinions by V_{14}

Opinion	Frequency	Percent	Cumulative
1	29	2.5	2.5
2	50	4.4	6.9
3	605	53.0	59.9
4	336	29.4	89.3
5	122	10.7	100.0
Total	1142	100.0	

(s) Frequency and Percentage Distribution of Opinions by V_{15}

Opinion	Frequency	Percent	Cumulative
1	113	9.9	9.9
2	157	13.7	23.6
3	209	18.3	41.9
4	434	38.0	79.9
5	229	20.1	100.0
Total	1142	100.0	

Appendix B continued

(t) Frequency and Percentage Distribution of Opinions by V_{16}

Opinion	Frequency	Percent	Cumulative
1	32	2.8	2.8
2	76	6.7	9.5
3	286	25.0	34.5
4	486	42.6	77.1
5	262	22.9	100.0
Total	1142	100.0	

(u) Frequency and Percentage Distribution of Opinions by V_{17}

Opinion	Frequency	Percent	Cumulative
1	190	16.6	16.6
2	249	21.8	38.4
3	160	14.0	52.5
4	357	31.3	83.7
5	186	16.3	100.0
Total	1142	100.0	

(v) Frequency and Percentage Distribution of Opinions by V_{18}

Opinion	Frequency	Percent	Cumulative
1	387	33.9	33.9
2	226	19.8	53.7
3	230	20.1	73.8
4	199	17.4	91.2
5	97	8.5	99.7
Total	1139	99.7	
Missing	3	0.3	100.0
Total	1142	100.0	

Appendix B continued

(w) Frequency and Percentage Distribution of Opinions by V_{19}

Opinion	Frequency	Percent	Cumulative
1	23	2.0	2.0
2	41	3.6	5.6
3	200	17.5	23.1
4	361	31.6	54.7
5	517	45.3	100.0
Total	1142	100.0	

(x) Frequency and Percentage Distribution of Opinions by V_{20}

Opinion	Frequency	Percent	Cumulative
1	27	2.4	2.4
2	52	4.6	7.0
3	415	36.3	43.3
4	368	32.2	75.5
5	280	24.5	100.0
Total	1142	100.0	

(y) Frequency and Percentage Distribution of Opinions by V_{21}

Opinion	Frequency	Percent	Cumulative
1	25	2.2	2.2
2	46	4.0	6.2
3	499	43.7	49.9
4	355	31.1	81.0
5	217	19.0	100.0
Total	1142	100.0	

(z) Frequency and Percentage Distribution of Opinions by V_{22}

Opinion	Frequency	Percent	Cumulative
1	122	10.7	10.7
2	124	10.9	21.6
3	505	44.2	65.8
4	279	24.4	90.2
5	110	9.6	99.8
Total	1140	99.8	
Missing	2	0.2	100.0
Total	1142	100.0	

Appendix B continued

(α) Frequency and Percentage Distribution of Opinions by V_{23}

Opinion	Frequency	Percent	Cumulative
1	66	5.8	5.8
2	135	11.8	17.6
3	250	21.9	39.5
4	435	38.1	77.6
5	256	22.4	100.0
Total	1142	100.0	

(β) Frequency and Percentage Distribution of Opinions by V_{24}

Opinion	Frequency	Percent	Cumulative
1	28	2.5	2.5
2	100	8.7	11.2
3	507	44.4	55.6
4	364	31.9	87.5
5	141	12.3	99.8
Total	1140	99.8	
Missing	2	0.2	100.0
Total	1142	100.0	

(γ) Frequency and Percentage Distribution of Opinions by V_{25}

Opinion	Frequency	Percent	Cumulative
1	147	12.9	12.9
2	160	14.0	26.9
3	595	52.1	79.0
4	178	15.6	94.6
5	62	5.4	100.0
Total	1142	100.0	

Appendix B continued

(δ) Frequency and Percentage Distribution of Opinions by V_{26}

Opinion	Frequency	Percent	Cumulative
1	27	2.3	2.4
2	40	3.5	5.8
3	74	6.5	12.3
4	397	34.8	47.1
5	602	52.7	99.8
Total	1140	99.8	
Missing	2	0.2	100.0
Total	1142	100.0	

(ε) Frequency and Percentage Distribution of Opinions by V_{27}

Opinion	Frequency	Percent	Cumulative
1	22	1.9	1.9
2	35	3.1	5.0
3	357	31.3	36.3
4	409	35.8	72.1
5	319	27.9	100.0
Total	1142	100.0	

(θ) Frequency and Percentage Distribution of Opinions by V_{28}

Opinion	Frequency	Percent	Cumulative
1	22	1.9	1.9
2	30	2.7	4.6
3	79	6.9	11.5
4	505	44.2	55.7
5	506	44.3	100.0
Total	1142	100.0	

APPENDIX C

Communalities

Indicators	Initial	Extraction
V_1	1.000	0.462
V_2	1.000	0.701
V_3	1.000	0.579
V_4	1.000	0.450
V_5	1.000	0.441
V_6	1.000	0.394
V_7	1.000	0.574
V_8	1.000	0.534
V_9	1.000	0.530
V_{10}	1.000	0.524
V_{11}	1.000	0.450
V_{12}	1.000	0.702
V_{13}	1.000	0.729
V_{14}	1.000	0.392
V_{15}	1.000	0.559
V_{16}	1.000	0.530
V_{17}	1.000	0.605
V_{18}	1.000	0.588
V_{19}	1.000	0.466
V_{20}	1.000	0.719
V_{21}	1.000	0.411
V_{22}	1.000	0.466
V_{23}	1.000	0.609
V_{24}	1.000	0.479
V_{25}	1.000	0.354
V_{26}	1.000	0.499
V_{27}	1.000	0.500
V_{28}	1.000	0.458

APPENDIX D

Factor Score Coefficients

Indicators	Factors				
	1	2	3	4	5
V_1	-0.134	0.062	0.007	0.086	-0.069
V_2	-0.013	0.058	-0.044	-0.075	-0.053
V_3	0.026	-0.004	0.016	0.442	0.040
V_4	-0.034	-0.020	0.178	0.159	-0.115
V_5	0.021	-0.116	-0.048	0.062	-0.027
V_6	-0.137	0.059	0.060	-0.111	0.007
V_7	0.012	-0.059	0.387	-0.031	0.039
V_8	0.021	0.000	-0.048	-0.113	0.001
V_9	-0.069	0.429	-0.032	0.077	-0.106
V_{10}	0.004	-0.076	-0.076	0.364	0.030
V_{11}	0.295	0.043	-0.029	-0.020	-0.110
V_{12}	-0.036	-0.057	-0.010	-0.031	0.005
V_{13}	-0.041	-0.102	-0.014	0.027	0.513
V_{14}	0.287	-0.156	0.021	0.096	0.110
V_{15}	-0.022	0.026	0.003	0.035	-0.051
V_{16}	0.318	0.011	-0.025	-0.055	-0.101
V_{17}	-0.092	0.071	-0.049	0.386	-0.030
V_{18}	-0.069	0.077	0.411	-0.018	-0.093
V_{19}	-0.141	0.239	-0.001	-0.027	0.162
V_{20}	-0.024	-0.085	-0.030	0.017	0.492
V_{21}	-0.061	0.231	0.085	-0.045	0.106
V_{22}	0.030	-0.064	0.306	-0.028	0.069
V_{23}	0.048	-0.221	-0.046	-0.109	0.093
V_{24}	0.341	-0.185	0.014	0.003	-0.004
V_{25}	-0.046	0.014	0.252	-0.059	-0.058
V_{26}	-0.009	0.401	-0.020	-0.026	-0.140
V_{27}	0.238	0.162	-0.014	0.059	0.001
V_{28}	0.198	0.148	-0.049	-0.149	-0.073

APPENDIX E

Contingency Tables

(a) Influence on the Third Factor by Demographic Characteristics

Demographic characteristics	Level of influence			Total
	Least influenced	Moderately influenced	Strongly influenced	
Sex				
Male	18.1	55.6	26.3	100.0
Female	33.5	48.3	18.2	100.0
Age				
16 to 19yrs	24.4	52.2	23.4	100.0
20yrs and above	21.1	58.9	20.0	100.0
School				
First class sch.	19.3	53.6	26.7	100.0
Second class sch.	28.3	51.6	20.1	100.0
Region				
Zone 1	25.7	50.9	23.4	100.0
Zone 2	20.8	56.9	22.3	100.0
Zone 3	40.0	31.1	28.9	100.0

Appendix E continued

(b) Influence on the Fourth Factor by Demographic Characteristics

Demographic characteristics	Level of influence			Total
	Least influenced	Moderately influenced	Strongly influenced	
Sex				
Male	24.4	55.8	19.8	100.0
Female	20.5	55.4	24.1	100.0
Age				
16 to 19yrs	22.9	55.7	21.4	100.0
20yrs and above	22.1	54.7	23.2	100.0
School				
First class sch.	22.2	56.5	21.3	100.0
Second class sch.	23.4	54.9	21.7	100.0
Region				
Zone 1	24.7	56.3	19.0	100.0
Zone 2	20.4	54.6	25.0	100.0
Zone 3	26.7	57.8	15.5	100.0

Appendix E continued

(c) Influence on the Fifth Factor by Demographic Characteristics

Demographic characteristics	Level of influence			Total
	Least influenced	Moderately influenced	Strongly influenced	
Sex				
Male	22.8	54.0	23.2	100.0
Female	26.2	53.3	20.5	100.0
Age				
16 to 19yrs	24.3	53.3	22.4	100.0
20yrs and above	22.1	59.0	18.9	100.0
School				
First class sch.	20.9	50.9	28.2	100.0
Second class sch.	27.0	56.1	16.9	100.0
Region				
Zone 1	27.1	52.9	20.0	100.0
Zone 2	20.7	55.0	24.2	100.0
Zone 3	20.0	51.1	28.9	100.0