STATISTICAL MODELS FOR STOCKS AND FLOWS OF STUDENTS IN AN EDUCATIONAL SYSTEM

BY

JOHN OKOTH OWINO

This thesis is submitted in fulfilment for the degree of Doctor of Philosophy in Mathematical Statistics in the Department of Mathematics,

UNIVERSITY OF NAIROBI

November, 1989.

DECLARATION

This thesis is my original work and has not been presented for a degree in any other University.

Signature: ______________

JOHN OKOTH OWINO

This thesis has been submitted for examination with my approval as University supervisor.

Signature: Odhiambo 29/3/90

Prof. J.W. Odhiambo

DEPARTMENT OF MATHEMATICS,
UNIVERSITY OF NAIROBI,
P.O. BOX 30197,
NAIROBI,
KENYA.

CONTENTS

LIST OF TABLES

SUMMARY OF CONTENTS

In developing countries education patterns are constantly changing due to rapid population growth and other socio-economic factors. This trend calls for transition models which incorporate factors which are internal or external to the system. These models are used together with the theory of stochastic processes to define various measures of academic retention. Estimates of these measures are computed using the stocks and flows data of the primary school system in Kenya. An attempt is made to control the system in two ways. First, control is made via some quantifiable factors which affect the system, so as to achieve some future desired educational characteristics, Secondly, some desired educational characteristics are specified and the problem is to find the transition process that should be followed in order to achieve the targets.

In chapter I an overview of mathematical modelling as applied to hierarchical processes is given. A brief description of the work already done in the area of modelling hierarchical processes in general and education systems in particular is also given in this chapter.

Chapter II examines the homogeneity of the Kenyan primary education system between 1964 to 1980. This is done by partitioning the entire period into equal intervals and comparing the average education characteristics of each of these intervals. Any appreciable difference in these characteristics would suggest departure from homogeneity of the process over the considered period.

The results of chapter II suggest possible inhomogeneity in the Kenyan primary education system. For this reason in Chapter III the assumption of homogeneity in the education process is relaxed. It is suggested here that the transition process changes in time. In particular a study is made of a number of transition models which attempt to incorporate endogeneous factors in the system over a period of time by means of time dependent probability distribution functions. These models are used together with the theory of the time dependent Markov chains, to compute various measures of academic retention.

Chapter IV describes a model which traces the flow of a cohort of students through the Kenyan primary education system. For the purpose of this study the term cohort is used to denote a group of students regardless of age or socio-economic background, who enter the first grade in the same academic year. In particular the cohort transition model is used as an application of the more general Markov chain model described in chapter III.

In chapter V the transition process is modelled as a function of time dependent quantifiable factors. The proposed model is first used to describe some educational characteristics. Then an attempt is made to control some of the factors so as to achieve some desired future educational characteristics optimally.

Finally in Chapter VI we consider a control problem where the desired educational characteristics are specified and we aim at finding the transition process to be followed in order to achieve the desired targeted characteristics optimally.

The thesis ends with a few general remarks by way of conclusion, regarding the results obtained in the present work and possible problems for future research. These conclusions form the contents of Chapter VII.

Every model proposed in the thesis has been illustrated by computing numerical values of several educational characteristics. The results of such computations are given in tables throughout the thesis.

The theoretical contents of this thesis is mostly based on the theory of Markov chains, especially the time dependent Markov chains. Use is also made of linear regression models and statistical control theory in multivariate regression models.

ACKNOWLEDGEMENTS

My profound gratitude is due to my supervisor Prof. J.W. Odhiambo for all the constructive assistance and guidance throughout this study. His constant scrutiny of the work is surely what has led to its successful completion.

I am grateful to my lecturers Prof. M.S. Patel and Prof. J.A.M. Ottieno both of whom kindled my interest in the subject of Mathematical Statistics. Prior to taking up full proffessorial appointment at Moi University, Prof. M.S. Patel acted as one of my supervisors thereby setting up high standards which this work has all along strived to achieve.

My thanks are due to the British Council for the opportunity of a one year attachment at the London School of Economics between September 1986 to September, 1987. That period served as a very important external exposure time both academically and socially. Specifically, I am highly indebted to Dr. C.M. Philips who supervised my studies during that period. Her suggestions and guidance concerning this work proved to be very beneficial.

The completion of this work is also as a result of direct or indirect help of many others. I thank the entire staff of the mathematics department under the chairmanship of Prof. W. Ogana for all the encouragement and help rendered during the study period. I also thank the Kenya Bureau of Statistics for making the data on the Kenyan Primary education system accessible for analysis.

I highly appreciate the expert and diligent manner in which Mrs. Mary Okello typed this manuscript.

Lastly but not least I thank my entire family, the Owinos, for their assistance, understanding and cooperation during the time of this study.

CHAPTER I

INTRODUCTION

1.1 AN OVERVIEW OF EDUCATION PLANNING MODELS

Education activity is becoming more complex and planning its development more difficult. This complexity is due to a multiplicity of factors, such as population increase, economic constraints and so on. These factors are more pronounced in developing countries, such as Kenya, where population increase is rapid.

To evolve an efficient system, educational activities need to be coordinated, for which an abstract realisation of the real system is necessary. A good model will, in addition to adequately describing the past, also give reliable estimates for the future. In order to understand the function of an education system, it is useful to study the long-run implications of the present educational structure and parameters.

Mathematical modelling has gained prominence as a means of improving education planning. At present there are many types of models, some deal with the whole system, some with particular sectors of the system and others with specific institutions. The functional forms of the models are similarly varied. There are stochastic models Gani (1963), Thonstad (1967); mathematical programming models Marshall and Oliver (1970), Mc Namara (1973), Propoi (1978), Clowes (1972); regression models, Pickford (1974), Young and Vassiloiu (1974), Drui (1963); demographic models, Stone (1966), Young (1971); computer simulation models, Baisuck and Wallace (1970) and many other approaches.

In this study we consider stochastic models and in particular the markovian model as applied to education planning. A stochastic process is one which develops in time according to probabilistic laws. This means that we cannot predict its future behaviour with certainty ; the most that we can do is to attach probabilities to the various possible future states. Such processes occur widely in nature and their study has provided the impetus for the rapid development of the theory of stochastic processes in the past few decades.

Education can be considered as a hierarchical organisation. Students usually stay in a given grade for one academic year and then move to the next grade or leave the system as graduates or dropouts. This is the basic idea of the markovian model. When students graduate and leave the system or when they drop out due to illness, death or poor academic performance, the situation is akin to transition into absorbing states. Transition between grades is similar to that between non-absorbing states. Consequently, an absorbing Markov chain is often used, for which it is essential to identify the transition ratios between grades or from a grade into a final educational qualification. These ratios can be interpreted as transition probabilities and as such give rise to the usual stochastic matrix. The grades and final educations form the states of the process.

The states of the education system will be partitioned into two categories: non-absorbing states corresponding to the various grades within the system and absorbing states corresponding to the various final educations, mortality etc. Transition between grades is thus similar to that between non absorbing states. One important aspect of this model is that it gives a good study of the variate under consideration, the student, from entry to exit.

Suppose the states of the education system are denoted by the integers 1,2,3,...,N and the times by $t = 0,1,2,...$: We now let $p_{ij}(t)$ denote the probability that a student in state i at time t will transfer to state j at time $(t+1)$. This will give rise to the transition matrix

$$P(t) = \left(\left(p_{ij}(t)\right)\right), \; i,j = 1,2,\ldots,N$$

If we further assume that the system has r absorbing and s non-absorbing states, then the transition matrix will have the canonical form,

$$P(t) = \begin{bmatrix} I & 0 \\ & \\ G(t) & Q(t) \end{bmatrix}$$

where:

I is an r x r identity matrix giving transition probabilities between absorbing states;

0 is an r x s matrix of zeros giving the transition probabilities from absorbing to non-absorbing states;

$G(t) = \left(\left(g_{ik}(t)\right)\right)$ s x r, where $g_{ik}(t)$'s are the probabilities that a student in grade i at time t will graduate with final education k at time $(t+1)$, $i = 1,2,\ldots,s$ and $k = 1,2,\ldots,r$;

and

$Q(t) = \left(\left(q_{ij}(t)\right)\right)$ s x s, where $q_{ij}(t)$'s are the probabilities that a student in grade i at time t will be in grade j at time $(t+1)$; $i,j = 1,2,\ldots,s$.

The diagonal elements of $Q(t)$, $q_{ii}(t)$'s are the probabilities of a student repeating grade i, $i = 1,2,\ldots,s$. There is sometimes a tendency for a student obtaining a final education from the same school activity in which case we would have only one element in each column of $G(t)$. However, in most cases persons who successfully complete a given final education often try for a year or more at a more advanced school activity, thus, it is reasonable to classify them as having the education they finally successfully completed.

If a person is in one of the r final educations we shall say that he(she) is absorbed so that he(she) will not leave that state. This assumption could slightly be modified by introducing the probabilities of returning to certain types of schools after some time of work, illness etc. This may include, retraining of working staff, adult education programs and so on. It should be noted that the rows of $Q(t)$ and $G(t)$ all add to one, since a student either goes to one of the s school activities or leaves the school activities thus ending up in one of the r final education categories.

We now define,

$n_{ij}(t)$ to be the number of students in state i at time t who go to state j

at time t+1. We shall call these values the <u>flows</u> at t;

and

$n_i(t)$ to be the number of students in state i at time t. These we shall call the <u>stocks</u> at time t.

Then

$$\sum_{j=1}^{N} n_{ij}(t) = n_i(t)$$

and for each $n_{ij}(t)$ there is associated a probability $p_{ij}(t)$ of moving from state i at time t to state j at time t + 1, so that

$$\sum_{j=1}^{N} p_{ij}(t) = 1.$$

For a large population size we may therefore assume that the flow sequence at time t,

$$\{n_{i1}(t), n_{i2}(t), \ldots, n_{iN}(t)\}$$

has a multinomial distribution given by

$$f\left(n_{i1}(t), n_{i2}(t), \ldots, n_{iN}(t)\right) = \frac{n_i(t)!}{\prod_{j=1}^{N} n_{ij}(t)!} \prod_{j=1}^{N} p_{ij}(t)^{n_{ij}(t)}$$

So, the maximum likelihood estimate of $p_{ij}(t)$ will be

$$\hat{p}_{ij}(t) = n_{ij}(t)/n_i(t); \quad i,j = 1,2,\ldots,N,$$

which is the proportion of persons in state i at time t who go to state j at time t + 1.

Next, the probability distribution at time t is

$$\underline{p}(t) = \left(p_1(t), p_2(t), \ldots, p_N(t)\right)'$$

where,

$p_i(t)$ is the probability of an individual being in education category i at time t, i = 1,2,...,N.

If we consider the number of students or people in state i at time t as a random variable which takes the value $n_i(t)$ with probability $p_i(t)$ for i = 1,2,...,N; then the stock sequence at time t,

$$\{n_1(t), n_2(t), \ldots, n_N(t)\}$$

can be assumed to have a multinomial distribution given by

$$f\left(n_1(t), n_2(t), \ldots, n_N(t)\right) = \frac{N(t)!}{\prod_{j=1}^{N} n_j(t)!} \prod_1^N p_j(t)^{n_j(t)}$$

where,

$$N(t) = \sum_{j=1}^{N} n_j(t),$$

is the total number of people in the system at time t. The maximum likelihood estimate of $p_i(t)$ is therefore

$$\hat{p}_i(t) = n_i(t)/N(t), \quad i = 1,2,\ldots,N;$$

which is the proportion of persons in the various educational categories at time t.

When $t = 0$, for some base year, then

$$\underline{p}(0) = \big(p_1(0), p_2(0), \ldots, p_N(0)\big)'$$

is the initial probability vector.

An important property of an absorbing Markov chain is that the probability of the system being absorbed tends to one as the number of trials gets larger. That is, the components of $G(t)$, $g_{ik}(t)$'s, tend to one as the number of trials increase and so the components of $Q(t)$, $q_{ij}(t)$'s, tend to zero as the number of trials increase. If we therefore consider higher transitional probabilities and their limiting values with reference to the matrix $P(t)$ we can make the following interpretations:

(a) The probability that a student now in school grade i will be in any of the school grades n years later is interpreted as the fraction of students now in school grade i who will still be in school n years later. It is called the <u>School Staying Ratio</u>.

(b) The probability that a student in grade i will graduate with a final education k is interpreted as the proportion of students now in grade i who will graduate with final educational qualification, k, n years later. It is called the <u>drop-out ratio</u>. The proportion of students now in grade i who drop out from school with final educational qualification k within n years is called the <u>school completion ratio</u> and is an important factor in manpower supply.

(c) The expected number of years left for students in grade i before graduating with any of the final educations is interpreted as the average length of time remaining for students now in grade i before graduation. It is called the <u>school survival time</u>.

(d) The expected number of school years left for any of the students in school now, before graduating with any of the final educations is interpreted as the <u>average length of time remaining for any student in school now before graduating</u>. It is called the <u>expected length of schooling</u>.

Generally probabilities are interpreted as proportions or fractions while expected values are interpreted as average values. The characteristics of the type cited above, together with the grade structures, once obtained may be used to estimate additional educational characteristics. These may include; cost of education up to completion; staffing and capital requirements; pupil performances etc. Some of these measures for the Kenyan primary education system are illustrated in Owino (1982).

Most of the characteristics of the type cited above are directly or indirectly a consequence of the systems promotion criteria. In order to avoid ending up with undesired characteristics in the education system, there is a need to apply some control strategy in the system. This is because, for example, a reasonable looking promotion criterion in an education system could lead to unreasonable resource allocation within the states of the system. More specifically, some desired future characteristics may be specified and one would then wish to find the transition process to be adapted so as to reach them. This type of problem is of great importance since its implementation at an appropriate time would stop a bad situation from getting out of hand.

1.2 BRIEF LITERATURE REVIEW

One of the early markovian models for education planning was proposed by Gani (1963) who used it to forecast enrolment and degrees awarded in Australian Universities. Since then many similar models have been discussed. Among the more substantial contributions, Thonstad's (1969) book on education planning makes an extensive use of stochastic models. Prior to the publication of this book, Thonstad (1967) had applied the markovian model to the Norwegian education system and reported some of his findings in an O.E.C.D. meeting. Other contributions include Uche (1978a, 1978b, 1980, 1982) who applies markovian model to the Nigerian education system. Applications of the Markov chain model to the Kenyan primary education system include Owino (1982), Khogali (1982), Odhiambo and Owino (1985), Odondo (1985), Odhiambo and Khogali (1986) and Owino and Philips (1989). In these applications, several educational characteristics such as drop-out ratios, school survival time, staying ratios, cost of education and many others were computed. More specifically Owino (1982) compares the education characteristics between sexes, for the Kenyan primary education system; Khogali (1982) considers a cohort type analysis of the system; Odondo (1985) compares the educational characteristics within provinces of the country. Recently Owino and Philips (1989) have looked at the problem of homogeneity of the system by comparing the educational characteristics before and after 1972. This study suggested that the Kenyan primary education has not stayed homogeneous over time.

Other substantial contributions of stochastic models to educational planning include Armitage, Philips and Davies (1970); Johnstone and Philip (1973); Johnstone (1974); Moore (1975); Balinsky and Reisman (1973); Barthomomew (1975) and Mark Blang (1967). These works generally apply the probabilistic approach to study the stocks and flows patterns of some hierachical organizations. In particular Marshall (1973) compares enrolments at a university in California using both the Markov and the cohort model approaches. See also Marshall (1975) and Marshall (1971).

Apart from trying to identify the long run implications of the present propensities of an educational system, there is interest in what should be done to alter the future outcomes of the system. This is the process of exercising some control in the system. Several authors have looked at this problem of control for various hierachical systems.

These include; Davies (1973) and Davies (1975) who look at the control problem based on grade structures in manpower systems. In the later study he considers the maintainability of grade structures in a graded system through recruitment control. Other applications on control include, Grinold and Stanford (1978) and Uche (1984). These works consider the control problem as applied to some graded manpower systems. Bartholomew (1982) gives a detailed consideration of the grade structural control problem in hierachical organizations in terms of the flow in the system.

It should be noted that the educational system is just one of the fields of study in which stochastic models have been successfully applied. In the neighbouring area of demography Keyfitz (1968) has provided a basic source of methodology for mathematical analysis of population. White's (1970) work on chains of opportunity is a detailed study of job mobility based on the idea of modelling the flow of vacancies through the system. Buying behaviour has been studied as a stochastic process see for example Chatfield and Goodhart (1970). Quantitative geographers have shown a growing interest in stochastic methods, much of it stimulated by the pioneering work of Hugerstrand (1967) on the diffusion of innovations. The use of stochastic models in studying spread of epidemics and rumours is also important; Becker (1968), Daley and Kendall (1965) and so on.

The developments refered to above which have taken place in recent years have been almost overshadowed by the rise of interest in manpower planning. Most of the new work in this area has been published in volumes of conference proceedings including Wilson (1969), Smith (1971) and Bartholomew and Smith (1971); see also Bartholomew (1982), Bartholomew (1973), Bartholomew and Forbes (1979), Sales (1971), Glen (1977), Vassoliou (1976) and Butler (1971).

Another important area of application of stochastic models is ecology and pest control; see for example Pielou (1976).

1.3 OBJECTIVE OF THE STUDY

The aim of the present study is to assess the long run implications of the present educational propensities. For example what is the long run implications of the present educational properties? What proportion of today's students will finally emerge from the present lot with a given final educational qualification? And so on. All these

implications are important in a country's manpower and economic planning. In many educational planning exercises, for example, forecasting of enrolments, graduations, and so on, quite a number of ratios are used. For instance it is assumed that a given proportion will proceed to another school category and so on. Those ratios are obviously affected by a number of factors, like, the capacities of the schools, admission policies, availability of funds for running the schools etc. [see for example; the General discussions on projections of student numbers in higher education in the J.R. Statist. Soc. Soc.A (1985); Uche (1978); and Le'gare (1972)].
It is thus important to study the expected future characteristics in relation to some of these factors. More specifically the objective of this thesis is to develop statistical models, based on the theory of absorbing Markov chains, for studying:

(i) Enrolment patterns

(ii) The school staying ratios and drop-out rates

(iii) The school survival times

(iv) The expected length of schooling by any pupil in the system

(v) The possible plan of action associated with controlling future system outcomes.

The study is divided into five broad headings as follows:

(a) The homogeneity of the educational transition process.

(b) Models based on time dependent Markov chains.

(c) Models based on a generalised cohort analysis.

(d) Models for estimating and controlling academic survival.

(e) The problem of attainability and maintainability of characteristics of an education system.

1.4 BRIEF OUTLINE OF WORK DONE IN THIS THESIS

In this thesis we have used the Markov chain model to study various characteristics of an educational system. Most of the studies cited in section 1.2 assume that the transitional probabilities are constant, which is convenient for purposes of analysis. In this study we have adjusted this assumption and considered situations where the probabilities change with time. An attempt has been made to study the functional relationship between transition rates and factors such as admission policies, availability of scholarships, social and political factors. All these factors are bound to change with time.

To justify the above considerations, the study begins by examining the homogeneity of the Kenyan primary education system via the educational characteristics, of the types mentioned earlier, over various equal time periods. This enables us to apply the theory of homogeneous Markov chains to study the flow process in each of these periods, by computing the corresponding characteristics. Any appreciable difference in these characteristics suggests departure from homogeneity of the process, over the entire period of consideration. The Kenyan primary education system is found to be inhomogeneous over time. Furthermore, a chi-square test based on flow values shows that the transition process is inhomogeneous. This suggests therefore that the system changes with time, which leads to the study of the time dependent process.

In studying the time dependent models we have considered the transition probabilities $p_{ij}(t)$'s as probabilities of occurrence of a random process with a corresponding distribution function, F(t), defined on the time domain. We have suggested some probability transition models, which are found to be useful in describing the transition process on the basis of a goodness of fit test. These probability transition models are then used to obtain the educational characteristics using the theory of the inhomogeneous Markov chains.

A further consideration of the system is made via the generalised cohort analysis to study the flow of a particular cohort of students through the education system. For the purpose of this study, the term cohort is used to denote a group of students regardless of age or socio-economic background who enter the first grade in the same academic year. The applications of the generalised cohort model are given on the basis of the time dependent probability transition models.

The study then considers the inhomogeneous transition process in terms of the effects of quantifiable factors which may change in time. Measures of academic retentions are obtained under this consideration. Using the multivariate control theory as described in Press (1982), an attempt is made to control some of the factors so as to attain some desired future educational characteristics. The main advantage of this variable dependent approach is that it takes into account factors which affect the system's transitional process. This implies that, which ever assumption on the transition process we opt for in future, the procedure gives a plan of action to be taken on the controllable factors so as to attain the goal.

The other type of problem investigated in this thesis is that of attainability and maintainability of desired educational characteristics. In this case, desired characteristics are specified and the problem is to obtain the transition process to be followed so as to achieve them. A condition for maintainability for some characteristics of the process is obtained in the form of a system of linear equations dependent on the system growth rate. Following the identification of the maintainable characteristics, an investigation is made into the methods of solution to the problem of attainability of the characteristics.

Each of the models introduced in this thesis has been illustrated by computing numerical values of the characteristics derived under the models.

CHAPTER II

THE PROBLEM OF HOMOGENEITY IN AN EDUCATION SYSTEM

2.1 INTRODUCTION

In earlier studies on the Kenyan primary education system it was assumed that the transition process remained homogeneous over the entire time period, extending from 1964 to 1980. The changes that have taken place over this period include;

(i) making primary education available to all,

(ii) abolishing building funds,

(iii) formation of Parents Teachers Associations (PTAs) to help run schools,

and many others. These changes justify the need to check for any variations in the average education characteristics over various time intervals.

In this chapter we study the average flow rates for the periods 1964-1969, 1969-1974 and 1975-1980. This is done in order to examine the homogeneity of the transition process over the entire time period. It is assumed that within each of the three time periods, the transition process is homogeneous. This enables us to apply the theory of homogeneous Markov chains to study the flow process in each period. The homogeneity of the process over the entire period is examined in terms of the following educational characteristics:

(i) the school retention rates,

(ii) drop out rates,

(iii) completion rates

(iv) survival times, and

(v) expected length of schooling.

Any appreciable difference in these characteristics between the time periods will suggest departure from homogeneity of the process over the entire period.

The data used for the analysis is based on the first seven grades of the process, that is grades one to seven.

2.2 THE MODEL

The states of the education system are denoted by 1,2,...,N; where N is the number of possible states. Let the transition matrix be denoted by $P = \left(\left(p_{ij}(t)\right)\right)$ where $p_{ij}(t)$ is the probability of moving from state i to state j in the time interval (t,t+1). Suppose that observations are available over T time periods denoted by t = 1,2,...,T; representing the periods between 1964 and 1980. Suppose that the time period can be further partitioned into ℓ time intervals so that the u-th time interval is of size τ_u, allowing for overlapping if necessary. We can then write the ℓ time intervals as;

$t = T_1+1,\ T_1+2, \ldots, T_1+\tau_1;\ t = T_2+1, T_2+2, \ldots,$
$T_2+\tau_2;\ \ldots;\ t = T_\ell+1, T_\ell+2, \ldots, T_\ell+\tau_\ell$; where
$T_1+1 = 1$ and $T_\ell+\tau_\ell = T$.

The u-th time interval is then

$$t = T_u+1, T_u+2, \ldots, T_u+\tau_u \quad \text{for} \quad u=1,2,\ldots,\ell.$$

If $\tau_1 = \tau_2 = \ldots = \tau_\ell$, then the intervals are of equal size.

Suppose that the transition process is homogeneous over the u-th interval. The transition matrix for this interval is then given by

$$P_u = \left(\left({}_u p_{ij}\right)\right),\ u = 1,2,\ldots,\ell \qquad (2.1)$$

The maximum likelihood estimates of the u-th transition probabilities are given by

$${}_u\hat{p}_{ij} = \sum_{t=T_u+1}^{T_u+\tau_u} n_{ij}(t) \Bigg/ \sum_{t=T_u+1}^{T_u+\tau_u} n_i(t) \qquad (2.2)$$

where, $n_{ij}(t)$ is the number of pupils who move from grade i to grade j in the time interval $(t,t+1)$ and $n_i(t)$ is the number of pupils in grade i at time t. That is, $n_{ij}(t)$ and $n_i(t)$ are respectively the flows and stocks at time t. These estimates will be used to compute the educational characteristics mentioned above, for each time interval u.

Suppose the education system consists of r absorbing and s non-absorbing states, where $r + s = N$. Then the transition matrices for the u-th time period can be expressed in the following canonical form:

$$P_u = \begin{bmatrix} I & 0 \\ G_u & Q_u \end{bmatrix} \qquad (2.3)$$

for $u = 1,2,\ldots,\ell$. In this canonical representation, I is an $r \times r$ identity matrix of transitions between absorbing states; 0 is an $r \times s$ matrix of zeros giving transitions from absorbing to non-absorbing states; $G_u = (({}_u g_{ik}))$ is an $s \times r$ matrix which gives the transitions from non-absorbing to absorbing states and $Q_u = (({}_u q_{ij}))$ is an $s \times s$ matrix giving the transitions between non-absorbing states.

The n-step transition matrices according to each of the time periods $u = 1,2,\ldots,\ell$, are

$$P_u^{(n)} = (({}_u p_{ij}^{(n)})), \quad i,j = 1,2,\ldots,N. \qquad (2.4)$$

Assuming a one-step markov process, we have,

$$P_u^{(n)} = P_u^n$$

$$= \begin{bmatrix} I & 0 \\ G_u & Q \end{bmatrix}^n$$

$$= \begin{bmatrix} I & 0 \\ (I+Q_u+Q_u^2+\ldots+Q_u^{n-1})G_u & Q_u^n \end{bmatrix} \qquad (2.5)$$

The n-step transition matrix $P_u^{(n)}$ for the u-th period gives the n-step probabilities during this time interval.

2.3 RETENTION PROPERTIES OF THE EDUCATION SYSTEM WITHIN EACH OF THE TIME PERIODS.

In this section we obtain expressions for the school retention ratios, drop out and completion rates, the expected length of schooling and the school survival times, for each of the time periods. We examine the homogeneity of the education system by comparing the flow rates for the three time periods: 1964-1969, 1969-1974 and 1975-1980. This entails comparing the above educational characteristics for these periods.

School Retention Rates

During the u-th time period, the probability that a student in grade i will be in grade j, n years later, is the (i,j)-th entry of the matrix Q_u^n; that is ${}_uq_{ij}^{(n)}$. The probability that the student will be in any of the s school grades is the sum of the elements of the i-th row of Q_u^n, which we denote by ${}_uq_i^{(n)}$. That is

$$ {}_uq_i^{(n)} = \sum_{j=1}^{s} {}_uq_{ij}^{(n)} \quad ; \; i=1,2,\ldots,s; \; n=0,1,2.. \tag{2.6} $$

This is the i-th entry of the column vector $Q_u^n\underline{j}$, where $\underline{j}$ is an sx1 column vector of ones. It is called the <u>school retention rate</u>.

Tables 1(a)- 1(g) give the school retention rates for the Kenyan primary education system for the periods 1964 - 1969, 1969 - 1974 and 1975 - 1980. Each entry consists of three values corresponding to the three respective time periods.

Table 1(a): Fraction of pupils in grade 1 who will be in grade j, n years later and the retention rates.

grade j / n-years	1	2	3	4	5	6	7	retention rates
	.0336	.8824						.9160
1	.0323	.8164						.8487
	.0840	.7413						.8253
	.0011	.0618	.8219					.8848
2	.0010	.0589	.7594					.8193
	.0071	.1251	.6136					.7458
	.0000	.0033	.0880	.7597				.8510
3	.0000	.0032	.0876	.6936				.7844
	.0006	.0158	.1565	.5155				.6884
		.0002	.0063	.1110	.6658			.7833
4		.0001	.0068	.1118	.6100			.7287
		.0018	.0266	.1790	.4271			.6345
		.0000	.0004	.0101	.1199	.5999		.7303
5		.0000	.0004	.0113	.1276	.5455		.6848
		.0002	.0038	.0388	.1901	.3653		.5982
			.0000	.0007	.0130	.1395	.5358	.6890
6			.0000	.0009	.0161	.1518	.4935	.6623
			.0005	.0067	.0508	.2079	.2855	.5514
			.0000	.0005	.0011	.0190	.2077	.2283
7			.0000	.0001	.0016	.0248	.2141	.2406
			.0001	.0010	.0106	.0692	.2029	.2838
				.0000	.0001	.0020	.0492	.0513
8				.0000	.0001	.0031	.0558	.0590
				.0001	.0019	.0176	.0828	.1024
					.0000	.0002	.0094	.0096
9					.0000	.0003	.0151	.0154
					.0003	.0038	.0255	.0296
						.0000	.0016	.0016
10						.0000	.0021	.0021
						.0007	.0066	.0073

Table 1(b): Fraction of pupils in grade 2 who will be in grade j, n years later and the retention rates.

grade j / n-years	2	3	4	5	6	7	retention rates
	.0364	.9315					.9679
1	.0399	.9302					.9701
	.0847	.8277					.9124
	.0013	.0683	.8609				.9305
2	.0016	.0773	.8496				.9285
	.0072	.1415	.6955				.8442
	.0000	.0038	.0968	.7545			.8551
3	.0001	.0048	.1095	.7472			.8616
	.0006	.0181	.1830	.5761			.7778
	.0000	.0002	.0073	.1105	.6798		.7978
4	.0000	.0003	.0094	.1322	.6682		.8101
	.0001	.0021	.0321	.2080	.4929		.7352
		.0000	.0005	.0101	.1352	.6072	.7530
5		.0000	.0007	.0146	.1643	.6046	.7842
		.0002	.0047	.0470	.2390	.3851	.6760
			.0000	.0007	.0162	.2149	.2318
6			.0000	.0013	.0244	.2427	.2684
			.0006	.0085	.0698	.2413	.3202
			.0000	.0000	.0015	.0478	.0493
7			.0000	.0001	.0028	.0599	.0628
			.0001	.0013	.0159	.0887	.1060
				.0000	.0001	.0088	.0089
8				.0000	.0003	.0119	.0122
				.0002	.0031	.0250	.0283
					.0000	.0015	.0015
9					.0000	.0021	.0021
					.0006	.0060	.0066
					.0000	.0002	.0002
10					.0000	.0004	.0004
					.0001	.0013	.0014

Table 1(c): Fraction of pupils in grade 3 who will be in grade j, n years later and the retention rates.

n-years \ grade j	3	4	5	6	7	retention rates
1	.0369 .0432 .0863	.9243 .9133 .8403				.9612 .9565 .9266
2	.0014 .0019 .0074	.0702 .0813 .1499	.8100 .8032 .6960			.8816 .8864 .8533
3	.0001 .0001 .0006	.0040 .0054 .0201	.0891 .1101 .1924	.7298 .7183 .5955		.8230 .8339 .8086
4	.0000 .0000 .0001	.0002 .0003 .0024	.0065 .0101 .0355	.1186 .1480 .2383	.6519 .6499 .4653	.7772 .8083 .7416
5		.0000 .0000 .0003	.0004 .0008 .0055	.0121 .0192 .0599	.2070 .2350 .2521	.2195 .2550 .3178
6			.0000 .0001 .0008	.0010 .0020 .0121	.0429 .0540 .0825	.0439 .0561 .0954
7			.0000 .0000 .0001	.0001 .0002 .0021	.0075 .0102 .0211	.0076 .0104 .0233
8				.0000 .0000 .0003	.0012 .0018 .0047	.0012 .0018 .0050
9				.0000 .0000 .0001	.0002 .0003 .0009	.0002 .0003 .0010
10					.0000 .0000 .0002	.0000 .0000 .0002

Table 1(d): Fraction of pupils in grade 4, who will be in grade j, n years later and the retention rates.

grade j / n-years	4	5	6	7	retention rates
1	.0391 .0458 .0921	.8764 .8794 .8284			.9155 .9252 .9205
2	.0015 .0021 .0085	.0640 .0825 .1575	.7896 .7865 .7087		.8551 .8711 .8747
3	.0001 .0001 .0008	.0035 .0054 .0225	.0991 .1101 .2225	.7053 .7183 .5537	.8080 .8339 .7995
4	.0000 .0000 .0001	.0002 .0004 .0028	.0084 .0140 .0468	.1979 .2266 .2523	.2065 .2410 .3020
5		.0000 .0000 .0003	.0006 .0013 .0082	.0381 .0480 .0723	.0387 .0493 .0808
6			.0000 .0001 .0013	.0064 .0086 .0167	.0064 .0087 .0180
7			.0000 .0000 .0001	.0010 .0014 .0034	.0010 .0014 .0035
8				.0002 .0002 .0006	..0002 .0002 .0006
9				.0000 .0000 .0001	.0000 .0000 .0001

Table 1(e): Fraction of pupils in grade 5 who will be in grade j, n years later and the retention rates.

grade j / n-years	5	6	7	retention rates
1	.0340 .0480 .0980	.9010 .8943 .8555		.9350 .9423 .9535
2	.0012 .0023 .0096	.0779 .1047 .1898	.8048 .8091 .6684	.8839 .9161 .8678
3	.0000 .0001 .0009	.0051 .0093 .0317	.1943 .2206 .2430	.1994 .2300 .2756
4	.0000 .0000 .0001	.0003 .0007 .0047	.0347 .0427 .0592	.0350 .0434 .0640
5		.0000 .0001 .0007	.0057 .0073 .0121	.0057 .0074 .0128
6		.0000 .0000 .0001	.0009 .0012 .0022	.0009 .0012 .0023
7			.0001 .0002 .0004	.0001 .0002 .0004
8			.0000 .0000 .0001	.0000 .0000 .0001

Table 1(f): Fraction of pupils in grade 6 who will be in grade j, n years later and the retention rates.

grade j / n-years	6	7	retention rates
1	.0525 .0691 .1239	.8932 .9048 .7814	.9457 .9739 .9053
2	.0028 .0048 .0153	.1854 .2032 .2075	.1882 .2080 .2228
3	.0001 .0003 .0019	.0312 .0359 .0414	.0313 .0362 .0433
4	.0000 .0000 .0002	.0050 .0059 .0073	.0050 .0059 .0075
5		.0008 .0009 .0012	.0008 .0009 .0012
6		.0001 .0001 .0002	.0001 .0001 .0002

Table 1(g). Fraction of pupils in grade 7 who will be in grade j, n years later and the retention rates.

grade j / n-years	7	retention rates
1	.1550 .1555 .1416	.1550 .1555 .1416
2	.0240 .0242 .0201	.0240 .0242 .0201
3	.0037 .0038 .0028	.0037 .0038 .0028
4	.0006 .0006 .0004	.0006 .0006 .0004
5	.0001 .0001 .0001	.0001 .0001 .0001

Comments on Tables 1(a) - 1(g)

The average promotion rates for all the grades before 1969 was 0.8934; between 1969 and 1974 it was 0.8833 and after 1974 it was 0.8190. This indicates a drop in the average promotion rate of about 8%. The corresponding average repeat rates increased from 0.0554 to 0.0620 and then to 0.0895 over the periods. This is an overall increase in repeat rates of about 62%.

Before 1969 the proportion of grade one pupils still in school after the first year was 0.9160; between 1969 and 1974 this proportion was 0.8487 and after 1974 it was 0.8253. This indicates a high retention rate for the pre-1969 system during the first year of schooling. These proportions were 0.8848, 0.8193 and 0.7458 after two years of schooling for the three time periods respectively. After eleven years of schooling the proportions were respectively 0.0003, 0.0004 and 0.0016. This indicates that in the long run the post 1974 pupils stayed longer in school than their earlier counterparts. We can make similar deductions for pupils who were already in any of the other grades. For example, the proportion of grade seven pupils still in school after one year were respectively 0.1550, 0.1555 and 0.1416. After three years these proportions were respectively 0.0037, 0.0038 and 0.0028. This seems to reflect the fact that after 1974, grade seven pupils left the system faster than during the earlier periods.

School Drop-out and Completion Rates

The probability that a student entering grade i graduates n years later with final education k is given by

$$_u g_{ik}^{(n)} = \sum_{j=1}^{s} {_u q_{ij}^{(n-1)}} \; {_u g_{jk}}, \quad i=1,2,\ldots,s;\ k=1,2,..r \tag{2.7}$$

for the u-th time interval. This quantity is the (i,k)-th entry of the matrix product $Q_u^{n-1} G_u$, for n=1,2,..;. It is called the drop-out rate. Summing the left hand side of equation (2.7) from n=1 to n=w, we get the probability of graduating with final education k within w years. We denote this sum by $_u\bar{g}_{ik}^{(w)}$ and write

$$_u\bar{g}_{ik}^{(w)} = \sum_{n=1}^{w} {_u g_{ik}^{(n)}}, \quad i=1,2,..,s;\ k=1,2,\ldots,r \tag{2.8}$$

This is the (i,k)-th entry of the matrix sum,

$$\sum_{n=0}^{w-1} Q_u^n G_u ,$$

and is interpreted as the school completion rate.

It is an important factor in manpower planning. The sum to infinity

$$\sum_{n=1}^{\infty} {}_u g_{ik}^{(n)}$$

exists since for an absorbing Markov chain

$$\lim_{n\to\infty} Q_u^n = 0.$$

This means that the matrix series $\sum_{n=0}^{\infty} Q_u^n G_u$ converges. We call this infinite sum the <u>school absorbing rate</u> and denote it by ${}_u\bar{g}_{ik}$. That is

$${}_u\bar{g}_{ik} = \sum_{n=1}^{\infty} {}_u g_{ik}^{(n)} \qquad (2.9)$$

and so ${}_u\bar{g}_{ik}$ is simply the (i,k)-th entry of the matrix

$$\sum_{n=0}^{\infty} Q_u^n G_u .$$

Table 2 below gives the school absorbing rates for the Kenyan primary education system for the periods 1964-1969, 1969-1974 and 1975-1980 respectively.

Table 2. Fraction of pupils who drop out from grade j within x years.

x-years \ grade j	1	2	3	4	5	6	7
	.0840	.0321	.0388	.0845	.0650	.0543	.8450
1	.1513	.0299	.0435	.0748	.0577	.0261	.8445
	.1747	.0876	.0734	.0795	.0465	.0948	.8584
	.1152	.0695	.1184	.1449	.1161	.8118	.9760
2	.1807	.0715	.1136	.1289	.0839	.7920	.9758
	.2542	.1558	.1467	.1253	.1322	.7772	.9799
	.1490	.1449	.1770	.1920	.8006	.9687	.9963
3	.2156	.1384	.1661	.1661	.7700	.9638	.9962
	.3116	.2222	.1914	.2005	.7244	.9567	.9972
	.2167	.2022	.2228	.7935	.9650	.9950	.9994
4	.2713	.1891	.1917	.7590	.9566	.9941	.9994
	.3655	.2648	.2584	.6980	.9360	.9925	.9996
	.2697	.2470	.7805	.9613	.9943	.9992	.9999
5	.3152	.2158	.7450	.9507	.9926	.9991	.9999
	.4018	.3240	.6823	.9192	.9872	.9988	.9999
	.3110	.7682	.9561	.9936	.9991	.9999	1
6	.3377	.7316	.9439	.9913	.9988	.9999	1
	.4486	.6798	.9046	.9820	.9977	.9998	1
	.7717	.9507	.9924	.9990	.9999	1	1
7	.7594	.9372	.9896	.9986	.9998	1	1
	.7162	.8940	.9767	.9965	.9996	1	1
	.9487	.9911	.9988	.9998	1	1	1
8	.9410	.9878	.9982	.9998	1	1	1
	.8976	.9717	.9950	.9994	.9999	1	1
	.9904	.9985	.9998	1	1	1	1
9	.9846	.9979	.9997	1	1	1	1
	.9704	.9934	.9990	.9999	1	1	1
	.9984	.9998	1	1	1	1	1
10	.9979	.9996	1	1	1	1	1
	.9927	.9986	.9998	1	1	1	1

Comments on Table 2

From Table 2 we observe that, before 1969, the proportion of grade one pupils who dropped out from school after one year was 0.0840. This proportion was 0.1513 between 1969 and 1974 and it was 0.1747 after 1974. For a grade seven pupil, the proportions were 0.8450, 0.8445 and 0.8584, for the three respective periods, after one year. Similar values can be obtained for other years. For example, the proportions of grade one pupils who dropped out after eight years were 0.9487, 0.9410 and 0.8976 respectively. All the grade seven pupils would have dropped out of the system after six years.

The School Survival Time

Let ${}_uT_n$ be the number of years a student spends in grade j during the first n years after entering grade i, for each of the ℓ time intervals, $u=1,2,\ldots,\ell$. Then $E\left[{}_uT_n\right]$ is the expected length of time such a student spends in grade j, during the first n years of schooling. We shall denote this quantity by ${}_u\ell_{ij}^{(n)}$. Thus by definition of $E\left[{}_uT_n\right]$ we can write

$$ {}_u\ell_{ij}^{(n)} = E\left[{}_uT_n\right] $$

$$= \sum_{k=0}^{n} {}_u q_{ij}^{(k)} . \qquad (2.10)$$

Note that ${}_u\ell_{ij}^{(n)}$ is just the (i,j)-th element of the matrix sum $\sum_{h=0}^{n} Q_u^h$. Let

$${}_uL_n = (({}_u\ell_{ij}^{(n)})); \quad i,j=1,2,\ldots,s$$

Then we can write

$${}_uL_n = \sum_{h=0}^{n} Q_u^h .$$

The quantity ${}_u\ell_{ij}^{(n)}$ is called the expected length of stay in school grade j during the first n years of schooling, for a student entering school grade i. Summing the rows of ${}_uL_n$ we obtain the school survival times for students in grade i during the next n years of education.

The infinite sum

$$L_u = \sum_{h=0}^{\infty} Q_u^h , \qquad (2.11)$$

exists since $\lim_{h\to\infty} Q_u^h = 0$. The elements of L_u, which we denote by ${}_u\ell_{ij}$, give the expected lengths of stay in school grade j by those entering grade i. The sums of the rows of L_u give the expected survival times in school by those in grade i. Table 3(a) below gives the expected length of stay in grade j by those in grade i and the school survival times.

If we let ${}_u d_{ij}$ denote the probability that an entrant into grade i spends some time in grade j before leaving the system, then the expected length of stay in school grade j by those in grade i is given by

$$ {}_u \ell_{ij} = \left({}_u d_{ij}\right)\left({}_u \ell_{jj}\right) + \left(1 - {}_u d_{ij}\right) \times 0 \qquad (2.12) $$

from which we obtain,

$$ {}_u d_{ij} = {}_u \ell_{ij} \big/ {}_u \ell_{jj} \qquad (2.13) $$

Thus to obtain the probabilities of reaching grade j from grade i we need simply divide the elements of each column of L_u by the diagonal element of that column. Table 3(b) below gives these probabilities for each of the three time periods.

Table 3(a) Expected length of stay in grade j by pupils in grade i and the survival times.

grade j / grade i	1	2	3	4	5	6	7	Survival times
	1.0348	.9476	.9166	.8816	.7998	.7605	.8040	6.1449
1	1.0334	.8786	.8542	.8178	.7553	.7256	.7774	5.8423
	1.0918	.8842	.8009	.7413	.6808	.6646	.6051	5.4687
		1.0378	1.0038	.9655	.8759	.8329	.8804	5.5963
2		1.0415	1.0126	.9692	.8954	.8601	.9216	5.7004
		1.0926	.9897	.9160	.8411	.8214	.7476	5.4084
			1.0384	.9987	.9061	.8616	.9108	4.7156
3			1.0452	1.0004	.9242	.8878	.9512	4.8088
			1.0944	1.0129	.9302	.9083	.8268	4.7726
				1.0406	.9441	.8977	.9490	3.8314
4				1.0480	.9681	.9300	.9965	3.9426
				1.1015	1.0115	.9877	.8991	3.9998
					1.0352	.9843	1.0406	3.0601
5					1.0504	1.0091	1.0812	3.1407
					1.1086	1.0825	.9854	3.1765
						1.0554	1.1157	2.1711
6						1.0742	1.1509	2.2251
						1.1414	1.0390	2.1804
							1.1835	1.1835
7							1.1842	1.1842
							1.1650	1.1650

Table 3(b) The probability of reaching grade j from grade i.

grade j / grade i	1	2	3	4	5	6	7
	1	.9132	.8827	.8472	.7726	.7206	.6793
1	1	.8436	.8173	.7802	.7191	.6755	.6565
	1	.8093	.7318	.6730	.6141	.5824	.5194
		1	.9667	.9278	.8461	.7892	.7439
2		1	.9688	.9248	.8524	.8007	.7782
		1	.9043	.8316	.7588	.7196	.6418
			1	.9597	.8753	.8164	.7696
3			1	.9546	.8799	.8265	.8032
			1	.9196	.8391	.7958	.7097
				1	.9120	.8506	.8019
4				1	.9216	.8658	.8415
				1	.9124	.8653	.7718
					1	.9326	.8793
5					1	.9394	.9130
					1	.9484	.8458
						1	.9427
6						1	.9719
						1	.8918
							1
7							1
							1

Comments on Tables 3(a) and 3(b)

From Table 3(a) we notice that, a new entrant into school before 1969 took an average of 6.1449 years in primary school, while during the period 1969-1974 the average survival time in school was approximately 5.8423 years. After 1974 the average survival time was 5.4687 years. This seems to suggest that after 1974 pupils took a much shorter time in primary school than during the earlier periods. These observations can be made for pupils in the other grades. For example, for pupils in grade two the average survival times were 5.5963, 5.7004 and 5.4084 years respectively, for the three time periods. For a grade seven pupil the average survival times were respectively 1.1835, 1.1842 and 1.1650 years.

From table 3(b) we observe that before 1969 a new entrant in primary school had a probability of 0.6793 of reaching grade seven, while during the periods 1969-1974 and 1975-1980 the probabilities of a new entrant reaching grade seven were 0.6565 and 0.5194 respectively. For a pupil already in grade two, the probability of reaching grade seven was 0.7439 before 1969, it was 0.7782 between 1969-1974 and 0.6418 after 1974. Obviously for a grade seven pupil these probabilities were all equal to one, since they were already in grade seven.

Expected Length of Schooling

Let

$$\underline{p}(t) = \left(p_1(t),\ P_2(t), \ldots, P_N(t)\right)' \tag{2.14}$$

be the proportions of students enrolled in the various grades at time t. These proportions may be estimated from the stocks data by

$$\hat{p}_i(t) = n_i(t) \Big/ \sum_{i=1}^{N} n_i(t) \tag{2.15}$$

Writing absorbing states first, the vector in (2.14) may be partitioned as

$$\underline{p}'(t) = \left(\underline{u}'(t),\ \underline{q}'(t)\right) \tag{2.16}$$

where, $\underline{u}(t)$ is a vector of proportions in the final or absorbing grades and $\underline{q}(t)$ is the vector of proportions in the school grades, all quantities considered at time t.

The expected length of stay in the various school grades by any of these students during the next n years are the components of the vector

$$\underline{q}'(t)_u L_n.$$

The expected length of stay in school by any of the students during that period is therefore,

$$\underline{q}'(t)_u L_n \, \underline{j},$$

where $\underline{j}$ is an sx1 vector of ones. Taking limits as n becomes large, we have

$$\lim_{n\to\infty} \underline{q}'(t)_u L_n = \underline{q}'(t) L_u \tag{2.17}$$

Thus $\underline{q}'(t)L_u$ is a vector which gives the expected length of stay in the various grades for any of the students. It follows that

$$\underline{q}'(t) L_u \, \underline{j}$$

gives the <u>expected length of schooling</u> (E.L.S) in the school system by any of the students.
That is,

$$E.L.S = \underline{q}'(t) L_u \, \underline{j} \tag{2.18}$$

Table 4 below gives the length of schooling in grade j by any pupil and the expected length of schooling by any pupil for each of the three time periods.

Table 4. The expected length of schooling in grade j and the overall expected length of schooling by any student in the system.

Grade j	1	2	3	4	5	6	7	expected length of schooling
Expected length of schooling in grade j	.2388	.4096	.5410	.6518	.7117	.7968	.9482	4.2979
	.2385	.3943	.5292	.6389	.7123	.8065	.9700	4.2897
	.2520	.4050	.5195	.6199	.6983	.8116	.8430	4.1493

Comments on Table 4.

From Table 4 we observe that, for the period before 1969, any pupil considered at random was expected to take 0.2388 of a year in grade one. Between 1969-1974 and after 1974, any pupil considered at random was expected to take 0.2385 and 0.2520 of a year, respectively, in grade one. This time increases with the primary school grade so that any pupil considered at random is expected to take 0.9482, 0.9700 and 0.8430 of a year in grade seven, for the three respective time periods.

Any pupil considered at random for the period before 1969 was expected to take, on the average, 4.2979 years in primary school regardless of their initial grade. For the periods 1969-1974 and 1975-1980, the expected lengths of schooling were 4.2892 years and 4.1493 years respectively.

2.4 COMPARATIVE COMMENTS

From the results presented in tables 1-4 we are able to make the following conclusions. The period 1975-1980 had the least average promotion rate of approximately 82% as compared to the promotion rates of 1964-1969 and 1969-1974 which were approximately 89% and 88% respectively. On the other hand, the period 1975-1980 had the highest average repeat rate of about 9%. The earlier time periods 1964-1969 and 1969-1974 had average repeat rates of about 5% and 6% respectively. This indicates an appreciable difference in school repeat rates in that, the later time period had a higher repeat rate as compared to the earlier periods.

During the period 1964-1969 approximately 92% of grade one pupils were still in school after the first year of schooling. During the later periods 1969-1974 and 1975-1980, only approximately 85% and 82%, respectively, of the grade one pupils were still in school after the first year of schooling. This seems to suggest that more pupils in primary school were able to continue with primary education after their first year of schooling, before 1969 than after 1969. On the other hand we observe, for example that, after eight years of schooling 5% of the pupils enrolled in school in the period 1964-1969 were still in school. In the periods 1969-1974 and 1975-1980 approximately 6% and 10%, respectively, of the pupils were still in primary school after eight years of schooling. Infact, after eleven years of schooling, a few of the 1969-1974 and 1975-1980 pupils would still be found in school while all those of the period 1964-1969 would have dropped out of school. This seems to indicate that pupils of the 1975-1980 period were retained longer in primary school in the long run than those of the earlier periods. This may possibly be due to the high repeat rates during the period 1975-1980.

An entrant into primary school during the period 1964-1969 took an average of 6 years 2 months in primary school. During 1969-1974 and 1975-1980 an entrant into primary school took an average of 5 years 10 months and 5 years 6 months, respectively, in primary school. This implies that pupils in the time periods 1969-1974 and 1975-1980 spent, on average, a shorter time in primary school than those of the earlier period 1964-1969.

This may possibly be due to a higher demand for grade one places during the later years or due to a higher drop out from the system during the later years.

The remarks made in the preceding paragraph can also be made for any other grade in the school system. For example, a pupil already in grade five spent on average 3 years 2 months in primary school during all the three time periods.

During the period 1964-1969 a pupil entering the primary school system had a 68% chance of reaching the highest primary school grade, that is grade seven. In the periods 1969-1974 and 1975-1980 a new entrant into primary school had a 65% and a 52% chance respectively, of reaching grade seven. This may be because pupils stayed longer in school during the period 1964-1969 as compared to the later time periods 1969-1974 and 1975-1980. A pupil already in grade two had a 74% chance of reaching grade seven between 1964-1969. The chance of reaching grade seven for a pupil already in grade two were respectively 78% and 64% during the periods 1969-1974 and 1975-1980. Here we observe that during the period 1969-1974 a pupil already in grade two had the highest chance of reaching grade seven as compared to the other periods.

Infact for pupils already in grades two to six the chance of reaching grade seven was highest during the period 1969-1974 all through. Generally, for pupils in any of the primary school grades the chance of reaching the highest primary school grade was least during the period 1975-1980. This may be due to the high demand for primary school places during this period because of population expansion.

Any pupil considered at random, regardless of their grade, was expected to take on average 4 years 3 months in primary school during the time periods 1964-1969 and 1969-1974. During 1975-1980, any pupil considered at random was expected to take on average 4 years 2 months in primary school. Again this reflects the high demand for primary school places during the time period, 1975-1980, in that they took a shorter time in primary school during this period than during the earlier periods.

As portrayed by the results, the major changes that have taken place in the education system over the period, 1964 to 1980, are the following:

(i) The average repeat rates continued to rise as from the period 1964-1969 to 1969-1974 and were highest during the period 1975-1980.

(ii) A grade one pupil spent a comparatively longer time in primary school during the period 1964-1969 than during the later time periods, 1969-1974 and 1975-1980.

(iii) Generally, pupils in primary school during the period 1969-1974 had a higher chance of reaching the highest primary school grade than those in primary school during the periods 1964-1969 and 1975-1980.

(iv) On average, any pupil in primary school during the periods, 1964-1969 and 1969-1974, was expected to stay longer in primary school than any pupil in primary school during the period 1975-1980.

In view of the above remarks, it seems reasonable to conclude that the Kenyan primary education system did not stay homogeneous over the period from 1964 to 1980. Most of the changes in the educational characteristics may have been due to the effect of changes in education policies and other social and economic factors.

CHAPTER III

TIME DEPENDENT TRANSITION MODELS

3.1 INTRODUCTION

In chapter II it was found out that the Kenyan primary education system has not remained homogeneous over time. In order to study a system, one needs to take into account past behaviour of the system. This means that in studying the transition process of the Kenyan education system the changes in the system over time cannot be ignored.

The application of time homogeneous transition models, such as stationary Markov chain models, though simple to apply, would only be suitable for studying education systems which are fully developed and which have stablised. In the third world countries and especially in Africa, education patterns are constantly changing due to rapid population growth and other socio-economic factors. This trend calls for time dependent dynamic models which would incorporate important factors, which could be internal or external to the system over a period of time.

The purpose of the present chapter is to relax the assumption of constant transition probabilities and instead assume that they change with time according to some models. In section 3.3, a number of transition models are proposed. The models attempt to incorporate endegeneous factors in the system over a period of time by means of probability distribution functions.

These models are then used together with the theory of time-dependent Markov chains, to compute various measures of academic retention. The measures include school staying ratios, drop out and completion ratios and expected length of schooling. Estimates for these measures are computed using the stocks and flows data of the primary school system in Kenya.

3.2 TIME DEPENDENT MARKOV MODEL

Introduction

Suppose the states of the education process are denoted by integers 1,2,..,N, where N is the number of possible states of the system. We consider time dependent Markov chain models, which are based on the assumption that the flows out of a given state are governed by time dependent probabilities. That is, the probability of an individual in state i at time t moving to state j at time t+1 depends only on i,j and t and not on any previous moves the individual may have made. If the probability of moving from state i to state j in the time interval (t,t+1) is denoted by $p_{ij}(t)$, then the transition matrix of flows is given by

$$P(t) = \left(\left(p_{ij}(t)\right)\right) \quad i,j=1,2,\ldots,N \tag{3.1}$$

and it depends on t. We will further assume that all individuals behave independently.

Let $n_{ij}(t)$ denote the number of individuals who were in state i at time t and are in state j at time t+1; these numbers represent the flows between states in the system. The number of individuals in state i at time t is then given by

$$n_i(t) = \sum_{j=1}^{N} n_{ij}(t) \tag{3.2}$$

and is called the stock in that state at time t.

We can treat the states individually because of the assumption of independent flows. If data on stocks and flows are available over τ time periods then, assuming the multinomial distribution, we have

$$L \propto \prod_{t=0}^{\tau-1} \prod_{j=1}^{N} p_{ij}(t)^{n_{ij}(t)} \quad , \tag{3.3}$$

where L is the likelihood function for the i-th state. Maximizing L with respect to $p_{ij}(t)$ subject to the restraint $\sum_{j=1}^{N} p_{ij}(t) = 1$, we obtain

$$\hat{p}_{ij}(t) = n_{ij}(t) / n_i(t) \tag{3.4}$$

as the maximum likelihood estimate of $p_{ij}(t)$. We notice that this estimate depends only on the stocks and flows data for the time interval $(t, t+1)$. However if we assume that $p_{ij}(t)$'s are constant over time then $p_{ij}(t) = p_{ij}$ and its estimate is now given by

$$\hat{p}_{ij}(t) = \sum_{t=0}^{\tau-1} n_{ij}(t) \Big/ \sum_{t=0}^{\tau-1} n_i(t) \tag{3.5}$$

In the special case, $\tau=1$, when data is available over a single time period, the two estimates coincide.

<u>Testing For Homogeneity</u>

Before one arrives at the conclusion that $p_{ij}(t)$'s change with time, one needs to test whether these probabilities are time homogeneous or not. The hypothesis to be tested may be formally stated as

$$Ho : p_{ij}(t) = p_{ij}, \text{ for all } i,j \text{ and } t.$$

Once again it is better to treat each i separately so that Ho is really s hypotheses, where s is the number of non-absorbing states or school grades.

Suppose that stocks and flows data are available for τ time periods, $0,1,2,\ldots,\tau-1$, say. Then the flow data for grade i can be arranged in a contigency table as indicated in Table 5 below.

Table 5: Flow data for grade i

Time Period	States 1	2	...	N	Row Totals
0	$n_{i1}(0)$	$n_{i2}(0)$	....	$n_{iN}(0)$	$n_i(0)$
1	$n_{i1}(1)$	$n_{i2}(1)$	....	$n_{iN}(1)$	$n_i(1)$
2	$n_{i1}(2)$	$n_{i2}(2)$	....	$n_{iN}(2)$	$n_i(2)$
.	.	.		.	.
.	.	.		.	.
.	.	.		.	.
.	.	.		.	.
$\tau-1$	$n_{i1}(\tau-1)$	$n_{i2}(\tau-1)$	....	$n_{iN}(\tau-1)$	$n_i(\tau-1)$
					$\sum_{t=0}^{\tau-1} n_i(t)$

Depending on the education process, some of the transition numbers $n_{ij}(t)$'s will be zero. For instance if we assume that promotion is only to the next higher grade and that there is no demotion, then the only possible non-zero transition numbers will be

$$n_{ii}(t),\ n_{i,i+1}(t) \quad \text{or} \quad n_{ik}(t), \quad k = s+1,\ s+2, .., N.$$

The hypothesis Ho requires the allocation of transition numbers to be according to the assumption of constant $p_{ij}(t)$'s over time. Thus, for each i, the expected transition numbers are given by

$$\hat{n}_{ij}(t) = E\left[n_{ij}(t) \,/\, Ho\right]$$

$$= n_i(t)\ \hat{p}_{ij}\ , \qquad \begin{array}{l} t = 0,1,\ldots,\tau-1 \\ j = 1,2,\ldots,N \end{array} \tag{3.6}$$

where $\hat{p}_{ij}$ is as given in (3.5).

For each i and for fixed j there are τ pairs of values $(n_{ij}(t), \hat{n}_{ij}(t))$, $t=0,1,\ldots,\tau-1$. To carry out the homogeneity test we must first of all divide both $n_{ij}(t)$'s and $\hat{n}_{ij}(t)$'s into k disjoint non-empty intervals. Let $O_{ij}(u)$ be the number of $n_{ij}(t)$'s which fall in the u-th interval and let $E_{ij}(u)$ be the corresponding $\hat{n}_{ij}(t)$'s which fall in the u-th interval. Then the chi-square statistic for testing Ho is given by

$$D = \sum_{u=1}^{k} \left(O_{ij}(u) - E_{ij}(u)\right)^2 \Big/ E_{ij}(u) \tag{3.7}$$

To test Ho we compute D and reject Ho at α level of significance if the P-value

$$Pr\left[\chi^2(k-1) > D_c\right] \leq \alpha$$

where D_c is the computed value of D. This test is performed for each i and j.

In order to conclude that the hypothesis of homogeneity is false it is enough to find that $p_{ij}(t) \neq \hat{p}_{ij}$, for at least one j, on the basis of the above test. This test was carried out for the Kenyan primary education system. The promotion flows were found to be homogeneous with P-values between

0.90 to 0.95. However for the repeat flows it was found that grades one and two are highly inhomogeneous with P-values between 0.01 and 0.001. The repeat flows for grades three to six were weakly homogeneous with P-values between 0.2 and 0.1. Grade seven repeat flows were again homogeneous with P-value about 0.4.

We therefore reject the suggestion of homogeneity of the primary education system on the basis of the non-homogeneity of repeat flows in grades one and two. Accordingly we suggest that the transition process changes with time. This leads to the study of the time dependent models discussed below.

Applications of the Time Dependent Markov Model.

Suppose that the education system consists of r absorbing and s non-absorbing states, such that r+s=N. Then the transition matrix for flows in the interval (t,t+1) can be put in the canonical form

$$P(t) = \begin{bmatrix} I & 0 \\ & \\ G(t) & Q(t) \end{bmatrix} \qquad (3.8)$$

where, I is an r x r identity matrix giving transition probabilities between absorbing states; 0 is an r x s matrix of zeros, giving transition probabilities from absorbing to non-absorbing states; $G(t) = ((g_{ik}(t)))$ is an s x r matrix, $g_{ik}(t)$ being the probability of a student in grade i at time t graduating with final education k at time t+1; and $Q(t) = ((q_{ij}(t)))$ is an s x s matrix of transitions between school grades.

The n-step transition matrix is defined by

$$P^{(n)}(t) = \left(\left(p_{ij}^{n}(t)\right)\right), \quad i,j=1,2,\ldots,N \qquad (3.9)$$

where $p_{ij}^{(n)}(t)$ is the probability that a student in state i at time t will be in state j at time t+n, that is n years later.

Assuming that the Markov property is satisfied, it can easily be shown that

$$P^{(n)}(t) = \prod_{h=0}^{n-1} P(t+h)$$

$$= \begin{bmatrix} I & 0 \\ \sum_{h=0}^{n-1} Q^{(h)}(t)G(t+h) & Q^{(n)}(t) \end{bmatrix}$$

(3.10)

where

$$Q^{(h)}(t) = \prod_{\ell=0}^{h-1} Q(t+\ell) \tag{3.11}$$

and

$$Q^{(0)}(t) = I \tag{3.12}$$

We now obtain expressions for the school retention rates, the drop out and completion ratios, the expected length of schooling and the school survival times.

School Retention Rates.

The probability that a student in school grade i at time t will be in school grade j, n years later is the (i,j)-th entry of $Q^{(n)}(t)$, that is $q_{ij}^{(n)}(t)$. The probability that a student in school grade i at time t will still be in any of the s school grades n years later is therefore given by the sum

$$q_i^{(n)}(t) = \sum_{j=1}^{s} q_{ij}^{(n)}(t) \quad , \quad \begin{array}{l} i=1,2,\ldots,s \\ n=0,1,2,\ldots \end{array} \tag{3.13}$$

It is the i-th entry of the column vector $Q^{(n)}(t)\ \underline{j}$, where $\underline{j}$ is an (sx1) column vector of ones. This probability is interpreted as the proportion of students in grade i who will be in any of the s school grades at time t+n, and is called the <u>school retention rate</u>.

School Drop-out and Completion Rates.

The probability that a student who enters school grade i at time t will graduate n years later with final education k is given by

$$g_{ik}^{(n)}(t) = \sum_{j=1}^{s} q_{ij}^{(n-1)}(t)\, g_{jk}(t+n-1), \qquad \begin{array}{l} i=1,2,\ldots,s \\ k=1,2,\ldots,r \end{array} \tag{3.14}$$

It is apparent that $g_{ik}^{(n)}(t)$ is the (i,k)-th entry of the matrix product

$$Q^{(n-1)}(t)\; G(t+n-1),\; n=1,2,\ldots$$

It is called the <u>drop out rate</u>. Summing the drop out rates for n=1 to n=w gives us the probability for a student entering grade i at time t to graduate with final education k within w years, which we denote by $\bar{g}_{ik}^{(w)}(t)$.

That is

$$\bar{g}_{ik}^{(w)}(t) = \sum_{n=1}^{w} g_{ik}^{(n)}(t)\;, \quad \begin{array}{l} i=1,2,\ldots,s \\ k=1,2,\ldots,r \end{array} \tag{3.15}$$

It is the (i,k)-th entry of the matrix sum

$$\sum_{h=0}^{w-1} Q^{(h)}(t)\; G(t+h).$$

It is called the <u>school completion rate</u> and is an important parameter in manpower planning. The probability that a student in grade i at time t will sooner or later graduate with final education k is therefore the infinite sum

$$\bar{g}_{ik}(t) = \sum_{n=1}^{\infty} g_{ik}^{(n)}(t) \tag{3.16}$$

It is the (i,k)-th entry of the matrix series

$$\sum_{h=0}^{\infty} Q^{(h)}(t)\ G(t+h)$$

which converges, since for an absorbing Markov chain

$$\lim_{h\to\infty} Q^{(h)}(t) = 0 \tag{3.17}$$

The quantity $\bar{g}_{ik}(t)$ is called the <u>school absorbing rate</u>.

School Survival Time

The probability that a student in grade i at time t will be in grade j in exactly k years is $q_{ij}^{(k)}(t)$. We note that $q_{ij}^{(k)}(t)$ is the (i,j)-th entry of the matrix $Q^{(k)}(t)$. Let T_n be the number of years a student spends in grade j during the next n years after enrolling in grade i at time t. Let $\ell_{ij}^{(n)}(t)$ denote the expected length of stay in school grade j by students in grade i at time t, during the next n years of schooling. Then

$$\ell_{ij}^{(n)}(t) = E(T_n)$$

$$= \sum_{k=0}^{n} q_{ij}^{(k)}(t) \qquad (3.18)$$

which is the (i,j)-th entry of the matrix series

$$L_n(t) = I + Q(t) + Q^{(2)}(t)+\ldots+Q^{(n)}(t). \qquad (3.19)$$

The expected length of stay in school by a student in grade i at time t, during the next n years, is the sum of the i-th row of the above series, that is $\sum_{j=1}^{s} \ell_{ij}^{(n)}(t)$.

The expected length of stay in school grade j by those in grade i at time t is given by

$$\ell_{ij}(t) = \lim_{n\to\infty} E[T_n]$$

$$= \sum_{k=0}^{\infty} q_{ij}^{(k)}(t) \qquad (3.20)$$

It is the (i,j)-th entry of the convergent matrix series

$$L(t) = \sum_{k=0}^{\infty} Q^{(k)}(t) \qquad (3.21)$$

The expected length of stay in school by a student entering grade i at time t, before graduating with any of the r final educations is the sum of the i-th row of the matrix L(t) given in (3.21). It is called the <u>school survival time</u>.

Suppose we let $d_{ij}(t)$ denote the probability that an entrant into grade i at time t spends some time in grade j before leaving the system, then the expected length of stay in school grade j by those in grade i at time t is given by

$$\ell_{ij}(t) = \big(d_{ij}(t)\big)\big(\ell_{jj}(t)\big) + \big(1 - d_{ij}(t)\big) \times 0, \tag{3.22}$$

from which we obtain

$$d_{ij}(t) = \ell_{ij}(t)/\ell_{jj}(t) \tag{3.23}$$

Thus to obtain the probabilities of reaching grade j after having joined grade i at time t, we simply need to divide the elements of each column of the matrix L(t) by the corresponding diagonal element.

Expected Length of Schooling.

The probability distribution of students in the various states at time t is given by

$$\underline{P}(t) = (P_1(t), P_2(t), \ldots, P_N(t))' \tag{3.24}$$

where $p_i(t)$ is the probability of a student being in grade i at time t, i=1,2,...,N. We can estimate $p_i(t)$ from the stocks data as

$$P_i(t) = n_i(t) / \sum_{i=1}^{N} n_i(t) .$$

If we write the absorbing states first, then the probability vector $\underline{P}(t)$ may be partitioned as

$$\underline{p}'(t) = (\underline{u}'(t), \underline{q}'(t)) \tag{3.25}$$

where $\underline{u}'(t)$ is the vector of proportions in the r absorbing states at time t; and $\underline{q}'(t)$ is the vector of proportions in the s non-absorbing states or school grades at time t.

The length of stay in school grade i by any student who is in school at time t during the next n years of schooling is therefore the j-th component of the vector

$$\underline{\ell}_n' = \underline{q}'(t)\, L_n(t), \tag{3.26}$$

where $L_n(t)$ is as in (3.19). The expected length of schooling by any student in school at time t, during the next n years of schooling is given by

$$\ell_n = \underline{q}'(t)\, L_n(t)\, \underline{j} \tag{3.27}$$

where $\underline{j}$ is an s x 1 column vector of ones.

The length of stay in school grade j by any student in school at time t is the j-th entry of the vector

$$\underline{\ell}' = \underline{q}'(t)\, L(t) \tag{3.28}$$

where L(t) is as in (3.21). The expected length of schooling by any student who is in school at time t is therefore given by

$$\ell = \underline{q}'(t)\, L(t)\, \underline{j}. \tag{3.29}$$

This is a good indicator of the retention property of the education system at any particular time t.

3.3 TRANSITION MODELS BASED ON PROBABILITY DISTRIBUTIONS.

In this section it is assumed that the transition probabilities may be regarded as probabilities of occurrence of random events, which characterise the transition process. In particular we assume that $p_{ij}(t)$ is the probability of occurrence of random events generated by a random process with a corresponding continuous distribution function, $F(t)$, defined on the time domain.

Assuming that $p_{ij}(t)$ is a continuous function of time which increases with t such that

$$\lim_{t\to\infty} p_{ij}(t) = 1 \quad \text{and} \quad \lim_{t\to-\infty} p_{ij}(t) = 0,$$

we may write

$$p_{ij}(t) = F(t), \ -\infty<t<\infty \tag{3.30}$$

Here negative values of t mean the times before we start observing the process, assuming that observation starts at t=0. If on the other hand we assume that $p_{ij}(t)$ is a decreasing function of t that satisfies

$$\lim_{t\to\infty} p_{ij}(t) = 0 \quad \text{and} \quad \lim_{t\to-\infty} p_{ij}(t) = 1$$

then we may write

$$p_{ij}(t) = 1 - F(t), \quad -\infty<t<\infty \tag{3.31}$$

More generally, suppose that $p_{ij}(t)$ increases to a maximum value 'a' where $0<a\leq 1$, as $t\to\infty$. If $p_{ij}(t)$ is a decreasing function of t then we consider $1 - p_{ij}(t)$ which is also an increasing function. Let us define the function

$$\rho_{ij}(t) = \begin{cases} p_{ij}(t)/a \ , & \text{if } p_{ij}(t) \text{ is increasing} \\ (1-p_{ij}(t))/a, & \text{if } p_{ij}(t) \text{ is decreasing} \end{cases}$$

Then

$$\lim \rho_{ij}(t) \to 1 \qquad \text{as} \quad t\to\infty$$

and

$$\lim \rho_{ij}(t) \to 0 \qquad \text{as} \quad t\to-\infty$$

We can then carry out the modelling process in terms of $\rho_{ij}(t)$'s. That is, we write

$$\rho_{ij}(t) = F(t), \quad -\infty<t<\infty \tag{3.32}$$

where F(t) is a probability distribution function. Note that $\rho_{ij}(t)$ is always an increasing function of t.

We shall now consider three transition models based on probability distribution functions. These are; the Normal transition model, based on the normal distribution; the Logistic transition model, based on the logistic distribution and the Exponential transition model, based on the exponential distribution.

<u>The Normal Probability Transition Model</u>.

Suppose that the education process is observed over a continuous time scale. We assume that the transition process is characterised by a random process whose distribution function over t is given by

$$F(t) = \frac{1}{\sigma_{ij}\sqrt{2\pi}} \int_{-\infty}^{t} \exp\{-(x-\mu_{ij})^2/2\sigma_{ij}^2\}dx, \quad -\infty<t<\infty$$

$$= \Phi\left(\frac{t - \mu_{ij}}{\sigma_{ij}}\right) \tag{3.33}$$

where $\Phi(\cdot)$ is the standardized normal distribution function. Since $\rho_{ij}(t)$ as defined in (3.32) always increases with time to a maximum value of one, we can write

$$\rho_{ij}(t) = \Phi\left(\frac{t - \mu_{ij}}{\sigma_{ij}}\right) \tag{3.34}$$

Let us make the probit transformation

$$y = \frac{t - \mu_{ij}}{\sigma_{ij}} + 5$$

$$= \alpha_{ij} + \beta_{ij}\, t$$

where,

$$\alpha_{ij} = 5 - \mu_{ij}/\sigma_{ij} \quad \text{and} \quad \beta_{ij} = 1/\sigma_{ij}$$

Then,

$$\rho_{ij}(t) = \Phi(\alpha_{ij} + \beta_{ij}t - 5). \tag{3.35}$$

Equation (3.35) defines a Normal probability model for calculating $\rho_{ij}(t)$ and hence $p_{ij}(t)$ over desired time intervals. To fit the model, the values of the parameters α_{ij} and β_{ij} can be estimated from the straight line regression equation

$$\text{Probit } \hat{\rho}_{ij}(t) = \alpha_{ij} + \beta_{ij}t \tag{3.36}$$

where $\hat{\rho}_{ij}(t)$ is the maximum likelihood estimate of $\rho_{ij}(t)$. It is obtained from the maximum likelihood estimate of $p_{ij}(t)$ using the invariance property of maximum likelihood estimators.

The rate of change of $\rho_{ij}(t)$ with respect to time is the density function of F(t), that is

$$\frac{d\ \rho_{ij}(t)}{dt} = F'(t)$$

$$= f(t), \text{ say.}$$

Thus the mean rate of change of $\rho_{ij}(t)$ is given by

$$E[f(t)] = \int_{-\infty}^{\infty} f(t)\ dF(t)$$

$$= 1/2\sigma_{ij}\sqrt{\pi}$$

$$= \beta_{ij}/2\sqrt{\pi} \tag{3.37}$$

and the quantity

$$\gamma_{ij} = 1/E[f(t)]$$

$$= 2\sqrt{\pi}/\beta_{ij} \tag{3.38}$$

is a rough measure of the average amount of time for the transition process to stabilize. This measure may be estimated by

$$\hat{\gamma}_{ij} = 2\sqrt{\pi}/\hat{\beta}_{ij} \tag{3.39}$$

where $\hat{\beta}_{ij}$ is the estimated value of β_{ij} obtained from (3.36).

The Logistic Probability Transition Model

Here we assume that the transition process can be described in terms of a random process with a continuous distribution given by

$$F(t) = 1/\left[1+\exp\{-(\alpha_{ij}+\beta_{ij}t)\}\right], \quad -\infty<t<\infty; \beta_{ij}>0 \tag{3.40}$$

where α_{ij} and β_{ij} are parameters of the process. Then we may write

$$\rho_{ij}(t) = F(t)$$

$$= \exp(\alpha_{ij}+\beta_{ij}t) / \{1+\exp(\alpha_{ij}+\beta_{ij}t)\} \tag{3.41}$$

From (3.41) it follows that

$$\rho_{ij}(t)/(1-\rho_{ij}(t)) = \exp(\alpha_{ij} + \beta_{ij}t)$$

which leads to the logit transformation

$$\log\{\rho_{ij}(t)/(1-\rho_{ij}(t))\} = \alpha_{ij} + \beta_{ij}t \qquad (3.42)$$

The parameters α_{ij} and β_{ij} may be estimated from the regression equation

$$\text{logit } \hat{\rho}_{ij} = \alpha_{ij} + \beta_{ij}t \qquad (3.43)$$

The density function of $F(t)$ is given by

$$f(t) = \beta_{ij} F(t)(1-F(t)) \qquad (3.44)$$

This is the rate of change of $\rho_{ij}(t)$ with respect to time. Hence the average rate of change of $\rho_{ij}(t)$ is given by

$$E[f(t)] = \beta_{ij} \int_0^1 F(t)(1-F(t))\, dF(t)$$

$$= \beta_{ij}/6 \qquad (3.45)$$

The average amount of time required for the transition process to stabilize is given by

$$\gamma_{ij} = 6/\beta_{ij} \qquad (3.46)$$

with an estimated value of

$$\hat{\gamma}_{ij} = 6/\hat{\beta}_{ij} \qquad (3.47)$$

where $\hat{\beta}_{ij}$ is an estimate of $\hat{\beta}_{ij}$ obtained from (3.43).

The Exponential Probability Transition Model

Here we shall assume that the transition probabilities are generated by a random process with distribution function

$$F(t) = 1 - \exp[-(\alpha_{ij} + \lambda_{ij}t)]; \quad t>0, \ \lambda_{ij}>0 \tag{3.48}$$

where λ_{ij} is the rate parameter of the process and $-\alpha_{ij}/\lambda_{ij}$ can be considered as the time when the process began evolving. Then we may write

$$\rho_{ij}(t) = 1 - \exp[-(\alpha_{ij} + \lambda_{ij}t)] \tag{3.49}$$

The parameters λ_{ij} and α_{ij} may be estimated as the gradient and intercept of the line

$$-\log(1 - \hat{\rho}_{ij}(t)) = \alpha_{ij} + \lambda_{ij}\, t \tag{3.50}$$

The density function of F(t) is given by

$$\begin{aligned} f(t) &= \lambda_{ij} \exp[-(\alpha_{ij} + \lambda_{ij}t)] \\ &= \lambda_{ij}(1 - F(t)). \end{aligned} \tag{3.51}$$

It is the rate of change of $\rho_{ij}(t)$ with respect to time. Hence, the mean rate of change of $\rho_{ij}(t)$ is given by

$$\begin{aligned} E[f(t)] &= \lambda_{ij} \int_0^1 (1 - F(t))dF(t) \\ &= \lambda_{ij}/2 \end{aligned} \tag{3.52}$$

and the average amount of time required for the transition process to stabilize is given by

$$\gamma_{ij} = 2/\lambda_{ij} \tag{3.53}$$

with a corresponding estimate,

$$\hat{\gamma}_{ij} = 2/\hat{\lambda}_{ij} \tag{3.54}$$

where $\hat{\lambda}_{ij}$ is an estimate of λ_{ij} obtained from (3.50).

3.4. RESULTS ON THE APPLICATION OF THE PROBABILITY MODELS.

In this section the results of the application of the models proposed in section 3.3 to the Kenyan primary school data are given. In order to assess how well the models fit the observed stocks and flows data, a goodness of fit test must be performed. If the goodness of fit test indicates that a model fits the stocks and flows data sufficiently well then, the model may be used to project the future transition rates.

It is useful to accompany every projected rate with its standard error, measured around the assumed model. Before using the models proposed in section 3.3 to project the transition rates and other measures of academic retention, it is necessary to explain how to perform the goodness of fit test and how to calculate the standard error of the estimates.

Testing for Goodness of Fit

Suppose that data is available over τ time periods, $t = 0, 1, 2, \ldots, \tau-1$, then the expected transition numbers in the time interval $(t, t+1)$ under the model assumptions is given by

$$n^*_{ij}(t) = E[n_{ij}(t)/\text{Model}]$$

$$= n_i(t)\ p^*_{ij}(t) \qquad (3.55)$$

for each i and j where $n_i(t)$ is the total stock in grade i and $p^*_{ij}(t)$ is the fitted value of $p_{ij}(t)$, corresponding to the time interval $(t, t+1)$. Then proceeding as in the case of testing for homogeneity, the chi-square statistic for testing the goodness of fit is

$$D = \sum_{u=1}^{k} \left(O_{ij}(u) - E^*_{ij}(u)\right)^2 / E^*_{ij}(u) \qquad (3.56)$$

where $O_{ij}(u)$ is the number of $n_{ij}(t)$'s which fall in the u-th interval and $E^*_{ij}(u)$ the number of $n^*_{ij}(t)$'s which fall in the corresponding u-th interval. This statistic has $(k-1-d)$ degrees of freedom, where d is the number of unknown parameters involved in the model.

Standard Error of Fitted Transition Rates

Consider a continuous function $g(x)$ defined for each value of a random variable X, with a finite mean μ and a finite variance. Assuming that the random function is repeatedly differentiable, it can be expanded in a Taylor series about the mean $x = \mu$ as follows:

$$g(x) = g(\mu) + g'(\mu)(x-\mu) + g''(\mu)(x-\mu)^2/2! + \ldots$$

A first approximation for the mean and variance of $g(x)$ is given by

$$\begin{aligned} E[g(X)] &\simeq E[g(\mu) + g'(\mu)(X-\mu)] \\ &= g(\mu) \end{aligned} \tag{3.57}$$

and

$$\begin{aligned} Var[g(X)] &\simeq Var[g(\mu) + g'(\mu)(X-\mu)] \\ &= [g'(\mu)]^2 \; Var\; X \end{aligned} \tag{3.58}$$

We will use the expressions (3.57) and (3.58) to calculate the standard errors of the estimates for each model.

RESULTS ON THE NORMAL PROBABILITY TRANSITION MODEL

Goodness of Fit

The goodness of fit test is carried out to assess the closeness of fit between the observed and estimated flow values for all the seven grades. This is done when the system is assumed to evolve according to the Normal probability transition model. It is found that the differences between the observed and fitted flow values are not significant with P-values ranging between 0.6 to 1.0. This therefore suggests that the Normal probability transition model may be used to study the Kenyan primary education system.

Standard Error of Estimates.

When we assume that the Normal transition model is correct the future transition probabilities can be obtained from

$$p^*_{ij}(t) = \Phi(\hat{\alpha}_{ij} + \hat{\beta}_{ij}t - 5) \tag{3.59}$$

using (3.35) with $\hat{\alpha}_{ij}$ and $\hat{\beta}_{ij}$ as the least square estimates of α_{ij} and β_{ij} obtained from (3.36). Let us define a random variable X by

$$X = \hat{\alpha}_{ij} + \hat{\beta}_{ij}t - 5 \tag{3.60}$$

Then

$$\begin{aligned} \mu &= E(X) \\ &= \alpha_{ij} + \beta_{ij}t-5 \end{aligned} \tag{3.61}$$

and

$$\text{Var } X = \text{Var}(\hat{\alpha}_{ij} + \hat{\beta}_{ij} t)$$

$$= \left(\sum_{t=0}^{\tau-1} t^2/\tau + t^2 - 2t\bar{t} \right)\sigma^2 \Big/ \sum_{t=0}^{\tau-1} (t - \bar{t})^2 \tag{3.62}$$

where

$$\bar{t} = (1/\tau) \sum_{t=0}^{\tau-1} t = (\tau-1)/2$$

and σ^2 is estimated as

$$\hat{\sigma}^2 = \sum_{t=0}^{\tau-1} (y_t - \hat{y}_t)^2 \Big/ (\tau-2) \tag{3.63}$$

with

$$y_t = \text{Probit } \hat{\rho}_{ij} \tag{3.64}$$

and $\hat{y}_t$ is the fitted value of y_t under the model.

Now

$$\rho^*_{ij}(t) = \Phi(X)$$

using (3.59) and (3.60). Thus

$$E[\rho^*_{ij}(t)] = E[\Phi(X)] \tag{3.65}$$

$$\simeq \Phi(\alpha_{ij} + \beta_{ij} t\text{-}5) \tag{3.65}$$

and

$$\text{Var}[\rho^*_{ij}(t)] = \text{Var}[\Phi(X)]$$

$$\simeq [\Phi'(\mu)]^2 \text{Var } X. \tag{3.66}$$

using (3.59) and (3.58).

Hence the standard error of $\rho^*_{ij}(t)$ is given by

$$\text{S.E. } \left(\rho^*_{ij}(t)\right) \simeq \Phi'(\hat{\mu})\ \sigma^*$$

$$= \exp(-\hat{\mu}^2/2)\sigma^*/\sqrt{2\pi} \qquad (3.67)$$

where

$$\hat{\mu} = \hat{\alpha}_{ij} + \hat{\beta}_{ij}\, t - 5$$

and

$$\sigma^* = \left[\Big(\sum_{t=0}^{\tau-1} t^2/\tau + t^2 - 2t\bar{t} \Big)\hat{\sigma}^2 \Big/ \sum_{t=0}^{\tau-1} (t - \bar{t})^2 \right]^{1/2} \qquad (3.68)$$

The standard error of the transition probabilities $p^*_{ij}(t)$ is simply obtained by multiplying (3.67) by the maximum value of $\rho_{ij}(t)$ namely 'a' which is mentioned earlier in section 3.3. Infact when a=1 the value remains the same. Table 6 below gives the percentage relative error of the transition ratios for the Kenyan primary education system under the Normal probability transition model.

TABLE 6: *The percentage standard error of the transition ratios under the Normal probability transition model*

transition ratio p_{ij}	time t: 0	1	2	3	4	5	6	7	8	9	10	...	16
p_{12}	2.2	2.5	2.9	3.3	3.8	4.2	4.8	5.3	5.9	6.6	7.3	...	12.5
p_{23}	1.2	1.4	1.6	1.9	2.2	2.5	2.8	3.2	3.6	4.0	4.5	...	8.2
p_{34}	0.8	1.0	1.1	1.3	1.4	1.6	1.8	2.0	2.2	2.5	2.7	...	4.6
p_{45}	1.2	1.4	1.6	1.8	1.9	2.1	2.3	2.5	2.8	3.0	3.2	...	4.9
p_{56}	0.8	0.8	0.9	1.0	1.1	1.3	1.4	1.5	1.6	1.8	2.0	...	2.9
p_{67}	2.0	2.2	2.6	2.9	3.3	3.7	4.1	4.6	5.1	5.6	6.2	...	10.5
p_{78}	0.7	0.8	0.9	0.9	1.0	1.0	1.1	1.2	1.2	1.3	1.4	...	1.7
p_{11}	21.9	23.7	25.2	26.5	27.6	29.2	30.4	31.6	32.8	33.8	35.0	...	39.6
p_{22}	19.5	21.1	22.3	23.6	25.0	26.2	27.4	28.6	30.2	30.8	31.8	...	37.0
p_{33}	17.0	18.0	19.0	20.4	21.5	22.0	22.9	24.5	25.4	26.3	27.2	...	31.5
p_{44}	14.3	15.4	16.2	17.1	18.0	18.9	19.8	20.7	21.5	22.3	23.0	...	26.7
p_{55}	10.0	10.8	11.4	12.1	12.8	13.4	14.3	14.2	15.0	15.8	15.9	...	17.9
p_{66}	9.4	10.0	10.7	11.4	12.0	12.6	13.0	13.6	14.1	14.6	15.0	...	17.2
p_{77}	4.0	4.4	4.7	5.1	5.5	5.9	6.2	6.7	7.0	7.4	7.9	...	10.3

Comments on Table 6

The relative standard error when using the Normal probability model to estimate $p_{12}(t)$ during the initial time period is 2.2%. It is around 1.2% for estimating $p_{23}(t)$ during the initial time period. These relative standard errors are given in the first column of Table 6. In fact the maximum relative standard error when using the Normal probability model to estimate $p_{ij}(t)$'s during the initial time is about 22%. When the Normal probability model is used to describe the transition process after one year, the maximum relative percentage error is about 24%. During the second year the maximum relative standard error is about 25% and so on. For example, when the Normal probability model is used to describe the transition process, during the fifth year, the maximum relative standard error in the transitions is about 29%. After 10 years, the maximum relative standard error in estimating the transitions is about 35%. We see that as is expected, the relative standard error in estimation increases with time. This suggests that the Normal probability model is useful for projecting transition rates especially for time periods near the initial time period.

Another observation from Table 6 is that the relative standard errors seem to be relatively higher for repeat rates than for the other transition rates. This could possibly be due to misreporting of repeat data in primary schools.

Measures of Academic Retention based on the Normal probability transition model.

Tables 7(a) - 7(g) give the projected school retention rates for the Kenyan primary education system under the Normal probability model during the projection period.

<u>Tables 7(a) - 7(g)</u>: The School retention rates under the Normal probability model.

<u>Table 7(a)</u>: Fraction of pupils in grade one who will be in grade j, n years later and the staying ratios.

grade j / n years	1	2	3	4	5	6	7	staying ratios
1	.0935	.7331						.8266
2	.0094	.1363	.6025					.7482
3	.0010	.0204	.1750	.4940				.6904
4	.0001	.0029	.0363	.2005	.3959			.6357
5		.0004	.0067	.0544	.2171	.3298		.6084
6		.0001	.0012	.0126	.0769	.2388	.2466	.5762
7			.0002	.0027	.0228	.1087	.2099	.3443
8				.0006	.0062	.0404	.1075	.1547

Table 7(b): Fraction of pupils in grade two who will be in grade j, n years later and the staying ratios.

grade j / n years	2	3	4	5	6	7	staying ratio
1	.0880	.8328					.9208
2	.0083	.1556	.6901				.8540
3	.0008	.0234	.2044	.5572			.7858
4	.0001	.0033	.0432	.2392	.4668		.7526
5		.0005	.0081	.0692	.2781	.3543	.7102
6		.0001	.0015	.0173	.1070	.2572	.3831
7			.0003	.0041	.0345	.1144	.1533
8				.0009	.0102	.0407	.0518

Table 7(c): Fraction of pupils in grade three who will be in grade j, n years later and the staying ratios.

grade j / n years	3	4	5	6	7	staying ratio
1	.0923	.8374				.9297
2	.0092	.1637	.6809			.8538
3	.0010	.0258	.2177	.5737		.8182
4	.0001	.0039	.0501	.2734	.4417	.7692
5		.0006	.0104	.0879	.2693	.3682
6		.0001	.0021	.0243	.1032	.1297
7			.0004	.0063	.0322	.0389
8			.0001	.0016	.0090	.0107

Table 7(d): Fraction of pupils in grade four who will be in grade j, n years later and the staying ratios.

grade j / n years	4	5	6	7	staying ratio
1	.0975	.8187			.9162
2	.0102	.1745	.6934		.8781
3	.0011	.0301	.2505	.5414	.8231
4	.0001	.0050	.0650	.2690	.3391
5		.0008	.0151	.0870	.1029
6		.0001	.0034	.0234	.0269
7			.0008	.0057	.0065
8			.0002	.0013	.0015

Table 7(e): Fraction of pupils in grade five who will be in grade j, n years later and the staying ratios.

grade j / n years	5	6	7	staying ratio
1	.1076	.8516		.9592
2	.0125	.2106	.6738	.8969
3	.0016	.0420	.2597	.3033
4	.0002	.0080	.0688	.0770
5		.0015	.0157	.0172
6		.0003	.0033	.0036
7		.0001	.0007	.0008
8			.0001	.0001

Table 7(f): Fraction of pupils in grade six who will be in grade j, n years later and the staying ratios.

grade j / n years	6	7	staying ratio
1	.1315	.8014	.9329
2	.0184	.2178	.2362
3	.0028	.0452	.0480
4	.0004	.0085	.0089
5	.0001	.0015	.0016
6		.0003	.0003

Table 7(g): Fraction of pupils in grade seven who will be in grade j, n years later and the staying ratios.

grade j / n years	7	staying ratio
1	.1429	.1429
2	.0203	.0203
3	.0029	.0029
4	.0004	.0004
5	.0001	.0001

Comments on Table 7(a) - (g)

If the Kenyan primary education system evolved according to the Normal probability model then 0.8266 of the new entrants into primary school would still be in school after one year. After two years of schooling 0.7482 of these pupils would still be in school. Similar figures can be obtained for the other years. For example, after eight years of schooling only 0.1547 of the grade one pupils would still be in school.

For pupils already in grade two, 0.9208 of the pupils would still be in primary school after one year. After two years of schooling 0.8540 of the grade two pupils would still be in school and so on. Infact after eight years of schooling only 0.0518 of the grade two pupils would still be in school. We can obtain similar proportions for pupils in any of the other school grades. For example, for a pupil enrolled in grade seven at our initial time, 0.1429 of them would still be in primary school after one year. After two years of schooling 0.0203 of these pupils would still be in school, and so on. In fact, after six years of schooling all the grade seven pupils would have left the primary school system according to the Normal probability model.

Table 8 below gives the school drop out rates for the Kenyan primary education system under the Normal probability transition model.

Table 8: The school drop-out rate within n years of schooling, under the Normal probability model.

grade j / n years	1	2	3	4	5	6	7
1	.1734	.0792	.0703	.0838	.0408	.0671	.8571
2	.2518	.1460	.1452	.1219	.1031	.7638	.9797
3	.3096	.2142	.1818	.1769	.6967	.9520	.9971
4	.3643	.2474	.2308	.6609	.9230	.9911	.9996
5	.3916	.2898	.6318	.8971	.9828	.9984	.9999
6	.4238	.6169	.8703	.9731	.9964	.9997	1
7	.6557	.8467	.9611	.9935	.9992	1	1
8	.8453	.9482	.9893	.9985	.9999	1	1

Comments on Table 8

According to the Normal probability model 0.1734 of grade one pupils enrolled at the initial time would have dropped out of the system after the first year of schooling. After two years of schooling 0.2518 of these grade one pupils would have dropped out of primary school and so on.

In fact during the seventh year the drop out from primary school is highest since after seven years 0.6557 will have left the system.

We can obtain similar rates for pupils enrolled in any of the other school grades, during the initial time. For example, 0.8571 of those in grade seven would have dropped out of the system after one year. After two years of schooling 0.9797 of these grade seven pupils would have left the primary school system. All the grade seven pupils would have dropped out of primary school after six years of schooling.

In Tables 9(a) - (b) we give the school survival times and the probabilities of promotions for the Kenyan primary education system, under the Normal probability transition model.

Table 9(a): The expected length of stay in grade j by pupils in grade i and the school survival times within eight years of schooling, under the Normal probability model.

grade i \ grade j	1	2	3	4	5	6	7	survival time
1	1.1041	.8932	.8219	.7643	.7127	.6772	.4564	5.4298
2		1.0972	1.0157	.9476	.8869	.8864	.7259	5.5597
3			1.1027	1.0315	.9616	.9657	.8464	4.9079
4				1.1089	1.0293	1.0282	.9265	4.0929
5					1.1219	1.1140	1.0220	3.2579
6						1.1532	1.0747	2.2279
7							1.1665	1.1665

Table 9(b): The probability of a pupil, enrolled in grade i at the initial time, reaching grade j within eight years of schooling, under the Normal probability model.

grade j / grade i	1	2	3	4	5	6	7
1	1	.8141	.7454	.6892	.6353	.5872	.3913
2		1	.9211	.8545	.7905	.7686	.6223
3			1	.9302	.8571	.8374	.7256
4				1	.9175	.8916	.7943
5					1	.9660	.8761
6						1	.9213
7							1

Comments on Tables 9(a)-(b)

If the education system is assumed to evolve according to the Normal probability model, a pupil who entered grade one at the initial time would be expected to take an average of 5.4298 years in primary school within the next eight years of schooling. A pupil in grade two at the initial time would be expected to take an average of 5.5597 years in primary school within the next eight years and so on. The time taken in primary school, by a pupil enrolled in grade seven during the initial time, is on average 1.1665 years [see Table 9(a)].

A pupil enrolled in grade one during the initial time had a 0.3913 chance of reaching the highest primary school grade within the next eight years of schooling. The probability of a pupil in grade two at the initial time, reaching grade seven was 0.6223, within the next eight years. Similar values can be obtained for pupils enrolled in any of the other school grades during the initial time. For example a grade six pupil had a 0.9213 chance of reaching grade seven within eight years of schooling [see Table 9(b)].

In Table 10 below, we give the expected length of schooling (E.L.S) by any pupil in the school system at the initial time during an eight-year period under the Normal probability transition model.

Table 10: The expected length of schooling in grade j and the overall expected length of schooling by any pupil in primary school at the initial time, during an eight-year period under the Normal probability model.

grade j	1	2	3	4	5	6	7	Expected length of schooling (ELS)
Length of stay in grade j	.2548	.4079	.5303	.6346	.7222	.8446	.8194	4.2138

Comments on Table 10

Under the Normal probability model, any pupil in the system at the initial time was expected to spend 0.2548 of a year in grade one during an eight-year period. The expected time spent by any of these pupils increases with grade level. For example, any pupil in the system at the initial time was expected to spend 0.4079 of a year in grade two during an eight-year period. On the other hand, any of these pupils was expected to take 0.8194 of a year in grade seven during an eight year period. We note here that the expected length of stay in grade six is 0.8446 which is higher than that for grade seven. This indicates that pupils seem to stay longer in grade six but after leaving this grade they take a shorter time in grade seven before leaving the system. Generally any pupil in primary school at the initial time is expected to spend approximately 4.2138 years in primary school during an eight-year period.

RESULTS ON THE LOGISTIC PROBABILITY TRANSITION MODEL.

Goodness of Fit

The goodness of fit test is carried out to assess the closeness of fit between the observed and estimated flow values for all the seven grades when the system is assumed to evolve according to the Logistic probability model. It is found that the differences between the observed and the fitted flow values are not significant, with P-values ranging between 0.6 to 1.0. This therefore suggests that the Logistic probability transition model may also be used to study the Kenyan primary education system.

Standard Error of Estimates

When we assume that the Logistic transition model is correct then the future transition probabilities can be obtained from

$$\rho^*_{ij}(t) = 1/\left[1 + \exp\{-(\hat{\alpha}_{ij} + \hat{\beta}_{ij}t)\}\right] \qquad (3.69)$$

using (3.40) with $\hat{\alpha}_{ij}$ and $\hat{\beta}_{ij}$ as the least squares estimates of α_{ij} and β_{ij} obtained from (3.43).
The mean of $\rho^*_{ij}(t)$ is now given by

$$E\left[\rho^*_{ij}(t)\right] \simeq 1/\left[1 + \exp\{-(\alpha_{ij} + \beta_{ij}t)\}\right] \qquad (3.70)$$

and its standard error is given by

$$S.E.(\rho^*_{ij}(t)) \simeq \sigma^*\exp(-\hat{\mu}) / (1 + \exp(-\hat{\mu}))^2 \qquad (3.71)$$

where

$$\hat{\mu} = \hat{\alpha}_{ij} + \hat{\beta}_{ij}\,t$$

and σ^* is obtained using (3.63) and (3.68) with

$$y_t = \text{logit } \hat{\rho}_{ij}(t)$$

Again the standard error of the transition probabilities is simply obtained by multiplying (3.71) by the maximum value of ρ_{ij} namely 'a'. In Table 11 below we give the relative standard errors of the transition probabilities of the Kenyan primary education system under the Logistic probability transition model.

TABLE 11: *The percentage standard error of the transition model.*

transition ratio p_{ij}	0	1	2	3	4
p_{12}	2.0	2.4	2.7	3.1	3.6
p_{23}	1.1	1.3	1.5	1.8	2.1
p_{34}	0.7	0.8	1.0	1.1	1.3
p_{45}	1.3	1.5	1.7	1.9	2.1
p_{56}	0.7	0.8	0.9	1.0	1.1
p_{67}	1.8	2.1	2.4	2.8	3.2
p_{78}	0.7	0.8	0.9	0.9	1.0
p_{11}	18.7	20.3	22.1	23.6	25.2
p_{22}	16.9	18.1	19.7	21.0	22.4
p_{33}	14.3	15.5	16.7	17.9	19.2
p_{44}	12.1	13.1	13.9	15.0	15.9
p_{55}	8.0	8.7	9.6	10.2	10.7
p_{66}	7.9	8.4	9.1	9.8	10.5
p_{77}	4.1	4.4	4.7	5.1	5.5

transition ratios under the Logistic probability

time t

5	6	7	8	9	10	...	16
4.1	4.7	5.3	5.9	6.6	7.4	...	13.5
2.4	2.7	3.1	3.6	4.0	4.6	...	9.0
1.4	1.6	1.8	2.0	2.3	2.5	...	4.6
2.3	2.5	2.7	2.9	3.2	3.4	...	5.3
1.2	1.3	1.4	1.6	1.7	1.8	...	2.9
3.6	4.0	4.5	5.1	5.6	6.3	...	11.2
1.0	1.1	1.2	1.2	1.3	1.3	...	1.7
26.9	28.5	30.1	31.8	33.3	34.9	...	43.2
23.7	25.3	26.8	28.2	29.6	30.9	...	38.6
20.3	21.6	22.7	24.0	25.1	26.3	...	32.8
17.1	18.1	19.1	20.1	21.1	22.1	...	27.4
11.5	12.3	12.8	13.4	14.1	14.8	...	18.1
11.1	11.7	12.3	13.1	13.6	14.2	...	17.4
5.9	6.2	6.6	7.0	7.4	7.8	...	10.2

Comments on Table 11.

When the Logistic probability transition model is used to describe the transition process, the relative standard error in estimating $p_{12}(t)$ during the initial time $(t = 0)$ is 2.0%. It is about 1.1% for estimating $p_{23}(t)$ during the initial time. These relative standard errors are given in the first column of Table 11. Infact the maximum relative standard error in estimating $p_{ij}(t)$'s, during the initial time, according to the Logistic probability model is about 19%. When the Logistic probability transition model is used to describe the transition process after one year, the maximum relative standard error is about 20%. During the second year, the maximum relative standard error is about 22% and so on. In the fifth year, for example, the maximum relative standard error is about 27%. After ten years, the maximum relative standard error is about 35%. Again as expected these relative standard errors of the estimates increase with time. This therefore suggests that the Logistic probability model is useful in projecting transition rates especially for time periods near the initial time.

Another observation from table 11 is that the relative standard errors again seem to be relatively higher for repeat rates than for the other transition rates. This is possibly due to misreporting of repeat data in primary schools.

Measures of Academic Retention based on the Logistic probability transition model.

Tables 12(a) - 12(g) below give the projected retention rates for the Kenyan primary education system under the Logistic probability transition model.

Tables 12(a) -12(g): The school retention rates under the Logistic probability model.

Table 12(a): Fraction of pupils in grade one who will be in grade j, n years later and the staying ratios.

grade j / n years	1	2	3	4	5	6	7	staying ratio
1	.1285	.6742						.8027
2	.0168	.1714	.5255					.7137
3	.0022	.0332	.1998	.4320				.6672
4	.0003	.0058	.0513	.2199	.3493			.6266
5		.0010	.0111	.0707	.2231	.2896		.5955
6		.0002	.0022	.0183	.0861	.2304	.2097	.5469
7			.0004	.0042	.0260	.1076	.1945	.3327
8			.0001	.0009	.0067	.0385	.1036	.1498
9				.0002	.0016	.0117	.0416	.0551
10					.0003	.0032	.0139	.0174
11					.0001	.0006	.0043	.0050
12						.0002	.0011	.0013

Table 12(b): Fraction of pupils in grade two who will be in grade j, n years later and the staying ratio.

grade j / n years	2	3	4	5	6	7	staying ratio
1	.1239	.7829					.9068
2	.0157	.1949	.6447				.8553
3	.0020	.0369	.2431	.5218			.8038
4	.0003	.0063	.0617	.2641	.4332		.7656
5		.0010	.0132	.0842	.2871	.3141	.6996
6		.0002	.0025	.0216	.1150	.2497	.3890
7			.0005	.0049	.0360	.1164	.1578
8			.0001	.0010	.0097	.0415	.0523
9				.0002	.0024	.0125	.0151
10					.0005	.0034	.0039
11					.0001	.0009	.0010
12						.0002	.0002

Table 12(c): Fraction of pupils in grade three who will be in grade j, n years later and the staying ratios.

grade j / n years	3	4	5	6	7	staying ratio
1	.1239	.8252				.9491
2	.0156	.2070	.6688			.8914
3	.0020	.0394	.2537	.5560		.8511
4	.0003	.0067	.0647	.2978	.4039	.7734
5		.0010	.0138	.1005	.2698	.3851
6		.0002	.0027	.0273	.1087	.1389
7			.0005	.0065	.0342	.0412
8			.0001	.0014	.0093	.0108
9				.0003	.0023	.0026
10				.0001	.0005	.0006
11					.0001	.0001

Table 12(d): Fraction of pupils in grade four who will be in grade j, n years later and the staying ratios.

n years \ grade j	4	5	6	7	staying ratio
1	.1255	.8117			.9372
2	.0160	.2055	.6761		.8976
3	.0021	.0394	.2766	.4924	.8105
4	.0003	.0068	.0761	.2667	.3499
5		.0011	.0176	.0908	.1095
6		.0002	.0037	.0248	.0287
7			.0007	.0060	.0067
8			.0001	.0013	.0014
9				.0003	.0003
10				.0001	.0001

Table 12(e): Fraction of pupils in grade five who will be in grade j, n years later and the staying ratios.

grade j / n years	5	6	7	staying ratio
1	.1266	.8347		.9613
2	.0162	.2353	.6100	.8615
3	.0021	.0502	.2531	.3054
4	.0003	.0096	.0703	.0802
5		.0017	.0164	.0181
6		.0003	.0034	.0037
7		.0001	.0007	.0008
8			.0001	.0001

Table 12(f): Fraction of pupils in grade six who will be in grade j, n years later and the staying ratios.

grade j / n years	6	7	staying ratio
1	.1545	.7343	.8888
2	.0240	.2115	.2355
3	.0038	.0458	.0496
4	.0006	.0089	.0095
5	.0001	.0016	.0017
6		.0003	.0003
7		.0001	.0001

Table 12(g): Fraction of pupils in grade seven who will be in grade j, n years later and the staying ratios.

grade j / n years	7	staying ratio
1	.1346	.1346
2	.0181	.0181
3	.0024	.0024
4	.0003	.0003
5	-	-

Comments on Tables 12(a) - 12(g)

If the Kenyan primary education system evolves according to the Logistic probability model, then 0.8027 of the new entrants into primary school would still be in primary school after one year of schooling. After two years of schooling 0.7137 of these pupils would still be in school. Similar figures can be obtained for the other years. For example, after eight years of schooling 0.1498 of the grade one pupils would still be in the system.

For pupils already in grade two, 0.9068 of the pupils would still be in primary school after one year. After two years, 0.8553 of the grade two pupils would still be in the system and so on. Infact after eight years of schooling 0.0523 of the grade two pupils would still be in school. We can obtain similar proportions for pupils in any of the other school grades. For example, for pupils enrolled in grade seven during the initial time, 0.1346 of these pupils would still be in primary school after one year. After two years of schooling 0.0181 of these pupils would still be in school and so on. After five years of schooling, all the grade seven pupils would have left the primary school system if it evolves according to the Logistic probability transition model.

In Table 13 below, we give the school drop-out rates for the Kenyan primary education system under the Logistic probability transition model.

Table 13: The school drop-out rates within n years, under the Logistic probability model.

grade j / n years	1	2	3	4	5	6	7
1	.1973	.0932	.0509	.0628	.0387	.1113	.8654
2	.2863	.1447	.1087	.1025	.1385	.7645	.9819
3	.3327	.1962	.1490	.1896	.6946	.9504	.9976
4	.3733	.2344	.2266	.6502	.9198	.9906	.9997
5	.4045	.3004	.6148	.8905	.9818	.9983	1
6	.4532	.6110	.8612	.9713	.9963	.9997	1
7	.6673	.8424	.9588	.9933	.9993	1	1
8	.8502	.9477	.9892	.9985	.9999	1	1
9	.9450	.9849	.9974	.9997	1	1	1
10	.9825	.9961	.9994	.9999	1	1	1
11	.9950	.9990	.9999	1	1	1	1
12	.9987	.9998	1	1	1	1	1
13	.9997	1	1	1	1	1	1
14	.9999	1	1	1	1	1	1
15	1	1	1	1	1	1	1

Comments on Table 13

When the Logistic probability model is used to describe the education process, 0.1973 of grade one pupils, enrolled at the initial time, would have dropped out of school after the first year of schooling. After two years of schooling 0.2863 of these students would have dropped out from primary school and so on. The rest of these proportions are obtained from the first column of Table 13. For example, the highest drop out from primary school for these grade one pupils is in the seventh year of schooling when 0.6673 would have dropped out of the system.

Of the pupils enrolled in grade two during the initial time, 0.0932 would have dropped out after one year if the system is described by the Logistic probability model. After two years of schooling 0.1447 of these grade two pupils would have dropped out of the system and so on. We can obtain similar proportions for pupils enrolled in any of the other school grades, during the initial time. For example, of pupils in grade seven during the initial time, 0.8654 would have left primary school after one year. After two years of schooling, 0.9819 of these grade seven pupils would have left the system. Infact all the students in primary school during the initial time would have left the system after fifteen years, according to the Logistic probability transition model.

Tables 14(a) - (b) below give the school survival times and the probabilities of promotions for the Kenyan primary education system, under the Logistic probability transition model.

Table 14(a): The expected length of stay in grade j by pupils in grade i and the school survival times, under the Logistic probability model.

grade i \ grade j	1	2	3	4	5	6	7	Survival time
1	1.1478	.8858	.7903	.7462	.6934	.6819	.5689	5.5143
2		1.1419	1.0223	.9657	.8978	.8841	.7386	5.6504
3			1.1417	1.0795	1.0042	.9899	.8288	5.0441
4				1.1438	1.0646	1.0508	.8823	4.1415
5					1.1451	1.1319	.9541	3.2311
6						1.1829	1.0024	2.1853
7							1.1555	1.1555

Table 14(b): The probability of a pupil, enrolled in grade i at the initial time, ever reaching grade j according to the Logistic probability model.

grade j / grade i	1	2	3	4	5	6	7
1	1	.7757	.6922	.6523	.6054	.5765	.4923
2		1	.8953	.8443	.7840	.7473	.6392
3			1	.9438	.8770	.8368	.7173
4				1	.9297	.8883	.7636
5					1	.9569	.8257
6						1	.8675
7							1

Comments on Tables 14(a)-(b)

If the education system evolves according to the logistic probability model, then a pupil in grade one at the initial time would be expected to take an average of 5.5143 years in primary school. A pupil enrolled in grade two at the initial time would be expected to take an average of 5.6504 years in primary school according to this model. We can obtain these figures for pupils enrolled in any of the other school grades. For example, a pupil in grade seven at this initial time would have taken approximately 1.1555 years in the system. [Refer to

Table 14(a)].

A pupil enrolled in grade one during the initial time had a 0.4923 chance of reaching the highest primary school grade according to the Logistic probability model. The probability of a pupil in grade two, at the initial time, reaching grade seven was 0.6392. Similar probabilities are obtainable for pupils of the other grades. For example, a grade six pupil would have had a 0.8675 chance of reaching grade seven, according to the Logistic probability transition model. [See Table 14(b)]

In Table 15 below, we give the expected length of schooling (ELS) by any pupil in school at the initial time when the system evolves according to the Logistic probability transition model.

Table 15 The expected length of schooling in grade j and the overall expected length of schooling by any pupil in primary school at the initial time according to the Logistic probability model.

grade j	1	2	3	4	5	6	7	Expected length of schooling (ELS)
Length of stay in grade j	.2649	.4144	.5297	.6449	.7328	.8569	.8226	4.2662

Comments on Table 15

When the system is assumed to evolve according to the Logistic probability model, any pupil in the system at the initial time was expected to spend 0.2649 of a year in grade one. Any of these pupils was expected to spend 0.4144 of a year in grade two and so on. These values increase with grade size to the highest value in grade six. That is, any of these pupils was expected to take 0.8569 of a year in grade six. We note again that for grade seven, the expected length of stay was 0.8226 which is lower than that for grade six. This indicates that pupils seem to stay longer in grade six but after leaving this grade they take a shorter time in grade seven before leaving the system. Generally any pupil in primary school at the initial time would be expected to spend approximately 4.2662 years in primary school.

RESULTS ON THE EXPONENTIAL PROBABILITY TRANSITION MODEL.

Goodness of Fit

The goodness of fit test is carried out to assess the closeness between the observed and fitted flow values for all the seven grades when the system is assumed to evolve according to the Exponential probability transition model. It is found that the differences between the observed and fitted flow values are not significant, with P-values ranging between 0.5 to 1.0. This therefore suggests that the Exponential probability transition model may be used to study the Kenyan primary education system.

Standard Error of Estimates

When we assume that the Exponential probability transition model is correct, then the future transition probabilities can be obtained from

$$p^{*}_{ij}(t) = 1 - \exp[-(\hat{\alpha}_{ij} + \hat{\lambda}_{ij} t)] \qquad (3.72)$$

using (3.49) with $\hat{\alpha}_{ij}$ and $\hat{\lambda}_{ij}$ as the least squares estimates of α_{ij} and λ_{ij} obtained from (3.50).

The mean of $\rho_{ij}^{*}(t)$ is now given by,

$$E\left[\rho_{ij}^{*}(t)\right] \simeq 1 - \exp\left[-(\alpha_{ij} + \lambda_{ij} t)\right] \qquad (3.73)$$

and its standard error is given by

$$S.E.\left(\rho_{ij}^{*}(t)\right) \simeq \sigma^{*}\exp\left[-(\hat{\alpha}_{ij} + \hat{\lambda}_{ij} t)\right] \qquad (3.74)$$

where σ^{*} is estimated using (3.63) and (3.68) with

$$y_t = -\log\left(1 - \hat{\rho}_{ij}(t)\right).$$

As before the standard error of the transition probabilities is simply obtained by multiplying (3.74) by the maximum value of ρ_{ij} namely 'a'. Table 16 below gives the standard errors of the transition probabilities of the Kenyan primary education system under the Exponential probability transition model.

Table 16: The percentage standard error of the transition ratios under the Exponential probability transition model.

t	time t											
transition ratio p_{ij}	0	1	2	3	4	5	6	7	8	9	10	... 16
p_{12}	1.9	2.2	2.7	3.0	3.4	3.9	4.4	5.1	5.8	6.5	7.4	... 15.4
p_{23}	1.1	1.2	1.5	1.7	2.0	2.3	2.6	3.0	3.5	4.0	4.5	... 9.8
p_{34}	0.7	0.9	0.9	1.1	1.2	1.4	1.6	1.8	2.0	2.3	2.5	... 5.0
p_{45}	1.3	1.5	1.7	1.8	2.0	2.2	2.5	2.7	2.9	3.2	3.5	... 5.5
p_{56}	0.7	0.8	0.9	1.0	1.1	1.2	1.3	1.4	1.5	1.7	1.8	... 3.0
p_{67}	1.8	2.0	2.4	2.7	3.1	3.5	3.9	4.4	5.0	5.6	6.3	... 12.0
p_{78}	0.7	0.8	0.9	0.9	1.0	1.0	1.1	1.2	1.2	1.3	1.3	... 1.7
p_{11}	18.1	19.7	21.3	22.8	24.3	26.2	27.8	29.3	31.0	32.9	34.5	... 44.8
p_{22}	15.9	17.2	18.6	19.9	21.6	22.9	24.5	25.8	27.3	28.9	30.3	... 39.4
p_{33}	13.6	14.7	15.8	17.7	18.2	19.6	20.7	22.0	23.2	24.4	25.7	... 33.4
p_{44}	11.3	12.3	13.2	14.2	15.2	16.2	17.2	18.4	19.5	20.5	21.5	... 27.9
p_{55}	7.5	7.9	8.5	9.3	10.1	10.8	11.4	12.1	12.7	13.5	14.2	... 18.4
p_{66}	7.1	7.9	8.4	9.0	9.6	10.3	11.0	11.7	12.3	13.0	13.7	... 17.7
p_{77}	4.1	4.4	4.8	5.1	5.5	5.8	6.2	6.6	7.0	7.4	7.7	... 10.1

Comments on Table 16

If the Exponential probability transition model is used to describe the transition process, the relative standard error in estimating $P_{12}(t)$ during the initial time is 1.9%. It is about 1.1% for estimating $P_{23}(t)$ during the initial time. These values can be seen in the first column of Table 16. Infact the maximum relative standard error in estimating the transitions during the initial time, according to the Exponential probability transition model, is about 18.1%. When the Exponential probability transition model is used to describe the transition process, after one year, the maximum relative standard error is around 20%. During the second year it is about 21% and so on. For example, when the Exponential probability model describes the transition process, the maximum relative standard error is about 26% during the fifth year. After ten years it is about 35%. Again we observe here that the relative standard errors increase with time as is expected. This suggests that the Exponential probability model is useful in projecting transition rates especially for time periods near the initial time.

Another observation from Table 16 is that the relative standard errors again seem to be relatively higher for the repeat rates than for the other transition rates. This may be due to misreporting of repeat rates in primary schools as mentioned earlier.

Measures of Academic Retention based on the Exponential probability transition model.

Tables 17(a) - 17(g) below give the projected school retention rates for the Kenyan primary education system under the Exponential probability transition model.

Tables 17(a) - 17(g): The school retention rates under the Exponential probability model.

Table 17(a): Fraction of pupils in grade one who will be in grade j,n years later and the staying ratios.

grade j / n years	1	2	3	4	5	6	7	staying ratio
1	.1306	.6714						.8020
2	.0172	.1738	.5220					.7130
3	.0023	.0341	.2013	.4292				.6669
4	.0003	.0060	.0521	.2209	.3470			.6263
5		.0010	.0113	.0715	.2235	.2877		.5950
6		.0001	.0022	.0186	.0868	.2304	.2082	.5463
7			.0004	.0043	.0263	.1081	.1943	.3334
8			.0001	.0008	.0069	.0388	.1040	.1506
9				.0002	.0016	.0118	.0419	.0555
10					.0004	.0032	.0141	.0177
11					.0001	.0007	.0043	.0051
12						.0001	.0011	.0012

Table 17(b): Fraction of pupils in grade two who will be in grade j, n years later and the staying ratios.

grade j / n years	2	3	4	5	6	7	staying ratio
1	.1272	.7797					.9069
2	.0164	.1975	.6420				.8559
3	.0021	.0379	.2452	.5195			.8047
4	.0003	.0065	.0629	.2654	.4311		.7662
5		.0010	.0135	.0852	.2878	.3123	.6998
6		.0002	.0026	.0220	.1159	.2499	.3906
7			.0005	.0050	.0364	.1171	.1590
8			.0001	.0010	.0099	.0419	.0529
9				.0002	.0024	.0127	.0153
10					.0005	.0035	.0040
11					.0001	.0009	.0010
12						.0002	.0002

Table 17(c): Fraction of pupils in grade three who will be in grade j, n years later and the staying ratios.

n years \ grade j	3	4	5	6	7	staying ratio
1	.1255	.8247				.9502
2	.0159	.2089	.6681			.8929
3	.0020	.0400	.2550	.5551		.8521
4	.0003	.0063	.0653	.2987	.4027	.7738
5		.0011	.0140	.1011	.2702	.3864
6		.0002	.0027	.0275	.1092	.1396
7			.0005	.0066	.0344	.0415
8			.0001	.0014	.0093	.0108
9				.0003	.0023	.0026
10				.0001	.0005	.0006
11					.0001	.0001

Table 17(d): Fraction of pupils in grade four who will be in grade j, n years later and the staying ratios

grade j / n years	4	5	6	7	staying ratio
1	.1269	.8111			.9380
2	.0162	.2067	.6749		.8978
3	.0021	.0397	.2773	.4905	.8096
4	.0003	.0068	.0765	.2667	.3503
5		.0011	.0177	.0910	.1098
6		.0002	.0037	.0249	.0288
7			.0007	.0060	.0067
8			.0001	.0013	.0014
9				.0003	.0003
10				.0001	.0001

Table 17(e): Fraction of pupils in grade five who will be in grade j, n years later and the staying ratios.

grade j / n years	5	6	7	staying ratio
1	.1271	.8335		.9606
2	.0163	.2357	.6072	.8592
3	.0021	.0504	.2526	.3051
4	.0003	.0096	.0703	.0802
5		.0017	.0164	.0181
6		.0003	.0034	.0038
7		.0001	.0007	.0008
8			.0001	.0001

Table 17(f): Fraction of pupils in grade six who will be in grade j, n years later and the staying ratios.

grade j / n years	6	7	staying ratio
1	.1550	.7306	.8856
2	.0241	.2110	.2351
3	.0038	.0458	.0496
4	.0006	.0089	.0095
5	.0001	.0016	.0017
6		.0003	.0003
7		-	-

Table 17(g): Fraction of pupils in grade seven who will be in grade j, n years later and the staying ratios.

grade j / n years	7	staying ratio
1	.1346	.1346
2	.0181	.0181
3	.0024	.0024
4	.0003	.0003
5	-	-

Comments on Tables 17(a) - 17(g)

If the Kenyan primary education system evolves according to the Exponential probability model then 0.8020 of the new entrants would still be in the system after one year of schooling. After two years of schooling 0.7130 of these pupils would still be in school. Similar figures can be obtained for the other years. For example, after eight years of schooling 0.1506 of the grade one pupils would still be in primary school.

For pupils already in grade two, 0.9069 of the pupils would still be in primary school after one year. After two years, 0.8559 of the grade two pupils would still be in the system and so on. Infact after eight years of schooling 0.0529 of the grade two pupils would still be in school. We can similarly obtain the proportions for pupils in any of the other school grades. For example, for pupils enrolled in grade seven during the initial time, 0.1346 of these pupils would still be in primary school after one year. After two years of schooling 0.0203 of these pupils would still be in the system and so on. After five years of schooling all the grade seven pupils would have left primary school if the system evolves according to the Exponential probability transition model.

Table 18 below gives the school drop-out rates for the Kenyan primary education system under the Exponential probability transition model.

Table 18: The school drop-out rates within n years, under the Exponential probability model.

n years \ grade j	1	2	3	4	5	6	7
1	.1980	.0931	.0498	.0620	.0393	.1144	.8654
2	.2870	.1441	.1072	.1022	.1409	.7649	.9819
3	.3321	.1953	.1479	.1904	.6949	.9504	.9976
4	.3737	.2338	.2262	.6497	.9198	.9905	.9997
5	.4050	..3001	.6136	.8902	.9819	.9983	. 1
6	.4536	.6095	.8605	.9712	.9962	.9997	1
7	.6666	.8410	.9585	.9933	.9993	.9999	1
8	.8794	.9471	.9891	.9986	.9999	. 1	1
9	.9445	.9847	.9974	.9997	. 1	1	1
10	.9823	.9960	.9994	.9999	1	1	1
11	.9949	.9990	.9999	1	1	1	1
12	.9987	.9998	. 1	1	1	1	1
13	.9997	1	1	1	1	1	1
14	.9999	1	1	1	1	1	1
15	1	1	1	1			

Comments on Table 18

When the Exponential probability model is used to describe the education process, 0.1980 of grade one pupils, enrolled at the initial time, would have dropped out of school after the first year. After two years of schooling 0.2870 of these pupils would have dropped out of school and so on. The rest of these proportions are in the first column of Table 18. For example, the highest drop out from primary school for these grade one pupils is in the seventh year when 0.6666 would have dropped out of the system.

Of the pupils enrolled in grade two during the initial time, 0.0931 would have dropped out of school after one year. After two years of schooling 0.1441 of these grade two pupils would have dropped out of the system and so on. Similar values can be obtained for pupils enrolled in any of the other school grades. For example, of the pupils in grade seven during the initial time, 0.8654 would have dropped out of the system after one year. After two years of schooling 0.9819 of these grade seven pupils would have left the system. Infact all the pupils in primary school during the initial time would have left the system after fifteen years, according to the Exponential probability transition model.

Tables 19(a) - (b) below give the school survival times and the probabilities of promotion between the grades of the Kenyan primary education system, under the Exponential probability transition model.

Table 19(a): The expected length of stay in grade j by pupils in grade i and the school survival times under the Exponential probability model.

grade j / grade i	1	2	3	4	5	6	7	survival time
1	1.1504	.8863	.7894	.7455	.6925	.6812	.5681	5.5134
2		1.1461	1.0229	.9667	.8983	.8842	.7384	5.6566
3			1.1437	1.0817	1.0057	.9908	.8287	5.0506
4				1.1456	1.0656	1.0509	.8808	4.1429
5					1.1458	1.1313	.9508	3.2279
6						1.1836	.9982	2.1818
7							1.1555	1.1555

Table 19(b): The probability of a student, enrolled in grade i at the initial time, ever reaching grade j according to the Exponential probability model.

grade i \ grade j	1	2	3	4	5	6	7
1	1	.7733	.6902	.6508	.6044	.5754	.4916
2		1	.8944	.8438	.7840	.7470	.6390
3			1	.9442	.8777	.8371	.7172
4				1	.9300	.8879	.7623
5					1	.9558	.8228
6						1	.8639
7							1

Comments on Tables 19(a)-(b)

When the education system evolves according to the Exponential probability transition model, then a pupil in grade one at the initial time would be expected to take an average of 5.5134 years in primary school. A pupil in grade two during the initial time would be expected to take an average of 5.6566 years in primary school according to this model. We can obtain similar values for pupils in any of the other school grades. For example a pupil in grade seven during the initial time would have taken an average of 1.1555 years in the system

[Refer to Table 19(a)].

A pupil enrolled in grade one during the initial time had a 0.4916 chance of reaching the highest primary school grade. The probability of a pupil in grade two, at the initial time, reaching grade seven was 0.6390. Similar probabilities are obtainable for pupils of the other grades. For example, a grade six pupil would have had a 0.8639 chance of reaching grade seven, according to the Exponential probability model [see Table 19(b)].

In Table 20 below we give the expected length of schooling (ELS) by any pupil in school at the initial time when the system evolves according to the Exponential probability transition model.

Table 20.

grade j	1	2	3	4	5	6	7	Expected length of schooling (ELS)
length of stay in grade j	.2655	.4153	.5298	.6454	.7332	.8569	.8213	4.2674

Comments on Table 20

When the system is assumed to evolve according to the Exponential probability model, any pupil in the system at the initial time is expected to spend 0.2655 of a year in grade one. Any of these pupils is expected to spend 0.4153 of a year in grade two and so on. These values increase with grade size to the highest value in grade six. That is, any of these pupils is expected to take 0.8569 of a year in grade six. We note again that for grade seven, the expected length of stay is 0.8213 which is lower than that for grade six. Generally any pupil in primary school at the initial time would be expected to take approximately 4.2674 years in the system if the system evolves according to the Exponential probability model.

3.5 CONCLUSIONS

This section starts with a summary of some comparative remarks on the three probability transition models discussed in this chapter. This is followed by some general conclusions concerning the application of the time dependent probability transition models to describe the educational process.

COMPARATIVE REMARKS

On the Average Time to Stabilize

When the transition process is assumed to evolve according to the Normal probability model, the transition process will take between 137 and 1848 time periods to stabilize. According to the

Logistic probability model, the transition process will take between 62 and 984 time periods to stabilize. Finally the Exponential transition process is expected to take between 22 and 385 time periods before stabilizing.

In all the three models the transitions from grade seven take longest to stabilize. This could possibly be due to high repeats in this grade especially by pupils who fail the secondary school qualifying examination and have to resit the examination. It is evident that when the system evolves according to the Exponential probability model it will reach a stable level faster than in the other two models.

On the Goodness of Fit

Both the Normal and the Logistic probability models fit the transition process with P-values between 0.6 to 1.0. The Exponential probability model, fits the transition process with a slightly lower P-value of between 0.5 to 1.0. This seems to suggest that the Normal and the Logistic probability models are better fits to the process than the Exponential probability model. However this difference in the P-values does not seem to be significant. All the three models may therefore be used to study the transition process of the Kenyan primary education system.

On the Standard Error of Estimation

The relative standard error in using the Normal probability model to describe the transition process during the initial time is at most 22%. It is at most 19% when using the Logistic probability model to describe the transition process, and at most 18% when using the Exponential probability model. It seems that during the initial time the Exponential probability model describes the transition process better than the other two probability models, even though it is quite comparable with the Logistic probability model. During the first year, again the Exponential and the Logistic probability models have maximum relative standard errors of 20% and 21% respectively. They seem to describe the process better than the Normal probability model which has a maximum relative error of 24%. In fact after ten years, the Exponential probability model still performs better with a maximum relative error of 34.5% followed by the Logistic probability model with a maximum of 34.9% and then the Normal probability model with a maximum of 35%. After sixteen years however, the model with the least maximum standard error is the Normal probability model with a value of 40% followed by the Logistic probability model with a maximum relative error of 43%, and finally the Exponential model with a maximum relative error of 45%.

On average therefore, it seems that the Logistic probability model performs consistently better as compared to the other two probability models. This is an observation based on both the goodness of fit test of the models and the standard error of estimates in using the models.

On the Measures of Academic Retention

It is observed that for new entrants in the education system, the Normal probability model generally predicts higher retention ratios compared to either the Logistic or the Exponential probability models, which have nearly the same retention ratios. For example, for the Normal probability model, 83% of the new entrants were still in school after one year. This figure is high compared to 80% who were still in school after one year according to either the Logistic or the Exponential probability models. We can compare the school retention rates for pupils in any of the other school grades during the initial time. For example for pupils of grade seven, again the Normal probability model predicts in general a higher retention tendency as compared to either the

Logistic probability model or the Exponential probability model.

As a consequence of the above observation we may conclude that in general the Normal probability model predicts the lowest drop-out tendency as compared to the other two probability models both of which predict approximately the same drop-out tendencies. For example, according to the Normal probability model all the grade seven pupils will have left the system after six years. On the other hand, according to either the Logistic or the Exponential probability model , all the grade seven pupils will have left primary school after five years of schooling. We particularly note that all the pupils enrolled in primary school during the initial time will have left primary school after fifteen years, if the system is described by either the Logistic or the Exponential probability models. Of course this time will be more for the Normal probability model because of its high pupil retention property.

Within an eight-year period a new entrant into primary school is expected to take 5 years 5months in the system according to the Normal probability model. According to the Logistic and the Exponential probability models a new entrant is expected to take

5 years 6 months before leaving the system. We can obtain similar comparative figures for pupils of the other school grades. For example, a grade seven pupil is expected to take approximately 1 year 2 months in primary school according to any of the three probability models.

On the other hand, a new entrant into primary school has a 39% chance of reaching grade seven within an eight year period, according to the Normal probability model. According to either the Logistic or the Exponential model, a new entrant has a 49% chance of ever reaching grade seven. For a pupil of grade six the chance of reaching grade seven is approximately 92% according to the Normal probability model. It is approximately 87% according to the Logistic probability model and approximately 86% according to the Exponential probability model. Here we observe that when the system evolves according to the Normal probability model then a grade six pupil has a higher chance of ever reaching grade seven than when the system evolves according to either the Logistic or the Exponential probability models.

When the education process is described according to the Normal probability model, any pupil in primary school during the initial time is expected to take approximately 4 years 2 months in primary school during the first eight years of schooling. According to either the Logistic or the Exponential probability models, any pupil is expected to take approximately 4 years 3 months before leaving the system.

GENERAL CONCLUSIONS

The following general remarks may be deduced from the foregoing discussions:

(i) The Kenyan primary education system is found to be time inhomogeneous on the basis of the homogeneity test.

(ii) The suggested time dependent probability transition models are all found to be useful in describing the transition process of the Kenyan primary education system on the basis of the goodness of fit tests. In particular these models may quite adequately describe the transition process for time periods close to the base year.

(iii) The Normal probability model predicts a higher pupil retention tendency compared to the two other probability models. However, we need to note that for the Normal probability model we assumed that transition rates converge to the maximum value of one. This may have led to the observed results. In both the Logistic and the Exponential probability models the optimum values of the transition rates were equal and less than one. It is therefore more appropriate for comparative purposes to look at the Logistic and the Exponential models. In this case we observe that the Exponential probability model and the Logistic probability model behave nearly the same in terms of their retention characteristics.

In the next chapter we shall study the Kenyan primary education system under the generalised cohort analysis. In particular we shall assume the system to evolve according to the Logistic probability transition model. The choice of the model used is due to the consistency of the relative standard errors of estimates and the goodness of fit.

CHAPTER IV

THE GENERALISED COHORT MODEL

4.1 INTRODUCTION

In this chapter we shall describe a model which traces the flow of a cohort of students through an education system. In particular we shall use a cohort transition model as an application of the more general Markov chain model, described in chapter three, in studying the flow of a group of students through the education system.

For our purpose, the term cohort will be used to mean a group of students regardless of age or socio-economic background, who enter the first grade in the same academic year. The cohort model is based on the fact that students flow through an education system in successive cohorts. Thus by tracing the progress of a student belonging to a given cohort, the enrolment in a given grade at a given time can be looked at as being composed of students belonging to various cohorts.

4.2 THE PROPOSED MODEL

Assumptions

Suppose the states of the education system are denoted by the integers $1,2,...,s$ where s is the number of non-absorbing states of the system.

We shall consider the time dependent Markov chain model, of the type suggested in chapter III, where the flows out of a given grade are governed by time dependent probabilities. That is, the probability of an individual in state i at time t moving to state j at time t+1 depends only on i,j and t and not on any previous moves the individual may have made. If the probability of moving from state i to state j in the time interval (t,t+1) is denoted by $p_{ij}(t)$ then the transition matrix of flows is given by

$$P(t) = ((p_{ij}(t))) \qquad [c.f.(3.1)]$$

and depends on time t. We shall further assume that

(i) new enrollment in the education system is only through the first grade of the education system;

(ii) a student can only be promoted to the next higher grade, so that

$$p_{ij}(t) = 0 \quad \begin{array}{l} \text{for all} \quad j = i+2, i+3, \ldots, s \quad \text{and} \\ \text{for all} \quad j < i \; ; \end{array}$$

(iii) a student can repeat a given grade any number of times, however the probability of repeating infinitely is zero;

(iv) all individuals behave independently.

It is therefore clear that at the end of each academic year some students pass and thus move to the next higher grade, during the following year, with probabilities

$$p_{i,i+1}(t), \quad \text{for} \quad i = 1,2,\ldots,s$$

and some fail to pass and thus repeat the same grade with probabilities

$$p_{ii}(t), \quad \text{for} \quad i = 1,2,\ldots,s$$

However these two events are not the only ones as far as movement of students through the school system is concerned, since a student may die or withdraw for whatever reason. Thus the transition probabilities satisfy

$$p_{ii}(t) + p_{i,i+1}(t) \leq 1; \quad 0 \leq p_{ij} \leq 1, \quad i < j.$$

If we let $n_{ij}(t)$ denote the number of individuals in grade i at time t who move to grade j at time (t+1); these numbers represent the flows between the states of the system. The stock of individuals in state i at time t is

$$n_i(t) = \sum_{j=1}^{N} n_{ij}(t) \qquad \text{[c.f. (3.2)]}$$

for $N > s$ where, states $s+1, s+2, \ldots, N$ are absorbing states. We can treat the states individually because of the assumption of independent flows. Thus assuming the multinomial distribution, the maximum likelihood estimate of $p_{ij}(t)$ is given by

$$\hat{p}_{ij}(t) = n_{ij}(t) \Big/ n_i(t) \qquad \text{[c.f. (3.4)]}$$

It is the proportion of students enrolled in the i-th school grade in year t who move to grade j in year (t+1).

The Logistic Probability Transition Model

In order to illustrate the applications of the generalised cohort model, we shall model the transition probabilities in terms of endogeneous variables which could be non-quantifiable but cause variation in the transition probabilities as time changes. We proceed by assuming that the transition probabilities may be regarded as probabilities of occurrence of random events which characterise the transition process. In particular as in chapter III we shall assume the transition process is characterised by a random process whose distribution function over time is

$$F(t) = 1\Big/\left[1 + \exp\{-(\alpha_{ij} + \beta_{ij}t)\}\right], \quad -\infty<t<\infty, \beta_{ij}>0 \qquad \text{[c.f. (3.40)]}$$

where α_{ij} and β_{ij} are as in (3.40). Defining $\rho_{ij}(t)$ as in chapter III we can then write

$$\rho_{ij}(t) = \exp(\alpha_{ij} + \beta_{ij}t) \Big/ \{1 + \exp(\alpha_{ij} + \beta_{ij}t)\},$$

[c.f. (3.41)]

The parameters α_{ij} and β_{ij} may be estimated from the regression equations

$$\text{logit } \hat{\rho}_{ij}(t) = \alpha_{ij} + \beta_{ij}t \qquad \text{[c.f. (3.43)]}$$

The transition probabilities $p_{ij}(t)$'s are then obtained by transforming back the values $\rho_{ij}(t)$'s to their original values.

In chapter III we saw that the Logistic probability transition model is suitable for studying the Kenyan primary education system. For illustration purposes we shall now give application of the generalised cohort analysis assuming that the system evolves according to the Logistic probability transition model.

The Time Dependent Cohort Model

In the cohort model we consider $n_i(t)$, the number of students enrolled in grade i at time t, as composed of students belonging to different cohorts. Let $q_i^j(0)$ denote the probability that a student belonging to an initial time cohort enrolls in the i-th school grade, j years after admission into the system. Then assuming new enrolment to be only through the first grade, we have

$$q_i^j(0) > 0 \quad \text{for} \quad j = i-1, i, i+1, \ldots$$
$$= 0 \quad \text{for} \quad j \leq i-2$$

Since new entrance into the system is through grade one and the fact that a new entrant at time zero will be expected to be in grade i, (i-1) years later if there is no repetition, it is evident that

$$q_1^o(0) = 1, \quad q_1^1(0) = p_{11}(0) \tag{4.1}$$

and

$$q_i^{i-1}(0) = p_{12}(0)p_{23}(1)p_{34}(2)\ldots p_{i-1,i}(i-2)$$

Generally starting from any base year, t, we may define $q_i^j(t)$ as the probability that a student of the time t cohort enrols in the i-th grade j years later. For $(i \geq 1)$ it can then be seen that

$$q_1^j(t) = p_{11}(t+j-1)\, q_1^{j-1}(t) \tag{4.2}$$

and

$$q_{i+1}^{j+1}(t) = p_{i,i+1}(t+j)q_i^j(t) + p_{i+1,i+1}(t+j)q_{i+1}^j(t)$$

Specifically when the base year is denoted by, t=0, we have

$$q_1^j(0) = p_{11}(j-1)q_1^{j-1}(0)$$

and (4.3)

$$q_{i+1}^{j+1}(0) = p_{i,i+1}(j)q_i^j(0) + p_{i+1,i+1}(j)q_{i+1}^j(0)$$

An alternative expression for $q_i^j(t)$'s may be obtained as follows. Expressing the number of students, $n_i(t)$ in grade i in year t in terms of the number enrolled in grade i and (i-1) in year (t-1) and the new enrollment N_t, we have

$$E\left[n_1(t)/N_t, n_1(t-1)\right] = N_t + p_{11}(t)n_1(t-1)$$

and (4.4)

$$E\left[n_i(t)/n_{i-1}(t-1), n_i(t-1)\right] = p_{i-1,i}(t-1)n_{i-1}(t-1)+p_{ii}(t)n_i(t-1)$$

which in matrix form becomes

$$E\left[\underline{n}(t)/N_t, \underline{n}(t-1)\right] = P'(t-1)\underline{n}(t-1) + \underline{e}\ N_t \qquad (4.5)$$

where $\underline{n}(t)$ and $\underline{n}(t-1)$ are enrollment vectors, in years t and (t-1) respectively; $P'(t-1)$ is the transpose of the transition matrix at time t-1, that is

$$P(t-1) = \left(\left(p_{ij}(t-1)\right)\right),$$

where $p_{ij}(t-1)$'s are the fitted transition probabilities under the assumed model; and

$\underset{-}{e}_{s\times 1} = (1,0,0,\ldots,0)'$. Substituting for $\underline{n}(t-1)$ reccursively in (4.5) we obtain

$$E\left[\underline{n}(t)/N_o,N_1,\ldots,N_t\right] = \sum_{j=0}^{t} P^{(j)'}(t-j)\ \underline{e}N_{t-j} \tag{4.6}$$

where

$$P^{(j)}(t-j) = \prod_{k=0}^{j} P(t-k)\ .$$

It is the j-step transition matrix from time t-j to time t.

We can also express $n_i(t)$ as the proportional sum of students belonging to all the preceding cohorts starting from time 0 upto time t as

$$E\left[n_i(t)/N_o,N_1,\ldots,N_t\right] = \sum_{j=0}^{t} q_i^j(t-j)N_{t-j} \tag{4.7}$$

which in vector form becomes

$$E\left[\underline{n}(t)/N_o,N_1,\ldots,N_t\right] = \sum_{j=0}^{t} \underline{q}^j(t-j)N_{t-j} \tag{4.8}$$

where

$$\underline{q}^j(t-j) = \left(q_1^j(t-j),q_2^j(t-j),\ldots,q_s^j(t-j)\right)'.$$

Comparing (4.6) and (4.8) we obtain

$$P^{(j)'}(t-j)\ \underline{e} = \underline{q}^j(t-j) \tag{4.9}$$

Thus $\underline{q}^j(t-j)$ is just the first column of $P^{(j)'}(t-j)$ or the first row of $P^{(j)}(t-j)$. For pupils joining school at any time t we can then arrange the probabilities $q_i^j(t)$ in a matrix Q(t) as,

$$Q(t) = \begin{bmatrix} q_1^o(t) & q_2^o(t) & \cdots & q_s^o(t) \\ q_1^1(t) & q_2^1(t) & \cdots & q_s^1(t) \\ \vdots & \vdots & & \vdots \\ q_1^n(t) & q_2^n(t) & \cdots & q_s^n(t) \end{bmatrix} \qquad (4.10)$$

The matrix Q(t) is called the retention probability matrix for pupils joining school at time t, and n is the maximum number of years a student stays in school

Next let

$$\underline{q}_k^j(t-j) = p^{(j)'}(t-j)\ \underline{e}_k$$

where $\underline{e}_k$ is a column vector of zeros except a one in the k-th position. Then $\underline{q}_k^j(t-j)$ is a column vector whose i-th component gives the probability that a student in grade k at time (t-j) will enrol in the i-th school grade j years later. It is denoted by $q_{ki}^j(t-j)$.

The vector $\underline{q}_k^j(t-j)$ gives the general retention probabilities. In fact it is just the k-th row of the j step transition matrix $p^{(j)}(t-j)$. Generally for pupils in any school grade k at time t we may define

$$\underline{q}_k^j(t) = P^{(j)'}(t)\ \underline{e}_k$$

where $\underline{e}_k$ is as before, then $\underline{q}_k^j(t)$ is a column vector whose i-th entry denoted, $q_{ki}^j(t)$, gives the probability of a student in grade k at time t enrolling in grade i, j years later. We can then arrange the probabilities $q_{ki}^j(t)$'s in a matrix of the type in (4.10) which is called a general retention probability matrix. That is

$$Q_k(t) = \begin{bmatrix} q_{k1}^o(t) & q_{k2}^o(t) & \cdots & q_{ks}^o(t) \\ q_{k1}^1(t) & q_{k2}^1(t) & \cdots & q_{ks}^1(t) \\ \vdots & \vdots & & \vdots \\ q_{k1}^n(t) & q_{k2}^n(t) & \cdots & q_{ks}^n(t) \end{bmatrix} \qquad (4.11)$$

When $k = 1$, (4.10) and (4.11) coincide. The reccursive formulae in (4.2) can now be generalised for any time t and grade k as,

$$q_{k,i+1}^{j+1}(t) = p_{i,i+1}(t+j)q_{ki}^j(t) + p_{i+1,i+1}(t+j)q_{k,i+1}^j(t) \qquad (4.12)$$

and

$$q_{kk}^j(t) = p_{kk}(t+j-1)\ q_{kk}^{j-1}(t)$$

The elements of the transition matrix $Q_k(t)$ are conditional probabilities in the sense that they are defined by considering students entering grade k at time t as constituting a cohort. We can however obtain the transition matrix of absolute probabilities, based on initial cohort of students as follows. Let $q^{*j}_{ki}(0)$ denote the probability that a pupil after entering grade k for the first time at time zero will enrol in grade i after j years. Suppose that $q_i(t)$ is the probability of a member of an arbitrary cohort of students enrolling in the i-th school grade for the first time at time t. Then

$$q_1(t) = 1, \tag{4.13}$$

since all the new enrolment is only in the first grade. For the other school grades we have

$$q_{i+1}(t) = p_{i,i+1}(t-1)\sum_{j=0}^{\infty} q_i^j(t-j-1), \quad i = 1,2,\ldots,s-1 \tag{4.14}$$

This value can be estimated from previously available transition data. We then easily see that

$$q^{*j}_{ki}(0) = q_k(0)q^j_{ki}(0); \quad i,k = 1,2,\ldots,s; \quad j = 0,1,2,\ldots,t \tag{4.15}$$

We can therefore write the new matrix of absolute probabilities as

$$Q_k^*(0) = \begin{bmatrix} q_{k1}^{*o}(0) & q_{k2}^{*o}(0) & \cdots & q_{ks}^{*o}(0) \\ q_{k1}^{*1}(0) & q_{k2}^{*1}(0) & \cdots & q_{ks}^{*1}(0) \\ \vdots & \vdots & & \vdots \\ q_{k1}^{*n}(0) & q_{k2}^{*n}(0) & \cdots & q_{ks}^{*n}(0) \end{bmatrix} \quad (4.16)$$

4.3 APPLICATIONS OF THE GENERALISED COHORT MODEL

In this section we give some applications of the model proposed in section 4.2 based on the Kenyan primary education system. We begin by estimating the probability of joining the i-th grade for the first time at time zero. These estimates are given in Table 21 below. The estimates are obtained using equation (4.14). Here time zero refers to any fixed reference time, in this case 1980.

Table 21: The probability of enrolling in grade i for the first time at time zero

grade (i)	1	2	3	4	5	6	7
probability of first enrolment at time zero	1	.6990	.7580	.7346	.6720	.6320	.5303

Comments on Table 21

As given in the above table the probability of an individual enrolling for the first time in grade one is 1. The probability that an individual after joining school will enroll in grade two for the first time at time zero is 0.6990. This probability is higher for grade three than for grade two since the probability that an individual after joining school enrolls in grade three for the first time at time zero is 0.7580. This may be a consequence of the higher promotion rate from grade two than from grades one or three. Similar values can be obtained for first enrolments in the rest of the school grades. For example, the probability that a pupil after joining school will enroll in grade seven for the first time at time zero is 0.5303. It is however evident that those probabilities are generally decreasing with grade size.

Tables 22(a) - (g) below give the probabilities of individuals of different cohorts, who after joining grade k for the first time at time zero, will enrol in grade i, j years later for $i,k = 1,2,\ldots,s$; $j = 1,2,\ldots$; when the system evolves according to the logistic probability transition model. These probabilities are $q^{*j}_{ki}(0)$'s which are elements of the retention matrix $Q_k^*(0)$. The sum over all the grades, of these probabilities, that is

$$r_k^j(0) = \sum_{i=1}^{s} q_{ki}^{*j}(0), \tag{4.17}$$

gives the fraction of students who after joining grade k for the first time at time zero, will still be in school j years later. It is comparable to the staying ratios mentioned earlier in chapter III. It is called the school staying ratio after joining grade k for the first time at time zero.

Tables 22(a) - (g) The school retention rates for the cohort of students who join grade k for the first time at time zero under the Logistic probability model.

Table 22(a) Fraction of pupils in grade one for the first time at time zero who will be in grade i,j years later.

grade i / year j	1	2	3	4	5	6	7	staying ratios
1	.1285	.6742						.8027
2	.0168	.1714	.5255					.7137
3	.0022	.0332	.1998	.4320				.6672
4	.0003	.0058	.0513	.2199	.3493			.6266
5		.0010	.0111	.0707	.2231	.2896		.5955
6		.0002	.0022	.0183	.0861	.2304	.2097	.5469
7			.0004	.0042	.0260	.1076	.1945	.3327
8			.0001	.0009	.0068	.0384	.1036	.1498
9				.0002	.0016	.0117	.0416	.0551
10					.0003	.0032	.0139	.0174
11					.0001	.0006	.0043	.0050
12						.0002	.0011	.0013

Table 22(b): Fraction of pupils in grade two for the first time at time zero who will be in grade i,j years later.

grade i / year j	2	3	4	5	6	7	staying ratio
1	.0866	.5473					.6339
2	.0111	.1362	.4506				.5979
3	.0014	.0258	.1699	.3648			.5619
4	.0002	.0044	.0432	.1846	.3028		.5352
5		.0007	.0092	.0589	.2006	.2196	.4890
6		.0001	.0017	.0151	.0804	.1746	.2719
7			.0003	.0034	.0252	.0814	.1103
8			.0001	.0007	.0068	.0290	.0366
9				.0001	.0018	.0087	.0106
10					.0003	.0024	.0027
11					.0001	.0006	.0007
12						.0001	.0001

Table 22(c): Fraction of pupils in grade three for the first time at time zero who will be in grade i,j years later

year j \ grade i	3	4	5	6	7	staying ratios
1	.0939	.6255				.7194
2	.0118	.1569	.5070			.6757
3	.0015	.0299	.1923	.4214		.6451
4	.0002	.0051	.0490	.2257	.3062	.5862
5		.0008	.0105	.0762	.2045	.2920
6		.0001	.0020	.0207	.0824	.1052
7			.0004	.0049	.0259	.0312
8			.0001	.0011	.0070	.0082
9				.0002	.0017	.0019
10				.0001	.0004	.0005
11					.0001	.0001

Table 22(d): Fraction of pupils in grade four for the first time at time zero who will be in grade i, j years later

grade i / year j	4	5	6	7	staying ratios
1	.0922	.5963			.6885
2	.0118	.1510	.4967		.6595
3	.0015	.0289	.2032	.3617	.5953
4	.0002	.0050	.0559	.1959	.2570
5		.0008	.0129	.0667	.0804
6		.0001	.0027	.0182	.0210
7			.0005	.0044	.0049
8			.0001	.0010	.0011
9				.0002	.0002
10				.0001	.0001

Table 22(e): Fraction of pupils in grade five for the first time at time zero who will be in i, j years later.

grade i / year j	5	6	7	staying ratio
1	.0851	.5609	.	.6460
2	.0109	.1581	.4099	.5789
3	.0014	.0337	.1701	.2052
4	.0002	.0065	.0472	.0539
5		.0011	.0110	.0121
6		.0002	.0023	.0025
7		.0001	.0005	.0006
8			.0001	.0001

Table 22(f): Fraction of pupils in grade six for the first time at time zero who will be in grade i, j years laters.

grade i / year j	6	7	staying ratio
1	.0976	.4641	.5617
2	.0152	.1336	.1488
3	.0024	.0289	.0313
4	.0004	.0055	.0059
5	.0001	.0010	.0011
6		.0002	.0002
7		.0001	.0001

Table 22(g): Fraction of pupils in grade seven for the first time at time zero who will be in grade i, j years later.

grade i / year j	7	staying ratio
1	.0714	.0714
2	.0096	.0096
3	.0013	.0013
4	.0002	.0002

Comments on Tables 22(a) - (g)

Table 22(a) gives the probabilities that a pupil after joining grade one for the first time at time zero will be in the other school grades j years later, when the system is assumed to evolve according to the Logistic probability transition model. Since new enrolment is only through the first school grade, Table 22(a) is identical to Table 12(a). For example for pupils entering grade one for the first time at time zero, 0.1285 of them will still be in grade one after one year while 0.6742 will be in grade two after two years. Similar values can be obtained for other grades, for pupils joining grade one for the first time at time zero. For example, after eight years of schooling, 0.1498 of the pupils who joined grade one for the first time at time zero will still be in school.

The probability of joining grade two for the first time at time zero is 0.6990, and the probability of still being in school one year later is 0.6339. The probability is 0.5979 that a pupil after joining grade two for the first time at time zero will still be in school two years later and so on. Similar probabilities can be obtained for pupils in the other school grades as well.

For example, the probability that a pupil after joining grade seven for the first time at time zero will still be in primary school one year later is 0.0714. The probability is 0.0096 that a pupil after joining grade seven for the first time at time zero will still be in school two years later.

The School Drop-out Ratio

The probability that a student from the initial cohort drops out of the education system j years after admission is given by

$$b_j(0) = 1 - \sum_{i=1}^{s} q_i^j(0) \tag{4.18}$$

The probability that a student in grade k at time zero will drop-out of the education system j years later is given by

$$b_j^k(0) = 1 - \sum_{i=1}^{s} q_{ki}^j(0) \tag{4.19}$$

This is a measure of drop-out from the various school categories after j years of schooling. The quantity

$$b_j^{*k}(0) = 1 - \sum_{i=1}^{s} q_{ki}^{*j}(0), \tag{4.20}$$

measures the proportion of students who drop-out of school j years after joining grade k for the first time at time zero. It is also a measure of the proportion of the initial stock of students who leave school with final education of at least grade k. The probability of drop-out within x years for a student in grade k at time zero is given by

$$b^k(0) = \sum_{j=1}^{x} b_j^k(0) \tag{4.21}$$

We can also define

$$b^{*k}(0) = \sum_{j=1}^{x} b_j^{*k}(0) \tag{4.22}$$

as the proportion who drop-out of school within x years after having joined grade k for the first time at time zero.

Table 23 below gives the drop-out rates within x years after joining grade k for the first time at time zero, when the system is assumed to evolve according to the logistic probability model.

Table 23: The school drop-out rates within x years after joining grade k for the first time at time zero under the logistic probability transition model.

grade k / x years	1	2	3	4	5	6	7
1	.1973	.3661	.2806	.3115	.3540	.4383	.9286
2	.2863	.4021	.3244	.3407	.4211	.8512	.9904
3	.3327	.4381	.3549	.4047	.7948	.9687	.9987
4	.3733	.4648	.4138	.7430	.9461	.9941	.9998
5	.4045	.5110	.7148	.9196	.9878	.9989	1
6	.4532	.7281	.8948	.9789	.9975	.9998	1
7	.6673	.8898	.9688	.9951	.9995	.9999	1
8	.8502	.9634	.9918	.9989	.9999	1	1
9	.9450	.9894	.9980	.9998	. 1	1	1
10	.9825	.9973	.9995	.9999	1	1	1
11	.9950	.9993	.9999	1	1	1	1
12	.9987	.9999	1	1	1	1	1
13	.9997	1	1	1	1	1	1
14	.9999	1	1	1	1	1	1
15	1	1	1	1			

Comments on Table 23

Table 23 gives the probabilities of dropping out from school within x years after joining grade k for the first time at time zero. Again, since new enrolment is only through the first school grade, it follows that the first column of Table 23 is just the same as the first column of Table 13. For example 0.1973 of the pupils joining grade one for the first time at time zero will have left school after one year. Within fifteen years of schooling all these pupils will have left primary school.

Of the pupils who join grade two for the first time at time zero 0.3661 will have dropped out within the first year. Of these pupils, 0.4021 will have dropped out of school within two years of schooling. Similar values may be obtained for pupils joining any of the other school grades for the first time at time zero. For example, of the pupils who join grade seven for the first time at time zero, 0.9286 will have dropped out of school within the first year of schooling. All these grade seven pupils will have left primary school after five years.

The School Survival time

The probability that a student in grade i at time zero will still remain in the education system j years after admission is $r_i^j(0)$, as given in equation (4.17). Consider the cohort $N_i(0)$ of students enrolled in grade i at time zero, and follow its members upto the last one in school. Then

$$\left(1 + r_i^1(0) + r_i^2(0) + \ldots + r_i^n(0)\right) N_i(0)$$

is the expected total number of school years of the cohort $N_i(0)$, where n is the maximum number of years they can remain in school. It follows that the expected length of schooling by a student in grade i at time zero, denoted by $\ell_i(0)$ is given by,

$$\ell_i(0) = 1 + r_i^1(0) + r_i^2(0) + \ldots + r_i^n(0) \tag{4.23}$$

Let $\ell_i^k(0)$ denote the average length of time spent in grade k by a student in grade i at time zero. Then,

$$\ell_i^k(0) = \sum_{j=0}^{n} q_{ik}^j(0) , \tag{4.24}$$

some of the $q_{ik}^j(0)$'s in (4.24) are zero. For example if $i > k$ then they are all zeros, but if $i \leq k$ then

$$\ell_i^k(0) = q_{ik}^{k-i-1}(0) + q_{ik}^{k-i}(0)+\ldots+q_{ik}^{k-i-1+n}(0) \tag{4.25}$$

The sum

$$\ell_i(0) = \sum_{k=i}^{s} \ell_i^k(0); \quad i = 1,2,\ldots,s \tag{4.26}$$

gives the expected remaining schooling time for a student in grade i at time zero.

We now consider some applications of the generalised cohort model based on the generalised retention matrix. Let $q_{ik}^{*j}(0)$ denote the probability that a student enrols in grade k, j years after joining grade i for the first time at time zero. Then,

$$\ell_i^{*k}(0) = q_{ik}^{*(k-i-1)}(0) + q_{ik}^{*(k-i)}(0)+\ldots+q_{ik}^{*(k-i-1+n)}(0) \tag{4.27}$$

is the average length of time spent in grade k after enrolling for the first time in grade i at time zero. In general $\ell_i^{*k}(0)$ is the sum of the k-th column of the generalised retention matrix given in (4.16). The expected school survival time after enrolling in grade i for the first time at time zero is therefore given by

$$\ell_i^*(0) = \sum_{k=i}^{s} \ell_i^{*k}(0) \tag{4.28}$$

We can now form a matrix of the expected survival times in grade k after entering grade i for the first time at time zero for i,k = 1,2,...,s as follows:

$$L^*(0) = \begin{bmatrix} \ell_1^{*1}(0) & \ell_1^{*2}(0) & \cdots & \ell_1^{*s}(0) \\ \ell_2^{*1}(0) & \ell_2^{*2}(0) & \cdots & \ell_2^{*s}(0) \\ \vdots & \vdots & & \vdots \\ \ell_s^{*1}(0) & \ell_s^{*2}(0) & \cdots & \ell_s^{*s}(0) \end{bmatrix} \tag{4.29}$$

The sum of the i-th row of $L^*(0)$ gives the expected remaining school years after having enrolled in grade i for the first time at time zero.

Table 24(a) below gives the school survival times after enrolling in grade i for the first time at time zero.

Now, let $\delta_i^k(0)$ denote the probability that a pupil enrols for the first time in grade i at time zero and spends some time in grade k before leaving the system. Then the expected length of stay in grade k by such a student is given by

$$\ell_i^{*k}(0) = \left(\delta_i^k(0)\right)\left(\ell_k^k(0)\right) + \left(1 - \delta_i^k(0)\right) \times 0 \tag{4.30}$$

from which we obtain

$$\delta_i^k(0) = \ell_i^{*k}(0)/\ell_k^k(0) \qquad (4.31)$$

We note here that $\ell_k^k(0)$ is the expected length of stay in grade k by those in grade k at time zero.

Table 24(b) below gives the probability of a pupil enrolling in grade i for the first time at time zero and reaching grade k before leaving school.

Table 24(a): The School Survival times for pupil who enrol in grade i for the first time at time zero.

grade k / grade i	1	2	3	4	5	6	7	survival time
1	1.1478	.8858	.7903	.7462	.6934	.6819	.5689	5.5143
2		.7982	.7146	.6750	.6276	.6179	.5163	3.9496
3			.8654	.8183	.7612	.7503	.6282	3.8234
4				.8402	.7821	.7719	.6481	3.0423
5					.7695	.7607	.6412	2.1714
6						.7476	.6335	1.3811
7							.6128	.6128

Table 24(b): The probability of joining grade i for the first time at time zero and ever reaching grade k before leaving school.

grade k / grade i	1	2	3	4	5	6	7
1	1	.7757	.6922	.6523	.6054	.5765	.4923
2		.6990	.6258	.5901	.5481	.5224	.4468
3			.7580	.7154	.6647	.6347	.5437
4				.7346	.6830	.6525	.5609
5					.6720	.6430	.5549
6						.6320	.5482
7							.5303

Comments on Tables 24(a) - (b)

The lengths of stay in school by pupils who join grade one for the first time at time zero are the same as those obtained earlier in Table 14(a). For example, a pupil joining grade one for the first time at time zero is expected to take an average of 5.5143 years in primary school. A pupil who will enter grade two for the first time at time zero is expected to take an average of 0.7982 of a year in that grade. Infact such a pupil is expected to take an average of 3.9496 years in primary school. Similar values may be obtained for pupils joining any of the other grades for the first time at time zero. For example, a pupil who enrols in grade seven for the first time at time zero is expected to take 0.6128 of a year in the school system. [See Table 24(a)].

As is the case above, the probabilities of ever reaching any of the school grades after joining grade one for the first time at time zero are identical to those of Table 14(b) in chapter III. For example, a pupil joining grade one for the first time at time zero, has a chance of 0.4923 of reaching the highest primary school grade. For a pupil joining school at some time, the probability of enrolling for the first time in grade two at time zero is 0.6990. This is the same as the probability of ever joining grade two. The probability of joining grade two for the first time at time zero and ever reaching grade seven is 0.4468. Similar probabilities can be obtained for the other grades. For example, the probability that a pupil in grade six for the first time at time zero will ever reach grade seven is 0.5482. [See Table 24(b)].

Expected Length of Schooling

Suppose that

$$\underline{p}^*(0) = \left(p_1^*(0),\ p_2^*(0),\ldots,p_s^*(0)\right)' \qquad (4.32)$$

is the proportion of pupils enrolled in the various school grades for the first time at time zero. These proportions may be estimated by

$$\hat{p}_i^*(0) = n_i^*(0) \Big/ \sum_{i=1}^{s} n_i^*(0) \qquad (4.33)$$

where $n_i^*(0)$ is the number joining grade i for the first time at time zero. It is not difficult to see that at any time

$$n_i^*(t) = n_i(t) - \left[\sum_{j=1}^{n} n_i(t-j)\ q_i^j(t-j)\right] \qquad (4.34)$$

The expected length of stay in the various school grades by any of the pupils in these grades for the first time at time zero is therefore

$$\underline{\ell}^{*\prime}(0) = \underline{p}^{*\prime}(0)\ L^*(0) \qquad (4.35)$$

where $L^*(0)$ is the matrix given in (4.29). The expected length of schooling $\left(ELS^*(0)\right)$ by any of the students joining the various grades for the first time at time zero is therefore the sum of the entries of the product in equation (4.35). That is

$$ELS^*(0) = \underline{p}^{*'}(0)\ L^*(0)\ \underline{j} \qquad (4.36)$$

where $\underline{j}$ is an (s x 1) column vector of ones. Table 25 below gives the expected length of schooling for pupils who enrol in the various school grades for the first time at time zero.

Table 25: The expected length of schooling for those who enrol in the various school grades for the first time at time zero.

grade i	1	2	3	4	5	6	7	ELS*(0)
ELS in i	.2609	.3516	.4360	.5181	.5713	.6461	.5969	3.3809

Comments on Table 25

Any student enrolled in the various grades for the first time at time zero is expected to spend approximately 0.2609 of a year in grade one. This value increases generally with the grade. For example, any pupil enrolled for the first time in any of the school grades at time zero is expected to take about 0.3516 of a year in grade two before leaving the system. The expected length of schooling by any of these pupils in grade six is approximately 0.6461 of a year. This value is greater than the expected length of schooling in grade seven, by these pupils, which is about 0.5969 of a year. This may be due to the high retention property of grade six. In general any pupil enrolled for the first time in the various grades at time zero is expected to spend approximately 3.3809 years in school before leaving the system.

4.4 CONCLUSIONS

Since new enrollment in primary school is assumed to be only through the first grade, there is complete similarity between the retention properties of new entrants into primary school under the general Markov process in chapter III and those of generalised cohort model. We now give some conclusions, based on the applications of the generalised cohort model to the Kenyan primary education system.

A pupil joining school for the first time at time zero has a 13% chance of still being in grade one after one year. We expected that about 67% of these grade one pupils will be in grade two after one year. This observation implies that approximately 80% of the pupils who join grade one for the first time at time zero will still be in primary school after one year. Similar values are obtainable during the other time periods of study. For example, after eight years of schooling approximately 15% of these pupils will still be in primary school and so on.

For the pupils joining grade two for the first time at time zero, approximately 63% will still be in school after one year. After two years of schooling about 60% of these pupils will still be in school and so on. We can obtain similar values for pupils joining any of the other school grades for the first time at time zero. For example, only about 7% of the pupils who are in grade seven for the first time at time zero will still be in school after one year. After two years of schooling only about 1% of these grade seven pupils will still be in school, and all of them will have left primary school after four years.

The school drop-out ratios for those enrolled in grade one for the first time at time zero are the same as those obtained earlier in chapter III for the Logistic probability model. For the pupils who after joining school enrol in grade two for the first time at time zero, about 37% will have dropped out of school within the first year. Within two years of schooling about 40% of these grade two pupils will have dropped out of school and so on. Infact all these grade two pupils will have left primary school within thirteen years. Drop out rates can be obtained for pupils in any of the other school grades for the first time at time zero. For example about 93% of the pupils who enrol for the first time in grade seven will have dropped out of primary school within the first year; 99% of these pupils will have left school within two years and so on. All these grade seven pupils will have left the system in four years.

The school survival times for new entrants into primary school at time zero are the same as those obtained for the generalised Markov chain model in chapter III. A pupil in grade two for the first time at time zero is expected to take an average of 10 months in grade two. These pupils are expected to take about 9 months in grade three and so on. In total a pupil who enrols in grade two for the first time at time zero is expected to take an average of 3 years 11 months in primary school. The school survival times can be obtained for pupils joining any of the other school grades at time zero for the first time. For example, it is approximately 3 years 1 month for pupils joining grade three for the first time at time zero. On the other hand a pupil enrolled in grade seven for the first time at time zero is expected to take about 7 months in school before leaving the system. As is expected the percentage of entrants into school at time zero who ever reach grade one is 100% since they all enter into that grade at time zero. About 49% of those pupils who join grade one for the first time at time zero will ever reach grade seven. These percentages are obtainable for pupils of the other school grades. For example 70% of the pupils in a cohort will join grade two for the first time at time zero. Of these pupils who join grade two for the first time at time zero about 45% of them will have a chance of ever reaching the highest primary school grade. Also 63% of a cohort of pupils will have a chance of joining grade six for the first time at time zero. Of these grade six pupils 55% will have a chance of ever reaching grade seven.

Any pupil enrolled at time zero for the first time in any of the various school grades is expected to take approximately 3 months in grade one. These pupils are expected to take about 4 months in grade two and so on. We note particularly that pupils who are enrolled for the first time at time zero in the various grades are expected to spend about 8 months in grade six and 7 months in grade seven respectively. This is possibly a consequence of the high retention property of grade six as compared to grade seven. In general those pupils who join the various school grades for the first time at time zero are expected to take approximately 3 years 5 months in primary school.

To summarize, we note that the generalised cohort model provides an alternative method of describing the flow of students in an educational system. Regarding the application of the generalised cohort model to the Kenyan primary education system, we make the following general remarks:

(i) Most of the results on the retention properties obtained by the generalised cohort model coincide with those of the general Markov chain model when we consider the new entrants into the first school grade at time zero.

(ii) In general the average repeat rates for pupils joining the school grades for the first time at time zero is approximately 9%. On the other hand the average promotion rate for these pupils is about 59%.

(iii) A pupil has on average a 67% chance of enrolling in any of the remaining school grades for the first time at time zero, after joining school.

(iv) In general any pupil enrolled in the education system for the first time at time zero, in any of the grades, is expected to take approximately 3 years 5 months in primary school.

CHAPTER V

PROBABILITY MODEL FOR ESTIMATING AND CONTROLLING ACADEMIC SURVIVAL

5.1 INTRODUCTION

In chapter III we suggest some time dependent transition models which incorporate endegenous factors over time by means of probability distributions. We shall now go further and assume that the transition probabilities change in time due to the effect of some changing social and economic factors. In particular we shall assume that the transition rates are affected by some quantifiable variables which change with time. The proposed model will then be used to obtain some measures of academic survival. In addition an attempt will be made to control some of these quantifiable variables in order to achieve transition rates which are as close as possible to some desired future values. Some consequences of the proposed model will be demonstrated, using data from the Kenyan primary education system.

In most developing countries, education patterns are constantly changing due to rapid population growth and other socio-economic factors. This trend calls for a time dependent model which would incorporate factors which are internal or external to the system. Factors like pupil-teacher ratios; cost of education; proportion of trained teachers; rate of inflation and many others have some effect on pupil performance and hence on the transition rate. In this chapter we propose a method of modelling the transition rates as responses dependent on quantifiable factors of the type mentioned above. Using multivariate statistical control theory, we shall attempt to obtain the most suitable values of these factors in order that some targeted transition rates are achieved at future time periods. The proposed model will then be used together with the theory of the non-homogeneous Markov chains to define some measures of academic survival. These measures will include; the school staying ratios; the school drop out and completion ratios; the school survival times and the expected length of schooling. Estimates of these measures will be computed using the stocks and flows data of the Kenyan primary education system. It is assumed in the calculations that the transition rates are dependent on, pupil-teacher ratios, cost of educating an individual per grade per year and sex ratios.

5.2 THE PROPOSED MODEL

Suppose the states of the education process are denoted by integers 1,2,...,N where N is the number of possible states in the system. The probability of moving from state i to j in the time interval (t,t+1) is denoted by $P_{ij}(t)$ and the transition matrix of flows is given by

$$P(t) = ((p_{ij}(t))) \qquad \text{[c.f. (3.1)]}$$

The maximum likelihood estimates of $p_{ij}(t)$ can be computed from the stocks and flows data using equation (3.4) of chapter III. In chapter III we modelled these transition rates as functions depending on time parameter t. In this chapter we shall not only be concerned with the time parameter but we shall go further and consider factors which cause these changes over time. We suggest factors such as policies of admission into the first grade; proportion of students per teacher; class densities; average experience of teachers; cost of education; sex ratios and many other factors.

Suppose that the transition rates $p_{ij}(t)$'s are affected by some quantifiable factors which change in time, denoted by

$$x_{1t}, x_{2t}, \ldots, x_{st}$$

We may then write

$$p_{ij}(t) = p_{ij}\left(x_{1t}, x_{2t}, \ldots, x_{st}\right) \tag{5.1}$$

Approximating this functional relationship by a linear function we may write

$$p_{ij}(t) = \sum_{\ell=1}^{s} \beta_{ij(\ell)} x_{\ell t} + \varepsilon_{ij}(t) \tag{5.2}$$

where $\varepsilon_{ij}(t)$ is the error in approximation at time t. To simplify notation, we assume that there are m possible combinations of ij's denoted by k = 1,2,..., m, so that equation (5.2) becomes

$$p_k(t) = \sum_{\ell=1}^{s} \beta_{k\ell} x_{\ell t} + \varepsilon_k(t); \quad \begin{array}{l} k = 1,2,\ldots,m \\ t = 1,2,\ldots,\tau \end{array} \tag{5.3}$$

where τ is the number of periods over which data is available. Combining the τ observations for each k we may write equation (5.3) in matrix form as;

$$\underset{\tau\times 1}{\underline{p}_k} = \underset{\tau\times s}{X} \; \underset{s\times 1}{\underline{\beta}_k} + \underline{\varepsilon}_k \tag{5.4}$$

or

$$\begin{bmatrix} p_k(1) \\ p_k(2) \\ \\ \\ p_k(\tau) \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1s} \\ x_{21} & x_{22} & \cdots & x_{2s} \\ \\ \\ x_{\tau 1} & x_{\tau 2} & \cdots & x_{\tau s} \end{bmatrix} \begin{bmatrix} \beta_{k1} \\ \beta_{k2} \\ \\ \\ \beta_{ks} \end{bmatrix} + \begin{bmatrix} \varepsilon_k(1) \\ \varepsilon_k(2) \\ \\ \\ \varepsilon_k(\tau) \end{bmatrix}$$

for $k = 1,2,\ldots,m$.

We shall assume further that there is no demotion in the system and that promotion is only to the next higher grade, so that

$$p_{ij}(t) = 0 \quad \text{for} \quad j \geq i+2 \quad \text{and} \quad j < i$$

where i and j are non-absorbing states. In this system there are only three types of possible transitions from a given grade, namely promotion, repeat and dropout. Since the transition ratios from each grade sum to one, we need only model two of the ratios as the other can be obtained once these two are known. In this case we shall for example consider only promotion and repeat rates. In cases where one may drop out or be promoted to more than one alternative states, the types of transitions will be more than three, in which case one has to model more than two transition rates.

Suppose that the total number of combinations of promotions and repeats are m, arranged so that the first m/2 are promotions in ascending order of grade. Schematically we have

$$\underbrace{p_1(t),p_2(t),\ldots,p_{m/2}(t)}_{\text{promotions}} \; ; \; \underbrace{p_{m/2+1}(t),p_{m/2+2}(t),\ldots,p_m(t)}_{\text{repeats}}$$

so that

$$p_k(t) = p_{i,i+1}(t) \quad \text{and} \quad p_{k+m/2}(t) = p_{ii}(t), 0<k\leq m/2$$

The problem now is to obtain the best estimate of $\underline{\beta}_k$, from the linear model defined in equation (5.4)

$$\text{subject to:} \quad p_\ell(t) + p_{\ell+m/2}(t) \leq 1 \text{ for } 0 < \ell \leq m/2$$

$$p_k(t) \geq 0 \text{ for all } k \text{ and } t \tag{5.5}$$

This requires minimizing

$$s(\underline{\beta}_k) = (\underline{p}_k - X\underline{\beta}_k)'(\underline{p}_k - X\underline{\beta}_k)$$

subject to;

$$\underline{p}_\ell + \underline{p}_{\ell+m/2} \leq \underline{1} \;, \quad 0 < \ell \leq m/2$$

$$\underline{p}_k \geq \underline{0} \;, \qquad k = 1,2,\ldots,m \tag{5.6}$$

where $\underline{1}$ and $\underline{0}$ are τx1 column vectors of ones and zeros respectively.

Combining all the m values, the equations in $\underline{\beta}_k$'s may be written as

$$\underset{\tau\times m}{P} = \left(\underline{p}_1, \underline{p}_2, \ldots, \underline{p}_m\right) = \underset{\tau\times s}{X} \; \underset{s\times m}{B} + \underset{\tau\times m}{E} \tag{5.7}$$

where

$$B = \left(\underline{\beta}_1, \underline{\beta}_2, \ldots, \underline{\beta}_m\right)$$

and

$$E = \left(\underline{\varepsilon}_1, \underline{\varepsilon}_2, \ldots, \underline{\varepsilon}_m\right).$$

We shall make an additional assumption that $\underline{\varepsilon}_k \sim N\left[\underline{0}, \sigma_k^2 I_\tau\right]$ where $\underline{0}$ is a $\tau\times 1$ column vector of zeros; σ_k^2 is an unknown constant and I_τ is an identity matrix of order τ. Using the least squares estimation procedure, the best estimate of B is given by

$$\hat{B} = \left(\hat{\underline{\beta}}_1, \hat{\underline{\beta}}_2, \ldots, \hat{\underline{\beta}}_m\right) = (X'X)^{-1}X'P \tag{5.8}$$

provided $(X'X)^{-1}$ exists, which is true as long as there is no high multicollinearity among the considered quantifiable variables. This estimate will also be optimal for our problem as long as the constraint set is satisfied. If however the constraint set is not satisfied, then alternative modelling procedures should be suggested for example the number of transition rates to be modelled may be adjusted. This adjustment may include using the average dropout rate and thus modelling for only the promotion rate.

Assuming that the constraint set is satisfied, then the fitted transition rates are

$$\underline{\hat{p}}_k = X\, \underline{\hat{\beta}}_k + \underline{\varepsilon}_k, \quad k = 1,2,\ldots,m \tag{5.9}$$

where for each t

$$\hat{\varepsilon}_k(t) = \underline{x}_t' \, (X'X)^{-1} \underline{x}_t \; \hat{\sigma}_k^2 \tag{5.10}$$

is the error in $p_k(t)$; $\underline{x}_t$ is the regressor vector at time t and $\hat{\sigma}_k^2$ is the mean square error in $p_k(t)$.

Testing for Goodness of Fit

Before proceeding to use the model we need to test how well the suggested model fits the observed values. Again as in chapter III suppose the data is available over τ time periods $t = 1,2,\ldots,\tau$. Then the expected transition numbers in the time interval $(t,t+1)$ under the model assumption is given by

$$\begin{aligned} n_{ij}^*(t) &= E\left[n_{ij}(t)/\text{Model}\right] \qquad [\text{c.f } (3.55)] \\ &= n_i(t)\, p_{ij}^*(t) \end{aligned}$$

for each i and j where $n_i(t)$ is the stock in grade i at time t and $p_{ij}^*(t)$ is the fitted value of $p_{ij}(t)$ corresponding to the time interval $(t,t+1)$. The chi-square statistic for testing the goodness of fit is

$$D = \sum_{u=1}^{k} \left(O_{ij}(u) - E^*_{ij}(u)\right)^2 / E^*_{ij}(u) \qquad [\text{c.f. } (3.56)]$$

where $O_{ij}(u)$ is the number of $n_{ij}(t)$'s which fall in the u-th interval and $E^*_{ij}(u)$ the corresponding number of $n^*_{ij}(t)$'s which fall in the u-th interval. This statistic has (k-1-s) degrees of freedom, where s is the number of unknown parameters and k, is as in chapter III.

The goodness of fit test is carried out to assess the closeness of fit between the observed and estimated flow values for all the seven grades, when the system is assumed to evolve according to the suggested model. It is found that the differences between the observed and fitted flow values are not significant with P-values ranging between 0.5 to 1.0. This suggests therefore, that the proposed model may be used to study the Kenyan primary education system.

USING THE MODEL TO PREDICT FUTURE TRANSITIONS

We have so far obtained a relationship between the transition rates and the quantifiable variables of the type

$$p_k(t) = \underline{x}_t' \underline{\beta}_k + \varepsilon_k(t); \quad \begin{array}{l} t = 1,2,\ldots,\tau \\ k = 1,2,\ldots,m \end{array} \qquad (5.11)$$

Suppose the regressor variable values are known at some future time, then the model may be used to estimate the future transition rates. For example, if the vector of values of the regressors at time $\tau+1$ is $\underline{x}_{\tau+1}$, the estimated transition rates at that time will be

$$\hat{p}_k(\tau+1) = \underline{x}'_{\tau+1}\hat{\underline{\beta}}_k + \hat{\varepsilon}_k(\tau+1); \quad k = 1,2,\ldots,m \qquad (5.12)$$

where $\hat{\varepsilon}_k(\tau+1) = \underline{x}'_{\tau+1}(X'X)^{-1}\underline{x}_{\tau+1}\hat{\sigma}_k$ is the error in the estimation. Generally we may estimate the transition rates h years later as

$$\hat{p}_k(\tau+h) = \underline{x}'_{\tau+h}\hat{\underline{\beta}}_k + \hat{\varepsilon}_k(\tau+h) \qquad (5.13)$$

whenever the regressor vector $\underline{x}_{\tau+h}$, $\tau+h$ years later, is known in advance and provided the constraints mentioned earlier are satisfied. In most cases, however, the future regressor variable values are usually unknown, in which case we may use some simulated or projected values based on their past trend. For example, if $\hat{\underline{x}}_{\tau+h}$ is the projected regressor vector h years later, then the estimated transition ratio will be

$$\hat{p}_k(\tau+h) = \hat{\underline{x}}'_{\tau+h}\hat{\underline{\beta}}_k + \hat{\varepsilon}_k(\tau+h) \qquad (5.14)$$

Results of the above type may in some cases be quite misleading since, in multiple linear regression, it is possible for a model to perform quite poorly at points far from the centre of the region constituting the original data. We may define such a region by considering the diagonal elements of the matrix

$$H = X(X'X)^{-1}X' \tag{5.15}$$

where, $\underset{s\times\tau}{X'} = (\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_\tau)$ is the matrix of the original data. The diagonal elements of H are of the type

$$h_{tt} = \underline{x}_t'(X'X)^{-1}\,\underline{x}_t \tag{5.16}$$

If $h_{max} = \max\{h_{tt};\ t = 1,2,\ldots,\tau\}$, then the region constituting the original data may be defined by the ellipsoid

$$R_{data} = \{\underline{x}_t;\ h_{tt} \leq h_{max}\} \tag{5.17}$$

It is the smallest convex set containing all the regressors constituting the original data and so it is the convex hull of the regressors.

In order to overcome this problem of regressor points falling far out of this region in the future time, we shall opt for a step by step procedure involving the improvement of the regression coefficients after each prediction step, as described below.

We have fitted the model

$$\underset{m\times 1}{\underline{p}(t)} = \underset{m\times s}{B_1} \; \underset{s\times 1}{\underline{x}_t} + \underset{m\times 1}{\underline{\varepsilon}(t)} , \quad t = 1,2,\ldots,\tau \qquad (5.18)$$

The first future transition ratios will be obtained as

$$\underset{m\times 1}{\hat{\underline{p}}(\tau+1)} = \underset{m\times s}{\hat{B}_1} \; \underset{s\times 1}{\hat{\underline{x}}_{\tau+1}} + \underset{m\times 1}{\hat{\underline{\varepsilon}}(\tau+1)} \qquad (5.19)$$

where,

$$\hat{B}_1' = \hat{B} = (Y_1'Y_1)^{-1}Y_1'P_1'$$

$$\underset{s\times\tau}{Y_1'} = X' = (\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_\tau) \quad \text{and}$$

$$P_1 = P' = (\underline{p}(1), \underline{p}(2), \ldots, \underline{p}(\tau))$$

To obtain the second future transitions, the estimated transitions at time $\tau+1$, $\hat{\underline{p}}(\tau+1)$ and its corresponding regressor vector $\hat{\underline{x}}_{\tau+1}$ are added to the original data bank so as to improve prediction. The second future transitions are therefore

$$\hat{\underline{p}}(\tau+2) = \hat{B}_2 \hat{\underline{x}}_{\tau+2} + \hat{\underline{\varepsilon}}(\tau+2) \qquad (5.20)$$

where now

$$\hat{B}_2' = \left(Y_2'Y_2\right)^{-1} Y_2'P_2'$$

$$\underset{s\times(\tau+1)}{Y_2'} = \left(Y_1';\ \hat{\underline{x}}_{\tau+1}\right) = \left(X';\ \hat{\underline{x}}_{\tau+1}\right)$$

and

$$\underset{m\times(\tau+1)}{P_2} = \left(P_1;\ \hat{\underline{p}}(\tau+1)\right) = \left(P';\hat{\underline{p}}(\tau+1)\right)$$

so that

$$\hat{B}_2' = \left(Y_1'Y_1 + \hat{\underline{x}}_{\tau+1}\hat{\underline{x}}_{\tau+1}'\right)^{-1}\left(Y_1'P_1' + \hat{\underline{x}}_{\tau+1}\hat{\underline{p}}'(\tau+1)\right) \tag{5,21}$$

Generally the k-th future transition ratios are obtained as

$$\hat{\underline{p}}(\tau+k) = \hat{B}_k\ \hat{\underline{x}}_{\tau+k} + \hat{\underline{\varepsilon}}(\tau+k) \tag{5.22}$$

where,

$$\begin{aligned}\hat{B}_k' &= (Y_k'Y_k)^{-1}Y_k'P_k' \\ &= \left(Y_{k-1}'Y_{k-1} + \hat{\underline{x}}_{\tau+k}\hat{\underline{x}}_{\tau+k}'\right)^{-1}\left(Y_{k-1}'P_{k-1}' + \hat{\underline{x}}_{\tau+k}\hat{\underline{p}}'(\tau+k)\right) \\ &= \left(X'X + \sum_{\alpha=1}^{k}\hat{\underline{x}}_{\tau+\alpha}\hat{\underline{x}}_{\tau+\alpha}'\right)^{-1}\left(X'P_1' + \sum_{\alpha=1}^{k}\hat{\underline{x}}_{\tau+\alpha}\hat{\underline{p}}'(\tau+\alpha)\right)\end{aligned} \tag{5.23}$$

and

$$\hat{\underline{\varepsilon}}(\tau+k) = \hat{\underline{x}}_{\tau+k}'\left(Y_k'Y_k\right)^{-1}\hat{\underline{x}}_{\tau+k}\ .\ \mathrm{diag}\left(\hat{\sigma}_{1k}, \hat{\sigma}_{2k}, \ldots, \hat{\sigma}_{mk}\right) \tag{5.24}$$

where $\hat{\sigma}_{\ell k}$ is the mean square error in $\hat{p}_{\ell}(t)$ for $t = 1,2,\ldots,\tau+k-1$.

We shall demonstrate the application of the regression model using data from the Kenyan primary education. For this purpose we use simulated values of some regressor variables. For our demonstration we shall fit the linear model

$$p_k(t) = \beta_{k1}x_{1t} + \beta_{k2}x_{2t} + \beta_{k3}x_{3t} + \varepsilon_k(t) \qquad (5.25)$$

where $k = (1,2,\ldots,14)$ represents all the possible promotions and repeats in the considered primary education system;

x_{1t} is the pupil teacher ratio at time t

x_{2t} is the cost of educating an individual per grade per year in K£.

and

x_{3t} is the pupil sex ratio at time t, taken as the proportion of males.

The observations of these quantifiable variables are available for $\tau = 17$ time periods and are given in table 26 below. The period considered is between 1964 to 1980 inclusive; x_{1t} and x_{3t} are extracted from the ministry of education annual report while x_{2t} is an approximation based on the actual fees paid by parents during the period 1964 - 1980.

Table 26: The regressor variables during the period 1964 - 1980 inclusive.

TIME (t) 1964-1980	pupil/teacher data value x_{1t}	cost/pupil (K£) approximate value x_{2t}	sex ratio Male proportion data value x_{3t}
1	36.3669	3	.6481
2	33.1699	4	.6359
3	31.1343	5	.6190
4	31.8258	6	.6087
5	31.9261	7	.5993
6	33.4726	8	.5949
7	34.4928	9	.5858
8	31.25	10	.5775
9	31.3271	11	.5708
10	32.1416	12	.5645
11	34.5594	13	.5512
12	33.4611	14	.5420
13	32.5281	15	.5369
14	31.1661	16	.5336
15	32.5541	17	.5911
16	39.8491	18	.5282
17	38.3122	19	.5253

The transition proportions $p_{ij}(t)$'s for the above periods are computed from the stocks and flows data available over that period. In order to estimate future transitions, the available regressors have been used to simulate values in the future by simple linear regression method. The i-th regressor value h years later is therefore

$$x_{i,\tau+h} = x_{i\tau} + r_i h \; ; \quad \begin{array}{l} i = 1,2,3 \\ h = 0,1,2,\ldots \end{array} \qquad (5.26)$$

where r_i is a rate estimated from the available regressor data.

5.3 APPLICATION OF THE MODEL

After fitting the model described above we can then estimate the elements of the time dependent transition matrix for each of the future time intervals $(t,t+1)$ for $t = \tau+h$, where $h = 1,2,\ldots$.

Suppose again that the education system consists of r absorbing and s non-absorbing states; $r+s = N$. Then the transition matrix for the flows can be put in the form given in equation (3.8) and the n-step transition matrix is given by equation (3.10). Expressions for the school retention rates, drop-out and completion rates, the expected length of schooling and the school survival times, also remain the same as those given in chapter III.

The School Retention Rates

We saw in chapter III that the school retention rate is simply the i-th entry of the column vector $Q^{(n)}(t)\ \underline{j}$, where $\underline{j}$ is an sx1 column vector of ones. That is, it is the probability that a student who is in grade i at time t will still be in school n years later. It is given by the sum

$$q_i^{(n)}(t) = \sum_{j=1}^{s} q_{ij}^{(n)}(t) \ ; \quad \begin{array}{l} i=1,2,\ldots,s \\ n=0,1,2,.. \end{array}$$

[c.f.(3.13)]

Tables 27(a) - (g) below give the school retention rates under the proposed model.

Table 27(a): Fraction of pupils in grade one who will be in grade j, n years later and the school retention rates.

n years \ grade j	1	2	3	4	5	6	7	School retention rates
1	.1358	.6798						.8156
2	.0152	.1655	.5558					.7365
3	.0018	.0285	.1966	.4612				.6881
4	.0002	.0044	.0456	.2161	.3722			.6385
5		.0006	.0089	.0633	.2220	.3095		.6043
6		.0001	.0016	.0151	.0804	.2336	.2276	.5584
7			.0003	.0032	.0231	.1049	.2037	.3352
8				.0006	.0058	.0367	.1062	.1493
9				.0001	.0013	.0111	.0425	.0550
10					.0003	.0030	.0145	.0178
11					.0001	.0008	.0045	.0054
12						.0002	.0013	.0015
13							.0003	.0003
14							.0001	.0001
15							-	-

Table 27(b): Fraction of pupils in grade two who will be in grade j, n years later and the school retention rates.

n years \ grade j	2	3	4	5	6	7	School retention rates
1	.1233	.7693					.8926
2	.0132	.1854	.6466				.8452
3	.0015	.0316	.2290	.5250			.7871
4	.0002	.0048	.0532	.2515	.4399		.7496
5		.0007	.0104	.0756	.2788	.3268	.6923
6		.0001	.0019	.0185	.1077	.2525	.3807
7			.0003	.0041	.0331	.1158	.1533
8			.0001	.0008	.0090	.0413	.0512
9				.0002	.0022	.0127	.0151
10					.0005	.0036	.0041
11					.0001	.0009	.0010
12						.0002	.0002

Table 27(c): Fraction of pupils in grade three who will be in grade j, n years later and the school retention rates.

grade j / n years	3	4	5	6	7	School retention rates
1	.1285	.8236				.9521
2	.0141	.2011	.6728			.8880
3	.0016	.0351	.2461	.5679		.8507
4	.0002	.0055	.0596	.2937	.4262	.7852
5		.0008	.0122	.0960	.2790	.3880
6		.0001	.0023	.0257	.1112	.1393
7			.0004	.0062	.0352	.0418
8			.0001	.0014	.0097	.0112
9				.0003	.0025	.0028
10				.0001	.0006	.0007
11					.0001	.0001

Table 27(d): Fraction of pupils in grade four who will be in grade j, n years later and the school retention rates.

grade j / n years	4	5	6	7	School retention rates
1	.1298	.8096			.9394
2	.0147	.2026	.6885		.9058
3	.0017	.0369	.2741	.5219	.8346
4	.0002	.0061	.0731	.2787	.3581
5		.0010	.0166	.0941	.1117
6		.0001	.0035	.0259	.0295
7			.0007	.0064	.0071
8			.0001	.0015	.0016
9				.0003	.0003
10				.0001	.0001

Table 27(e): Fraction of pupils in grade five who will be in grade j, n years later and the school retention rates.

grade j / n years	5	6	7	School retention rates
1	.1301	.8541		.9842
2	.0155	.2348	.6539	.9042
3	.0019	.0483	.2677	.3179
4	.0002	,0090	.0737	.0829
5		.0016	.0172	.0188
6		.0003	.0037	.0040
7		.0001	.0008	.0009
8			.0002	.0002
9			-	-

Table 27(f): Fraction of pupils in grade six who will be in grade j, n years later and the school retention rates.

grade j / n years	6	7	School retention rates
1	.1578	.7210	.8788
2	.0229	.2177	.2406
3	.0034	.0473	.0507
4	.0005	.0092	.0097
5	.0001	.0017	.0018
6		.0003	.0003
7		.0001	.0001

Table 27(g): Fraction of pupils in grade seven who will be in grade j, n years later and the school retention rates.

grade j / n years	7	School retention rates
1	.1329	.1329
2	.0179	.0179
3	.0025	.0025
4	.0003	.0003

Comments on Tables 27(a) - (g)

When the system is modelled according to the proposed linear regression model, 0.8156 of the pupils in grade one at the initial time will still be in primary school after one year. After two years of schooling 0.7365 of these pupils will still be in school and so on. In fact all the pupils in grade one at time zero will have left primary school after fifteen years, according to this regression model. Similar proportions can be obtained for pupils in any of the other school grades at time zero. For example, 0.8926 of the pupils, enrolled in grade two at time zero will still be in primary school after one year. After two years of schooling, 0.8452 of these pupils will still be in primary school and so on. On the other hand, 0.1329 of the pupils in grade seven at time zero will still be in primary school after one year. After two years 0.0179 of these grade seven pupils will still be in school. In fact all these grade seven pupils will have left primary school after five years.

The School Drop-out and Completion Rates

The drop out rate is the probability that a student entering grade i at time t will graduate n years later with final education k. As in chapter III, it is given by

$$g_{ik}^{(n)}(t) = \sum_{j=1}^{s} q_{ij}^{(n-1)}(t)\; g_{jk}(t+n-1); \quad \begin{matrix} i = 1,2,\dots,s \\ k = 1,2,\dots,r \end{matrix}$$

[c.f.(3.14)]

The rate of droping out from school within w years is obtained by summing the drop out rates for $n = 1$ to $n = w$. As given in chapter III, we have

$$\bar{g}_{ik}^{(w)}(t) = \sum_{n=1}^{w} g_{ik}^{(n)}(t) \;, \quad \begin{matrix} i = 1,2,\dots,s \\ k = 1,2,\dots,r \end{matrix}$$

[c.f.(3.15)]

as the drop out ratio within w years. The sum to infinity of the drop out rates gives the school absorbing rate, which is

[c.f.(3.16)]

$$\bar{g}_{ik}(t) = \sum_{n=1}^{\infty} g_{ik}^{(n)}(t)$$

Table 28 below gives the school drop out rates within x years under the proposed model.

Table 28: The fraction of pupils who drop out from grade j within x years.

grade j / x years	1	2	3	4	5	6	7
1	.1844	.1074	.0479	.0606	.0158	.1212	.8671
2	.2635	.1548	.1120	.0942	.0958	.7594	.9821
3	.3119	.2129	.1493	.1654	.6821	.9493	.9975
4	.3615	.2504	.2148	.6419	.9171	.9903	.9997
5	.3957	.3077	.6120	.8883	.9812	.9982	1
6	.4416	.6193	.8607	.9705	.9960	.9997	1
7	.6648	.8467	.9582	.9929	.9991	.9999	1
8	.8507	.9488	.9888	.9984	.9998	1	1
9	.9450	.9849	.9972	.9997	1	1	1
10	.9822	.9959	.9993	.9999	1	1	1
11	.9946	.9990	.9999	1	1	1	1
12	.9985	.9998	1	1	1	1	1
13	.9997	.9999	1	1	1	1	1
14	.9999	1	1	1	1	1	1
15	1	1	1	1			

Comments on Table 28

When the system is modelled according to the proposed linear regression model, 0.1844 of those in grade one at time zero will have dropped out of primary school within one year. Within two years of schooling, 0.2635 of these pupils will have dropped out and so on. The highest drop out rate for pupils in grade one at time zero is in the sixth year of schooling. In fact after fourteen years of schooling, practically all these pupils will have left primary school. Similar drop out rates can be obtained for pupils enrolled in any of the other school grades at time zero. For example, 0.1074 of those in grade two at time zero will have dropped out of primary school after one year. After two years of schooling 0.1548 of these grade two pupils will have left primary school and so on. On the other hand, 0.8671 of those in grade seven at time zero will have left primary school after one year. After two years of schooling 0.9821 of these grade seven pupils will have left the school system and so on. In fact all these grade seven pupils will have dropped out of primary school within five years.

The School Survival Time

The length of stay in grade j by a pupil enrolled in grade i at time t, during the next n years of schooling can be obtained as the (i,j)-th entry of the matrix $L_n(t)$ in equation (3.19) of chapter III. The expected length of stay in school grade j by those enrolled in grade i at time t can be obtained as the (i,j)-th entry of the infinite matrix series L(t) given in equation (3.21). The sum of the i-th row of L(t) gives the school survival time. This is the expected length of stay in school by a pupil in grade i at time t before graduating with any of the r final educations.

Finally the probability of ever reaching grade j after having enrolled in grade i at time t is given by

$$d_{ij}(t) = \ell_{ij}(t) / \ell_{jj}(t) \qquad \text{[c.f.(3.23)]}$$

Tables 29(a) and 29(b) below give the school survival times and the probabilities of promotion within the grades of the Kenyan primary education system under the proposed linear regression model.

Table 29(a): The expected length of stay in grade j by those in grade i at time zero and the survival times.

grade j / grade i	1	2	3	4	5	6	7	Survival times
1	1.1530	.8787	.8089	.7596	.7054	.7000	.6007	5.6063
2		1.1381	.9920	.9415	.8757	.8714	.7540	5.5727
3			1.1444	1.0662	.9935	.9913	.8645	5.0599
4				1.1464	1.0562	1.0567	.9288	4.1881
5					1.1478	1.1482	1.0172	3.3132
6						1.1848	.9973	2.1821
7							1.1536	1.1536

Table 29(b): The probabilities of ever reaching grade j after enrolling in grade i at time zero.

grade j / grade i	1	2	3	4	5	6	7
1	1	.7723	.7067	.6626	.6145	.5908	.5207
2		1	.8668	.8213	.7629	.7355	.6536
3			1	.9300	.8656	.8367	.7494
4				1	.9202	.8919	.8051
5					1	.9691	.8817
6						1	.8645
7							1

Comments on Tables 29(a) - (b)

If the system is assumed to evolve according to the proposed linear regression model, a pupil in grade one at the initial time takes an average of 5.6063 years in primary school before dropping out. A pupil in grade two at the initial time takes approximately 5.5728 years in primary school before leaving the system. We can obtain similar values for pupils enrolled in any of the other school grades at the initial time. For example, a pupil in grade seven at the initial time is expected to take 1.1536 years in primary school. [Refer to Table 29(a)].

Moreover, a pupil enrolled in grade one during the initial time has a 0.5207 chance of reaching the highest primary school grade according to this linear regression model. The probability of a pupil in grade two at the initial time, reaching grade seven is 0.6536 and so on. Similar probabilities can be obtained for pupils enrolled in the other school grades at the initial time. For example, a grade six pupil has a 0.8645 probability of reaching grade seven according to the proposed model. [See Table 29(b)].

Expected Length of Schooling (ELS)

The length of stay in school grade j by any student who is in school at time t during the next n years is the j-th component of the vector $\underline{\ell}_n$ given in equation (3.26). The expected length of schooling by any student in school at time t, during the next n years is given by ℓ_n defined in equation (3.27); and the length of stay in school grade j by any student in school at time t is the j-th entry of the vector $\underline{\ell}$ defined in equation (3.28). Furthermore the expected length of schooling (E.L.S) by any student in school at time t is given by equation (3.29).

Table 30 below gives the expected length of schooling by any of the pupils in school at the initial time under the proposed linear regression model.

Table 30: The expected length of schooling under the linear regression model.

grade i	1	2	3	4	5	6	7	Expected length of schooling (ELS)
ELS in i	.2661	.4122	.5287	.6420	.7293	.8618	.8501	4.2902

Comments on Table 30

If the system is assumed to evolve according to the proposed linear regression model, any pupil in the system at the initial time is expected to take approximately 0.2661 of a year in grade one. Any of these pupils is expected to take 0.4122 of a year in grade two and so on. These times increase with grade size, with the highest value in grade six. That is, any of these pupils is expected to take approximately 0.8618 of a year in grade six while they are expected to take approximately 0.8501 of a year in grade seven. This result may be a consequence of pupils dropping out from primary school mainly after grade six.

Generally, any pupil in primary school at the initial time is expected to take approximately 4.2902 years in primary school if the system evolves according to the linear regression model.

5.4 A STEP BY STEP CONTROL PROCEDURE FOR THE LINEAR REGRESSION TRANSITION MODEL

SINGLE PERIOD CONTROL

We have so far fitted a model of the type

$$\underset{m\times 1}{\underline{p}(t)} = \underset{m\times s}{B_1}\ \underset{s\times 1}{\underline{x}_t} + \underset{m\times 1}{\underline{\varepsilon}(t)}; \quad t = 1,2,\ldots,\tau \tag{5.27}$$

where,

$$\underline{p}(t) = \left(p_1(t), p_2(t), \ldots, p_m(t)\right)',$$

is the vector of all possible transition proportions in the system at time t. Let us make an additional assumption that the errors are normally distributed i.e. $\underline{\varepsilon}(t) \sim N\left[\underline{0}, \Sigma\right]$, where Σ is a symmetric positive definite matrix of covariances and B_1 is as given in equation (5.19).

It is often desired to control these relationships so as to optimize some objective function. For example, we may desire to exercise some control over the output values by appropriate adjustments on the input variables. Often there is a desired target level of the system output and any deviation from this target is thought of as an error, to be made as small as possible. In our problem the transition proportions are the output values while the regressor variables are the input values.

We shall use a Bayesian approach to develop the predictive distribution for the system output at some time in future. Optimal control is establ-shed by minimizing a loss function with respect to the predictive distribution. The loss function assumed is comprised of two parts; one part that is quadratic in the difference between actual and target system output and one part that is quadratic in the difference between present and future level of controllable variables. The latter part of the loss function is called the cost of control.

Suppose the first q_1 components of the regressors $\underline{x}_t$ are controllable and the remaining $q_2 = s - q_1$ are uncontrollable. For example, pupil-teacher ratios, pupils per class, cost of education per pupil per year, may be controllable; while time, drop-out rates, sex rates may be considered as uncontrollable.

Let the first future observation $\underline{p}(\tau+1)$ be estimated as before by

$$\underline{p}(\tau+1) = B_1 \underline{x}_{\tau+1} + \underline{\varepsilon}(\tau+1) \tag{5.28}$$

where B_1 may be estimated as before. Since we have assumed that $\underline{\varepsilon}(\tau) \sim N[\underline{0},\Sigma]$ it follows that for given B_1 and Σ, $\underline{p}(\tau+1) \sim N[B_1 \underline{x}_{\tau+1},\Sigma]$.

If prior information about (B_1, Σ) is available it may be used to obtain the predictive distribution of $\underline{p}(\tau+1)$. However in the absence of such information we may use a diffuse prior. Let us assume that the joint prior density of the parameters is given by

$$f\left(B_1, \Sigma^{-1}\right) \propto 1 \big/ \left|\Sigma^{-1}\right|^{(m+1)/2} \tag{5.29}$$

By definition, the predictive distribution of $\underline{p}(\tau+1)$ given $\underline{p}(1), \underline{p}(2), \ldots, \underline{p}(\tau)$ is then given by

$$f\left(\underline{p}(\tau+1) \big/ \underline{p}(1), \underline{p}(2), \ldots, \underline{p}(\tau)\right) = \iint f\left(\underline{p}(\tau+1) \big/ B_1, \Sigma^{-1}\right) . f\left(B_1, \Sigma^{-1} \big/ \underline{p}(1), \underline{p}(2), \ldots, \underline{p}(\tau)\right) dB_1 \, d\Sigma^{-1} \tag{5.30}$$

This integral can be evaluated to give the predictive density as

$$f\left(\underline{p}(\tau+1) \big/ \underline{p}(1), \underline{p}(2), \ldots, \underline{p}(\tau)\right) \propto 1 \Big/ \left[\nu_1 + \left(\underline{p}(\tau+1) - B_1 \underline{x}_{\tau+1}\right)' H\left(\underline{p}(\tau+1) - B_1 \underline{x}_{\tau+1}\right)\right]^{(\tau+1-s)/2} ; \tag{5.31}$$

which has the form of a multivariate $\underline{t}$ distribution with non-centrality parameter $B_1 \underline{x}_{\tau+1}$ and $\nu_1 = \tau - s - (m-1)$ degrees of freedom with

$$H = \nu_1 S_1^{-1} \Big/ \left(1 + \underline{x}_{\tau+1}' D_1^{-1} \underline{x}_{\tau+1}\right) \tag{5.32}$$

where

$$D_1 = \sum_{t=1}^{\tau} \underline{x}_t \, \underline{x}_t' \tag{5.33}$$

and

$$S_1 = \sum_{t=1}^{\tau} \left(\underline{p}(t) - B_1 \underline{x}_t\right)\left(\underline{p}(t) - B_1 \underline{x}_t\right)' \tag{5.34}$$

A description of control in the multivariate linear model is found in a book by Press (1982).

OPTIMUM CONTROL

Suppose the system output at time $\tau+1$ is to be kept as close as possible to a target value $\underline{a}_1$. Define the loss function

$$\ell\left(\underline{p}(\tau+1)\right) = \left(\underline{p}(\tau+1) - \underline{a}_1\right)' G \left(\underline{p}(\tau+1) - \underline{a}_1\right) + \left(\underline{x}_{\tau+1} - \underline{x}_\tau\right)' J \left(\underline{x}_{\tau+1} - \underline{x}_\tau\right) \tag{5.35}$$

where the first part represents the squared error due to deviation from the target and the second part is the squared error due to cost of controlling the system. The matrices G and J are pre-assigned symmetric matrices of non-negative constants.

The future risk ρ is therefore given by

$$\rho = E\left[\ell\left(\underline{p}(\tau+1) \,/\, \underline{p}(1), \underline{p}(2), \ldots, \underline{p}(\tau)\right)\right] \tag{5.36}$$

$$= \int \ell\left(\underline{p}(\tau+1)\right) \cdot f\left(\underline{p}(\tau+1) \big/ \underline{p}(1), \underline{p}(2), \ldots, \underline{p}(\tau)\right) d\underline{p}(\tau+1)$$

Substituting for $f\left(\underline{p}(\tau+1) \big/ \underline{p}(1), \underline{p}(2), \ldots, \underline{p}(\tau)\right)$, we obtain

$$\rho = \left[\text{tr } S_1 G/(\nu_1-2)\right]\left(1+\underline{x}'_{\tau+1} D_1^{-1} \underline{x}_{\tau+1}\right) + \left(B_1 \underline{x}_{\tau+1} - \underline{a}_1\right)' G\left(B_1 \underline{x}_{\tau+1} - \underline{a}_1\right)$$
$$+ \left(\underline{x}_{\tau+1} - \underline{x}_\tau\right)' J\left(\underline{x}_{\tau+1} - \underline{x}_\tau\right) \tag{5.37}$$

Partitioning the variables into controllable and uncontrollable variables we obtain,

$$\underline{x}_{\tau+1} = \begin{array}{c} \begin{bmatrix} \underline{x}_{1,\tau+1} \\ \underline{x}_{2,\tau+1} \end{bmatrix} \end{array} \begin{array}{c} q_1 \\ q_2 \end{array} ; \quad \underline{x}_\tau = \begin{bmatrix} \underline{x}_{1,\tau} \\ \underline{x}_{2\,\tau} \end{bmatrix} \begin{array}{c} q_1 \\ q_2 \end{array} ; \quad \underset{m \times s}{B_1} = \begin{array}{cc} \left[B_{11}\right. ; & \left.B_{12}\right] \\ q_1 & q_2 \end{array}$$

$$D_1^{-1} = \begin{array}{c} \begin{bmatrix} D_1^{11} & D_1^{12} \\ D_1^{21} & D_1^{22} \end{bmatrix} \\ \begin{array}{cc} q_1 & q_2 \end{array} \end{array} \begin{array}{c} q_1 \\ q_2 \end{array}, \quad \text{and} \quad J = \begin{array}{c} \begin{bmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{bmatrix} \\ \begin{array}{cc} q_1 & q_2 \end{array} \end{array} \begin{array}{c} q_1 \\ q_2 \end{array}$$

The risk is then minimized by differentiating with respect to the controllable variables and equating to zero. That is we solve the equations

$$\frac{\partial \rho}{\partial \underline{x}_{1,\tau+1}} = \underline{0} \tag{5.38}$$

These equations give the solution

$$\underline{x}^*_{1,\tau+1} = \left[B'_{11} G B_{11} + \{\mathrm{tr}\, S_1 G/(\nu_1 - 2)\} D_1^{11} + J_{11}\right]^{-1} \left[B'_{11} G\left(\underline{a}_1 - B_{12}\underline{x}_{2,\tau+1}\right)\right.$$

$$\left. - \{\mathrm{tr} S_1 G/(\nu_1-2)\} D_1^{12} \underline{x}_{2,\tau+1} + J_{11}\underline{x}_{1\tau} + J_{12}\left(\underline{x}_{2\tau} - \underline{x}_{2,\tau+1}\right)\right] \tag{5.39}$$

so that

$$\hat{\underline{x}}_{\tau+1} = \begin{bmatrix} \underline{x}^*_{1,\tau+1} \\ \\ \hat{\underline{x}}_{2,\tau+1} \end{bmatrix}$$

is the optimal value of the regressors at time $\tau+1$. This solution should give a value of $\underline{p}(\tau+1)$ which is as close as possible to our targeted output value $\underline{p}(\tau+1)_{TARGET} = \underline{a}_1$. Here $\hat{\underline{x}}_{2,\tau+1}$ is the vector of uncontrollable variables at time $\tau+1$ which may be estimated from past data. Theoretically, transitions at time $\tau+1$ are given by

$$\hat{\underline{p}}(\tau+1) = B_{11}\underline{x}^*_{1,\tau+1} + B_{12}\hat{\underline{x}}_{2,\tau+1} + \hat{\underline{\varepsilon}}(\tau+1) \tag{5.40}$$

even though for practical purposes we may use the targeted value $\underline{p}(\tau+1)_{TARGET} = \underline{a}_1$.

MULTIPLE PERIOD CONTROL

In general we may wish to control the system some $k \geq 1$ intervals of time later. In this case the $(\tau+k)$-th output, $\underline{p}(\tau+k)$, is to be made as close as possible to a target value. $\underline{p}(\tau+k)_{TARGET} = \underline{a}_k$. To achieve this we shall assume that there is a sequence of intermediate target values $\underline{a}_1, \underline{a}_2, \ldots, \underline{a}_{k-1}$ to pass through before attaining the final target value $\underline{a}_k$. The problem can then be looked upon as that of obtaining a sequence of optimal regressor-vectors $\hat{\underline{x}}_{\tau+1}, \hat{\underline{x}}_{\tau+2}, \ldots, \hat{\underline{x}}_{\tau+k}$ for the responses at each level. For example, the first optimal response will be attained by setting the controllable variable vector as given in (5.39). The optimal regressor vector at time $(\tau+1)$,

$$\hat{\underline{x}}_{\tau+1} = \begin{bmatrix} \underline{x}^*_{1,\tau+1} \\ \\ \hat{\underline{x}}_{2,\tau+1} \end{bmatrix}$$

and the corresponding target value $\underline{p}(\tau+1)_{TARGET} = \underline{a}_1$ are as before added to the original data bank to improve prediction in the next time step. The target for the second period control is now $\underline{p}(\tau+2)_{TARGET} = \underline{a}_2$ and following a similar procedure as before, the second optimum regressor vector is obtained as

$$\hat{\underline{x}}_{\tau+2} = \begin{bmatrix} \underline{x}^*_{1,\tau+2} \\ \\ \hat{\underline{x}}_{2,\tau+2} \end{bmatrix}$$

where the controllable vector is now

$$\underline{x}^*_{1,\tau+2} = \left[B_{21}' G B_{21} + \{trS_2G/(\nu_2-2)\}D_2^{11} + J_{11}\right]^{-1}\left[B_{21}' G(\underline{a}_2 - B_{22}\underline{x}_{2,\tau+2}) - \{trS_2G/(\nu_2-2)\}D_2^{22}\,\underline{x}_{2,\tau+2} + J_{11}\underline{x}^*_{1,\tau+1} + J_{12}(\underline{x}_{2,\tau+1} - \underline{x}_{2,\tau+2})\right] \tag{5.41}$$

It depends on the optimal regressors at time $\tau+1$. The matrix B_2 is obtained in a similar manner as given by equation (5.20); and

$$D_2 = \sum_{t=1}^{\tau+1} \underline{x}_t\, \underline{x}_t' = D_1 + \hat{\underline{x}}_{\tau+1}\, \hat{\underline{x}}_{\tau+1}' \tag{5.42}$$

$$S_2 = \sum_{t=1}^{\tau+1} \left(p(t) - B_2\underline{x}_t\right)\left(\underline{p}(t) - B_2\underline{x}_t\right)' = \sum_{t=1}^{\tau} \left(\underline{p}(t) - B_2\underline{x}_t\right)\left(\underline{p}(t) - B_2\underline{x}_t\right)' + \left(\underline{a}_1 - B_2\underline{x}_{\tau+1}\right)\left(\underline{a}_1 - B_2\underline{x}_{\tau+1}\right)' \tag{5.43}$$

and

$$\nu_2 = \nu_1 + 1.$$

Generally the target value at the k-th point in time i.e. $\underline{p}(\tau+k)_{TARGET} = \underline{a}_k$, will be attained optimally by setting the k-th controllable vector to be

$$\underline{x}^*_{1,\tau+k} = \left[B_{k1}' GB_{k1} + \{trS_kG/(\nu_k-2)\}D_k^{11} + J_{11}\right]^{-1}\left[B_{k1}' G(\underline{a}_k - B_{k2}\underline{x}_{2,\tau+k})\right.$$

$$-\{trS_k G/(\nu_k-2)\} D_k^{22}\underline{x}_{2,\tau+k} + J_{11}\underline{x}^*_{1,\tau+k-1} + J_{12}(\underline{x}_{2,\tau+k-1}-\underline{x}_{2,\tau+k})\Big] \tag{5.44}$$

which depends on all the intermediate optimals.

Theoretically the transition vector at this time is

$$\hat{\underline{p}}(\tau+k) = B_{k1}\underline{x}^*_{1,\tau+k} + B_{k2}\hat{\underline{x}}_{2,\tau+k} + \hat{\underline{\varepsilon}}(\tau+k) \tag{5.45}$$

Again as before $\underset{m\times s}{B_k} = \left[\underset{q_1}{B_{k1}} ; \underset{q_2}{B_{k2}}\right]$ is obtained as in equation (5.23); $\hat{\underline{x}}_{2,\tau+k}$ is a projected vector of the uncontrollable variables;

$$D_k = D_{k-1} + \hat{\underline{x}}_{\tau+k-1}\hat{\underline{x}}'_{\tau+k-1}; \tag{5.46}$$

$$S_k = \sum_{t=1}^{\tau}(\underline{p}(t)-B_k\underline{x}_t)(\underline{p}(t)-B_k\underline{x}_t)' + \sum_{t=\tau+1}^{\tau+k-1}(\underline{a}_t-B_k\underline{x}_t)(\underline{a}_t-B_k\underline{x}_t)'; \tag{5.47}$$

and

$$\nu_k = \nu_{k-1} + 1 \tag{5.48}$$

$$= \tau+k-s-(m-1).$$

Below we give some results on controlling the Kenyan primary education system over three periods with varying targets.

RESULTS ON CONTROLLING THE KENYAN PRIMARY EDUCATION SYSTEM OVER THREE PERIODS WITH VARYING TARGETS

These results are simply to serve as illustrative examples of the proposed model.

TRANSITION RATES AT THE INITIAL TIME, τ

p_{12}	p_{23}	p_{34}	p_{45}	p_{56}	p_{67}	p_{78}	p_{11}	p_{22}	p_{33}	p_{44}	p_{55}	p_{66}	p_{77}
.6798,	.7693,	.8232,	.8096,	.8541,	.7610,	.8671,	.1358,	.1233,	.1285,	.1298,	.1301,	.1578,	.1329

Regressor variables at time τ

Controllable ($\underline{x}_{1\tau}$)		Uncontrollable ($\underline{x}_{2\tau}$)	
Pupil-teacher ratio	= 38.3122	Sex ratio	= 0.5253
Cost	= K£19		

CASE I:- MAINTAINING A FIXED TRANSITION RATE TARGET OVER THE NEXT THREE YEARS

(a) After one year

(i) Target transition rates $\underline{p}(\tau+1)_{TARGET} = \underline{a}_1$

.6798, .7693, .8283, .8096, .8541, .7210, .8671, 1358, .1233, .1285, .1298, 1301, .1578, .1329

(ii) Theoretical values attained $\hat{\underline{p}}(\tau+1)$

.6951, .8206, .8385, .8212, .8516, .7732, .8649, .1097, .1044, .1073, .1104, .1167, .1428, .1351

(iii) Regressor variables at time $\tau+1$

Controllable $(\underline{x}^*_{1,\tau+1})$	Uncontrollable $(\hat{\underline{x}}_{2,\tau+1})$
Pupil-teacher ratio = 37.1758 Cost = K£20.1436	Sex ratio = 0.5164

(b) After <u>two years</u>

(i) Target transition rates $\underline{p}(\tau+2)_{TARGET} = \underline{a}_2$

.6798, .7693, .8283, .8096, .8541, .7210, .8671, .1358, .1233,

(ii) Theoretical values attained $\hat{\underline{p}}(\tau+2)$

.6786, .7961, .8230, .8080, .8415, .7488, .8569, .1173, .1102,

(iii) Regressor variables at time $\tau+2$

Controllable ($\underline{x}^*_{1,\tau+2}$)	Uncontrollable ($\hat{\underline{x}}_{2,\tau+2}$)
Pupil-teacher ratio =36.8818	Sex ratio = 0.5076
Cost =K£20.4473	

.1285, .1298, .1301, .1578, .1329

, .1136, .1162, .1214, ,1479, .1431

(c) After <u>three years</u>

(i) Target transition rates $\hat{\underline{p}}(\tau+3)_{\text{TARGET}} = \underline{a}_3$

.6798, .7693, .8283, .8096, .8541, .7210, .8671, .1358, .1233,

(ii) Theoretical values attained $\hat{\underline{p}}(\tau+3)$

.6701, .7795, .8109, .8002, .8347, .7337, .8516, .1228, .1139,

(iii) Regressor variables at time $\tau+3$

Controllable ($\underline{x}^*_{1,\tau+3}$)	Uncontrollable ($\hat{\underline{x}}_{2,\tau+3}$)
Pupil-teacher ratio =36.3867 Cost =K£20.9512	Sex ratio = 0.4987

.1285, .1298, .1301, .1578, .1329

, .1178, .1202, .1247, .1516, .1484

CASE II:- INCREASE IN PROMOTION RATES OF 0.01 PER YEAR WITH A DECREASE IN REPEAT RATES OF THE SAME AMOUNT PER YEAR, OVER THE NEXT THREE YEARS.

(a) After <u>one year</u>

(i) Target transition rates $\underline{p}(\tau+1)_{TARGET} = \underline{a}_1$

.6898, .7793, .8336, .8196, .8641, .7310, .8771, .1258, .1133, .1185, .1198, .1201, .1478, .1229

(ii) Theoretical values attained $\hat{\underline{p}}(\tau+1)$

.6974, .8213, .8384, .8224, .8523, .7747, .8655, .1097, .1043, .1072, .1103, .1167, .1428, .1345

(iii) Regressor variables at time $\tau+1$

Controllable ($x^*_{1,\tau+1}$)		Uncontrollable ($\hat{\underline{x}}_{2,\tau+1}$)	
Pupil-teacher ratio	=37.0703	Sex ratio	= 0.5164
Cost	= K£20.2490		

(b) After two years

(i) Target transition rates $\underline{p}(\tau+2)_{TARGET} = \underline{a}_2$

.6998, .7893, 8436, .8296, .8741, .7410, .8871, .1158, .1033,

(ii) Theoretical values attained $\hat{\underline{p}}(\tau+2)$

.6880, .8008, .8251, .8142, .8460, .7560, .8612, .1151, .1075,

(iii) Regressor variables at time $\tau+2$

Controllable $(\underline{x}^*_{1,\tau+2})$	Uncontrollable $(\hat{\underline{x}}_{2,\tau+2})$
Pupil-teacher ratio = 36.5273 Cost = K£20.8008	Sex ratio = 0.5076

.1085, .1098, .1101, .1378, .1129

.1110, .1136, .1190, .1456, .1388

(c) After <u>three years</u>

(i) Target transition rates $\underline{p}(\tau+3)_{TARGET} = \underline{a}_1$

.7098, .7993, .8536, .8396, .8841, .7510, .8971, .1058, .0933,

(ii) Theoretical values attained $\hat{\underline{p}}(\tau+3)$

.6887, .7896, .8164, .8131, 8445, .7483, .8612, .1169, .1074

(iii) Regressor variables at time $\tau+3$

Controllable ($\underline{x}^*_{1,\tau+3}$)	Uncontrollable ($\hat{\underline{x}}_{2,\tau+3}$)
Pupil-teacher ratio =35.7012 Cost = K£21.6348	Sex ratio = 0.4987

.0985, .0998, .1001, .1278, .1029

, .1114, .1137, .1186, .1457, .1388

CASE III:- *DECREASE IN PROMOTION RATES OF 0,01 PER YEAR WITH A REPEAT RATE INCREASE OF THE* SAME AMOUNT PER YEAR, OVER THE NEXT THREE YEARS.

(a) After one year

(i) Target transition rates $\underline{p}(\tau+1)_{TARGET} = \underline{a}_1$

.6698, .7593, .8136, .7996, .8441, .7110, .8571, .1458, .1333, .1385, .1398, .1401, .1678, .1429

(ii) Theoretical value attained $\hat{\underline{p}}(\tau+1)$

.6929, .8198, .8386, .8199, .8509, .7716, .8642, .1097, .1046, .1074, .1105, .1167, .1428, .1358

(iii) Regressor variables at time $\tau+1$

Controllable ($\underline{x}^*_{1,\tau+1}$)	Uncontrollable ($\hat{\underline{x}}_{2,\tau+1}$)
Pupil-teacher ratio =37.2813	Sex ratio = 0.5164
Cost = K£20.0381	

(b) *After <u>two years</u>.*

(i) Target transition rates $\underline{p}(\tau+2)_{TARGET} = \underline{a}_2$

.6598, .7493, .8036, .7896, .8341, .7010, .8471, .1558, .1433,

(ii) Theoretical value attained $\hat{\underline{p}}(\tau+2)$

.6691, .7912, .8208, .8017, .8369, .7414, .8525, .1198, .1129

(iii) Regressor variables at time $\tau+2$

Controllable ($\underline{x}^*_{1,\tau+2}$)	Uncontrollable ($\hat{\underline{x}}_{2,\tau+2}$)
Pupil-teacher ratio =37.2422 Cost = K£20.085	Sex ratio = 0.5076

.1485, .1498, .1501, .1778, .1529

, .1164, .1190, .1239, .1503, .1475

(c) *After three years*

(i) Target transition rates $\underline{p}(\tau+3)_{TARGET} = \underline{a}_3$

.6498, .7393, .7936, .7796, .8241, .6910, .8371, .1658, .1533,

(ii) Theoretical value attained $\hat{\underline{p}}(\tau+3)$

.6509, .7686, .8052, .7868, .8246, .7183, .8419, .1290, .1208

(iii) Regressor variables at time $\tau+3$

Controllable $(\underline{x}^*_{1,\tau+3})$	Uncontrollable $(\hat{\underline{x}}_{2,\tau+3})$
Pupil-teacher ratio = 37.0996 Cost = K£20.2363	Sex ratio = 0.4987

.1585, .1598, .1601, .1878, .1629

, .1246, .1269, .1311, .1577, .1581

Comments on controlling the system

Before giving any comments on the results we need to remember that, for suggested target transitions, we are looking for regressor values which will attain the target optimally. The regressors are partitioned into controllable and non-controllable variables. The controllable variables are computed using the multivariate control model while the uncontrollable ones are projected using previous observations.

We begin with case one, where during the next three years we intend to maintain fixed transition rates. If we are to maintain the fixed transition rates after one year then the pupil-teacher ratio will have to be reduced from 38.3122 to 37.1758. On the other hand, the cost of education per year should be increased from K£19 to K£20.1436. The uncontrollable variable, that is the sex ratio, is 0.5164 as estimated from past trends. After two years if we are still to maintain the fixed transition rates, the pupil-teacher ratio should be reduced further to 36.8818 and the cost increased further to K£20.4473. The sex ratio is now 0.5076, as estimated from past trends. If we still intend to maintain these fixed transition rates after three years, the pupil-teacher ratio needs to be reduced even further to 36.3867, and the cost increased further to K£20.9512. The sex ratio is now estimated at 0.4987 based on previous trends.

Next we consider case two where we require an increase of 0.01 in the promotion rates per year and a decrease of the same amount in the repeat rates, for the next three years. If we are to increase promotion rates by 0.01 and reduce repeat rates by the same amount after one year, we shall need to decrease the pupil-teacher ratio from 38.3122 to 37.0703. On the other hand, the cost of education should also increase from K£19 to K£20.2490. If during the second period we further wish to increase the promotion rates by 0.01 and decrease the repeat rates by a further 0.01, then the pupil-teacher ratio must be reduced further to 36.5273. On the other hand the cost also needs to be increased further in the second year to K£20.8008. In the third year if we wish to increase promotion rates further by 0.01 and decrease repeat rates further by 0.01 then, the pupil-teacher ratio needs to be reduced even further to 35.7012. The cost of education will also require to be increased further to a value of K£21.6348. The sex ratios are as in case one above.

Finally we consider case three where we require a decrease of 0.01 in the promotion rates and an increase of the same amount in the repeat rates, during the next three years. Under case three, during the first year we need to reduce the pupil-teacher ratio from 38.3122 to 37.2813. On the other hand the cost should increase from K£19 to K£20.0381. During the second year under case three adjustments in the transition ratios, the pupil-teacher ratio needs to be reduced slightly to 37.2422 while the cost should be increased to K£20.085. In the third year of control, the pupil-teacher ratio needs to be decreased again to a value of 37.0996 and the cost increased to K£20.2363. The sex ratios are again the same as in case one during the first, second and third years respectively.

5.5 CONCLUSIONS AND COMMENTS

On The Regression Model

These comments are based on Tables 27 - 30. The proportion of students in the system at the end of the first year of schooling is approximately 82%. Approximately 73% of these grade one pupils will still be in school after two years of schooling and so on. In fact according to the regression model, all the grade one pupils will have left primary school after fifteen years. Similar percentages are obtainable for pupils in any of the other school grades at the initial time. For example, approximately 89% of the grade two pupils will still be in school after one year. After two years of schooling about 85% of the grade two pupils will still be in school and so on. On the other hand for pupils in grade seven, at the initial time, about 13% will still be in school after one year. After two years of schooling only about 2% of the grade seven pupils will still be in school and so on. In fact practically all the grade seven pupils will have left primary school after five years.

According to the regression model, approximately 18% of the pupils in grade one at time zero drop out within one year of schooling. Within two years of schooling, approximately 26% of these pupils will have dropped out of primary school and so on. We can obtain drop out rates for pupils in any of the other school grades at time zero. For example about 11% of those in grade two at the initial time will have dropped out within one year. Within two years approximately 15% will have dropped out of school and so on. On the other hand if we consider those in grade seven at the initial time then about 87% will have dropped out of primary school within one year. Within two years of schooling 98% of the grade seven pupils will have left primary school and so on. In fact as before, all the grade seven pupils will have left primary school after five years.

A pupil in grade one at the initial time spends an average of about 5 years 7 months in primary school, according to the regression model. On the other hand a pupil in grade two at the initial time is expected to take 5 years 1 month in primary school and so on. We can obtain the survival times for pupils in any of the other school grades. For instance, a pupil in grade seven is expected to take approximately 1 year 2 months in primary school before leaving the system. On the other hand, a pupil in grade one at time zero has approximately a 52% chance of ever reaching the highest primary school grade. A pupil in grade two at this time has approximately a 65% chance of ever reaching grade seven and so on. We can obtain similar probabilities for pupils in any of the other grades at the initial time. For example a grade six pupil has a chance of about 86% of reaching grade seven.

Any pupil in the system at the initial time is expected to take an average of 3 months in grade one. These expected lengths of stay in the grades increase with grade to a maximum of about 10 months in grade seven. In general a student in primary school at time zero is expected to take approximately 4 years 3 months in primary school before leaving the system. This value gives a general measure of the time spent in primary school by any of the pupils in the system at time zero.

On The Regression Control Model

In controlling the system, for suggested target transition ratios, we seek values of the regressor variables which will help in attaining the suggested targets optimally.

In all the three cases of alternative targets used, it was found that, after the first year, the pupil-teacher ratio needs to be lowered in order to achieve the target optimally. This is possibly due to the high pupil-teacher ratio at the starting time τ. It is, however, noticed that the case of increasing promotions rates by 0.01 and lowering repeat rates by the same amount requires a smaller number of pupils per teacher among all the cases. This seems to imply that a low pupil-teacher ratio may lead to a better pupil performance, in that more pupils will be promoted and less will repeat any given grade. On the other hand the cost shows an increasing trend in all the cases despite the control action. We note however that the rate of increase in cost is highest for the case of increased promotions and decreased repeats sited above as compared to the other two cases. This may possibly be a consequence of the low pupil-teacher ratio caused by the control action.

General Conclusions

The main advantage of using the linear regression model over the direct time dependent models discussed in chapter III is that, it takes into account factors which affect the transitions in the education system. In addition, which ever assumption on the transition ratios we opt for, this method gives a plan of action to be taken on the controllable variables so as to achieve the desired goals optimally.

It is important to note that the problem of finding multiple period optimal solutions without prior fixed intermediate targets can be quite complex. In fact the two period case appears not to have a closed form solution.

Finally we point out that the suggested model need not only be applied to the primary education system. It may in fact be applied to the entire education system in a country. We may even use more variables to describe the transition process provided, with the available data, we can adequately derive the predictive distribution. In the present case the predictive distribution had $\nu = \tau - s-(m-1)$ degrees of freedom where, $\tau = 17$ and $m = 14$. The condition for the distribution to be adequately derived is that the degrees of freedom should be greater than zero, i.e. $(\nu > 0)$. This condition can easily be seen to imply that the number of regressor variables to be used must be at most three, i.e. $(s \leq 3)$.

CHAPTER VI

ATTAINABILITY AND MAINTAINABILITY OF GRADE STRUCTURES

6.1 INTRODUCTION

We have so far considered the effects of transition probabilities on the changing education structure. In particular the study is simplified when the transition process has stablized, in which case we would be dealing with constant transition probabilities. Some transition probabilities of a Markov system may lead to undesirable consequences. For example, a reasonable looking promotion rate in an education system may lead to growth of certain grades at the expense of others.

In some fields of application it is possible to exert some control over the system via transition rates. In fact manpower and education planning are largely concerned with how to construct and operate a system so as to meet specified targets. Such an exercise will usually begin with a forecast of what will happen if the present trend continues. The projected structure will rarely coincide with what is desired and so the question arises as to what can be done to alter things. It is at this stage that the need for a control in the transition process arises.

Previously we had transition probabilities and we used them to calculate educational characteristics. Now suppose the desired educational characteristics are specified, and that the problem is to find the transition process to be followed in order to achieve them. In practice only some flows can be controlled, and even then there may be limits to the degree of control which can be exercised.

6.2 CONTROLLING TRANSITIONS

The objective of controlling transitions varies. For example, we may wish to operate the system, subject to a fixed budget, and certain restrictions on grade sizes. Many such objectives can be expressed in terms of the expected stock vector $\underline{n}(T)$, for some future time T. We shall assume, there is some desired grade structure or sequence of structures which we wish to attain. In practice such structures may not be specified precisely. It will normally be sufficient if the actual structure is reasonably close to the target. In any event random variations will cause minor fluctuations even if the target is attained. It therefore seems more reasonable to specify our goal in the form of <u>acceptable structures</u> rather than as a fixed structure, say $\underline{n}^* = (n_1^*, n_2^*, \ldots, n_s^*)'$.

Here 's' is the total number of grades in the system.

The control in transitions may be exercised through flows which may be categorised as follows:

(*i*) *Promotion flows:* These include promotions, demotions and repeats represented by the transition matrix $P = (p_{ij})$, assuming a stabilized system and $i,j = 1,2,\ldots,s$.

(*ii*) *Recruitment flows:* These include total number of recruits at time T, denoted by M(T) and the manner of allocation of recruits through the hierarchies called the recruitment vector and denoted by $\underline{r}' = (r_1, r_2, \ldots, r_s)$; $0 \le r_i \le 1$

(*iii*) *Wastage flows:* These give the rate of loss of members of the system. We represent it by a vector $\underline{w}' = (w_1, w_2, \ldots, w_s)$.

We note that in the model, wastage and promotions are related since

$$\sum_{j=1}^{s} p_{ij} + w_i = 1 \qquad \text{for } i = 1,2,\ldots,s.$$

Of all these flows, wastage flows are least amenable to control. In fact promotion and recruitment flows can be controlled directly by management. In all, recruitment seems to provide the most effective means of control since decisions to recruit more or fewer people at a given time has no immediate impact on those already in any given hierarchical organization. However, control through wastage can at best be indirectly brought about through choice of promotion rates. If these promotion flows are too small, many persons are likely to leave the system out of frustration. We thus control the system through promotion and recruitment. We still note that even between these two a more precise control can be over recruitment since promotion, if not carefully managed will affect morale and general efficiency of the system.

Controlling flows in a system has two classifications, namely:

(*i*) <u>*Attainability*</u>: This is concerned with whether or not a desired goal can be reached and if so, by which means.

(*ii*) <u>*Maintainability*</u>: Which has to do with remaining at the goal structure once it has been attained.

Although the question of attainability comes prior to maintainability, in many cases it is preferrable to treat them in the reverse order. This is because there is usually little or no point in trying to attain a structure which is not maintainable. As such the solution to the maintainability problem partially solves that of attainability as well.

6.3 MAINTAINABILITY

Suppose $\underline{n}^* = \left(n_1^*, n_2^*, \ldots, n_s^*\right)'$ is the grade structure which we wish to maintain, then there must exist flow parameters P , $\underline{w}$ and $\underline{r}$ such that

$$\underline{n}^{*'} = \underline{n}^{*'} P + \underline{n}^{*'} \underline{w}\, \underline{r}' \tag{6.1}$$

This follows since if the structure $\underline{n}^*$ is to be maintained the total size is necessarily fixed. In this case the number of new recruits is then

$$M(T) = \underline{n}^{*'} \underline{w} .$$

Recruitment control therefore implies that P and $\underline{w}$ are fixed and the problem is to find a feasible recruitment vector $\underline{r}$ which satisfies (6.1). On the other hand, promotion control implies that $\underline{w}$ and $\underline{r}$ are fixed and we find a feasible promotion flow matrix, P, which satisfies condition (6.1).

There are many reasons why we may wish to maintain a given grade structure over a period of time. For example, the question may arise, as part of a planning exercise and it is required to know what promotion and recruitment policies are compartible with this structure. Not every structure can be maintained and an important part of the investigation is to delineate those structures which can be maintained from those which cannot. Suppose our primary interest concerning the grade structures is on the proportions enrolled in the various grades at any given time. The proportion of students enrolled in grade i at time t, denoted $q_i(t)$ is given by

$$q_i(t) = n_i(T)/N(T), \quad i = 1,2,\ldots,s \tag{6.2}$$

where $n_i(T)$ is the number of students enrolled in grade i at time T and N(T) is the total enrollment in the system at time T. That is

$$N(T) = \sum_{i=1}^{s} n_i(T).$$

The basic difference equation for the Markov model is then given by

$$\underline{q}'(T+1) = \underline{q}'(T)P + \underline{q}'(T)\underline{w}\,\underline{r}' \tag{6.3}$$

where,

$$\underline{q}'(T) = \left(q_1(T), q_2(T), \ldots, q_s(T)\right).$$

If the relative grade sizes are to remain constant, then we have

$$\underline{q}(T+1) = \underline{q}(T) = \underline{q}$$

and we are interested in flow parameters which satisfy

$$\underline{q}' = \underline{q}'P + \underline{q}'\underline{w}\,\underline{r}' \tag{6.4}$$

Recruitment Control

A grade structure $\underline{q}$ can be maintained by recruitment if we can find a recruitment parameter $\underline{r}$ satisfying equation (6.4) above. From this equation we have that

$$\underline{r}' = \underline{q}'(I - P)\,/\,\underline{q}'\underline{w} \tag{6.5}$$

We note here that although the elements of $\underline{r}'$ add to one, they may not all be positive. If they are all positive then (6.5) gives the unique policy meeting the requirements. Otherwise the structure $\underline{q}$ is not maintainable. From (6.5) it follows that the maintainable region is the set of $\underline{q}$'s for which

$$\underline{q}' \geq \underline{q}'P \tag{6.6}$$

Since $\underline{q} = (q_1, q_2, \ldots, q_s)'$ with $\sum_{i=1}^{s} q_i = 1$ and all q_i's ≥ 0, it follows that all the feasible structures $\underline{q}$ must lie on the hyperplane

$$\sum_{i=1}^{s} q_i = 1.$$

Now from equation (6.4) we have

$$\underline{q}'(I - P) = \underline{q}'\underline{w}\ \underline{r}' \tag{6.7}$$

implying that

$$\underline{q}' = \underline{q}'\underline{w}\ \underline{r}'(I - P)^{-1} \tag{6.8}$$

We may write $\underline{r} = \sum_{1}^{s} r_i \underline{e}_i$ where $\underline{e}_i$ is a column vector of zeros with one in the i-th place. Substituting for $\underline{r}$ in equation (6.4) we have

$$\underline{q}' = \underline{q}'\underline{w} \sum_{1}^{s} r_i \{\underline{e}_i'(I - P)^{-1}\} \tag{6.9}$$

Post multiplying both sides by an sx1 column vector of ones we have

$$1 = \underline{q}'\underline{w} \sum_{1}^{s} r_i \ell_i \tag{6.10}$$

where ℓ_i is the sum of the elements of the i-th row of $(I - P)^{-1}$. Therefore

$$\underline{q}'\underline{w} = 1 \Big/ \sum_{1}^{s} r_i \ell_i \tag{6.11}$$

Substituting for $\underline{q}'\underline{w}$ in equation (6.9) we get

$$\underline{q}' = \sum_{i=1}^{s} \{r_i \ell_i \Big/ \Big(\sum_1^s r_i \ell_i\Big)\} (1/\ell_i) \underline{e}_i' (I - P)^{-1} \tag{6.12}$$

Equation (6.12) reveals that the maintainable $\underline{q}$'s by recruitment are just convex combinations of the points of the type $\ell_i^{-1}\ \underline{e}_i'\ (I - P)^{-1}$, $i = 1,2,\ldots,s$. The vertices of the set of maintainable structures are easily computed by taking the rows of $(I - P)^{-1}$ and scaling the entries so the rows sum to one. We note here that the first row of $(I - P)^{-1}$ is of particular interest because it gives the structure which can be maintained by recruitment into the bottom grade in hierarchical organizations without demotions.

Suppose further that we allow the total system size to change in time at some rate say α. The total size of the system at time $T + 1$ is then

$$N(T + 1) = (1 + \alpha)\ N(T)$$

and the basic equation for an organization of fixed size is now

$$\underline{n}'(T+1) = \underline{n}'(T)P + \underline{n}'(T)\underline{w}\ \underline{r}' + \alpha N(T)\ \underline{r}' \tag{6.13}$$

Introducing the relative grade sizes

$$\underline{q}(T) = \underline{n}(T)/N(T)$$

then

$$(1+\alpha)\underline{q}'(T+1) = \underline{q}'(T)P + \underline{q}'(T)\underline{w}\ \underline{r}' + \alpha\ \underline{r}' \qquad (6.14)$$

The condition for maintaining the relative grade sizes to say $\underline{q}$, by recruitment, is then to find $\underline{r} \geq \underline{0}$ such that

$$(1+\alpha)\underline{q}' = \underline{q}'P + \underline{q}'\underline{w}\ \underline{r}' + \alpha\,\underline{r}' \qquad (6.15)$$

with $\sum_1^s r_i = 1$. If such an $\underline{r}$ exists it is given by

$$\underline{r}' = \{\underline{q}'(I - P) + \alpha\underline{q}'\}\big/(\underline{q}'\underline{w} + \alpha) \qquad (6.16)$$

It is clear that any $\underline{q}$ for which $\underline{r}$ given by (6.5) is non-negative will also yield a non-negative $\underline{r}$ in (6.16). This implies that if the system is expanding, the set of maintainable structures will increase. The converse is also true for $\alpha < 0$.

Following a similar procedure as in the case of a fixed size system it can be shown that the maintainable $\underline{q}$'s by recruitment, under expansion or contraction, are just the convex combination of vectors with coordinates proportional to $\underline{e}_i'\left[(1+\alpha)I - P\right]^{-1}$, $i = 1,2,\ldots,s$. That is

$$\underline{q}' = \sum_1^s K_i\underline{e}_i'\left[(1+\alpha)I - P\right]^{-1}\Big/\ell_i^* \qquad (6.17)$$

The constants K_i's are obtained subject to the condition that $\sum_{1}^{s} q_i = 1$ and $\sum_{1}^{s} K_i = 1$. Or simply, the vertices of the set of maintainable structures are computed by taking the rows of $[(1+\alpha)I - P]^{-1}$ and scaling their entries so that their rows sum to one. Here ℓ_i^* is the row sum of $[(1+\alpha)I - P]^{-1}$

Promotion Control

We now consider control of the system flows through promotion. For a fixed size organization the problem is to find promotion flow parameters P which satisfy equation (6.4). Here to delineate the set of maintainable structures we must determine the set of $\underline{q}$'s for which there exists an admissible P with

(i) p_{ij}'s ≥ 0 ,(non negativity condition)

(ii) $\sum_{1}^{s} p_{ij} = 1 - w_i$, $i = 1,2,\ldots,s$

From equation (6.4) we have

$$\underline{q}'P = \underline{q}' - \underline{q}'\underline{w}\,\underline{r}' \tag{6.18}$$

We then readily see that for $\underline{q}$ to be maintained we should have

$$\underline{q}' \geq \underline{q}'\underline{w}\,\underline{r} \tag{6.19}$$

We note that although P does not have a unique value, it is extremely rare for every manner of transition to be possible in an organization. In particular for hierarchical systems without demotions P is upper triangular. When P is super diagonal, for example in an educational system, then (6.18) can be written as follows:

$$\left.\begin{aligned} q_1p_{11} &= q_1 - r_1\sum_1^s q_iw_i \\ q_1p_{12} + q_2p_{22} &= q_2 - r_2\sum_1^s q_iw_i \\ \cdots\cdots &\cdots\cdots \\ q_{s-1}p_{s-1,s} + q_sp_{ss} &= q_s - r_s\sum_1^s q_iw_i \end{aligned}\right\} \qquad (6.20)$$

Substituting $p_{ii} = 1 - w_i - p_{i,i+1}$ for $i = 1,2,\ldots,s-1$ in (6.20) we obtain, and solving for $p_{i,i+1}$

$$p_{i,i+1} = \sum_{k=1}^{i} r_k\left(\sum_1^s q_jw_j\right)\Big/q_i - \left(\sum_1^i q_jw_j\right)\Big/q_i\,, \qquad i = 1,2,\ldots,s-1 \qquad (6.21)$$

The structure of the system, $\underline{q}$, is then maintainable by promotion if

$$0 \leq p_{i,i+1} \leq 1 - w_i$$

or if,

$$0 \leq \sum_{k=1}^{i} r_k\left(\sum_1^s q_jw_j\right) - \sum_1^i q_jw_j \leq q_i(1 - w_i) \qquad (6.22)$$

for $i = 1,2,\ldots,s-1$. As a special case, when recruitment is only into the lowest grade then $r_1 = 1$ and $r_k = 0$ for $k \neq 1$, in which case we have

$$p_{i,i+1} = \left(\sum_{j=i+1}^{s} q_j w_j\right) \Big/ q_i ; \quad i = 1,2,\ldots,s-1 \tag{6.23}$$

These are the promotion rates required to maintain the grade structure

$$\underline{q} = (q_1, q_2, \ldots, q_s)' .$$

Suppose we further allow the total size of the system to change in time at some rate α, say. The basic equation for an organization of fixed size is then given by equation (6.13). The condition for maintaining the relative grade size $\underline{q}$ by promotion is then to find a feasible promotion parameter P such that

$$(1+\alpha)\underline{q}' = \underline{q}'P + (\underline{q}'\underline{w} + \alpha)\underline{r}' \tag{6.24}$$

From which we have

$$\underline{q}'P = (1+\alpha)\underline{q}' - (\underline{q}'\underline{w} + \alpha)\underline{r}' \tag{6.25}$$

Thus the set of maintainable structures, $\underline{q}$, by promotion is defined by

$$\underline{q}' \geq (\underline{q}'\underline{w} + \alpha)\underline{r}'/(1+\alpha) \tag{6.26}$$

for general P. For super diagonal P, for example in an educational system, the elements of P must satisfy

$$q_1 p_{11} = q_1(1+\alpha) - r_1(\sum_1^s q_i w_i + \alpha)$$

and

$$q_i p_{i,i+1} + q_{i+1} p_{i+1,i+1} = q_{i+1}(1+\alpha) - r_{i+1}(\sum_1^s q_i w_i + \alpha)$$

$$(i = 1,2,\ldots,s-1) \quad (6.27)$$

Solving for $p_{i,i+1}$ and making use of the inequalities

$$0 \le p_{i,i+1} \le 1 - w_i$$

we obtain,

$$p_{i,i+1} = \sum_{k=1}^{i} r_k \{\sum_1^s q_j(w_j + \alpha)\}/q_i - \{\sum_1^i q_j(w_j + \alpha)\}/q_i$$

$$(i = 1,2,\ldots,s-1) \quad (6.28)$$

and the condition for a structure $\underline{q}$ to be maintainable is

$$0 \le \sum_{k=1}^{i} r_k \{\sum_1^s q_j(w_j + \alpha)\} - \sum_1^i q_j(w_j + \alpha) \le q_i(1 - w_i)$$

$$(i = 1,2,\ldots,s-1) \quad (6.29)$$

If recruitment is only into the lowest grade then we have

$$p_{i,i+1} = \{\sum_{j=i+1}^{s} q_j(w_j + \alpha)\}/q_i \quad (6.30)$$

The condition for maintainability of a structure $\underline{q}$ is then given by

$$0 \leq \sum_{j=i+1}^{s} q_j\left(w_j + \alpha\right) \leq q_i\left(1 - w_i\right) \qquad (6.31)$$

$$(i = 1,2,\ldots,s-1)$$

In order to obtain the set of maintainable structures, $\underline{q}$, we have to consider the $(s-1)$ inequalities in (6.31) and the additional condition

$$\sum_{1}^{s} q_i = 1 \qquad (6.32)$$

The set of maintainable structures is therefore the set of non-negative q_i's which satisfy the 's' equations

$$\begin{aligned}
q_1 + q_2 + \cdots + q_s &= 1 \\
q_2\left(w_2+\alpha\right) + q_3\left(w_3+\alpha\right) + \cdots + q_s\left(w_s+\alpha\right) &= \left(1-w_1-\delta_1\right)q_1 \\
q_3\left(w_3+\alpha\right) + \cdots + q_s\left(w_s+\alpha\right) &= \left(1-w_2-\delta_2\right)q_2 \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
q_s\left(w_s+\alpha\right) &= \left(1-w_{s-1}-\delta_{s-1}\right)q_{s-1}
\end{aligned} \qquad (6.33)$$

where $\delta_1, \delta_2, \ldots, \delta_{s-1}$ are non-negative proportions. It is easy to see that $\delta_i = p_{ii}$ for $i = 1,2,\ldots,s-1$. This follows because, for a hierachical organization with s states and promotion only to the next higher grade, we have

$$p_{i,i+1} = \begin{cases} 1 - w_i - p_{ii} , & i = 1,2,\ldots,s-1 \\ w_i , & i = s \end{cases}$$

The set of equations in (6.33) can now be written in matrix notation as

$$\underset{s\times s}{W} \; \underset{s\times 1}{\underline{q}} = \underset{s\times 1}{\underline{e}} \tag{6.34}$$

where,

$$W = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ -(1-w_1-\delta_1) & (w_2+\alpha) & (w_3+\alpha) & \cdots & (w_{s-1}+\alpha) & (w_s+\alpha) \\ 0 & -(1-w_2-\delta_2) & (w_3+\alpha) & \cdots & (w_{s-1}+\alpha) & (w_s+\alpha) \\ & & & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & & -(1-w_{s-1}-\delta_{s-1}) & (w_s+\alpha) \end{bmatrix}$$

$$\underline{q} = (q_1, q_2, \ldots, q_s)'$$

and

$$\underline{e} = (1,0,0,\ldots,0)'$$

For varying repeat rates δ_i's and change rate α we can obtain the maintainable structures as the solution of the system given by (6.34) where all

q_i's are non negative. If no feasible solution can be obtained for a given set of δ_i's and a change rate α then no structure is maintainable under these conditions.

In Table 31 below we give the maintainable grade structures for the Kenyan primary education system as the system growth rate α varies, and for some controlled repeat rates.

Table 31: Maintainable grade structures under varying growth rates α with controlled repeat rates.

Growth rate α	Repeat rates p_{ii} $(i=1,2,\ldots,6)$	q_1	q_2	q_3	q_4	q_5	q_6	q_7
$\alpha \leq -1$	0 .09	NO MAINTAINABLE STRUCTURES Since everyone leaves the system after one year.						
$\alpha=-0.8$	0	0	.0001	.0005	.0024	.0110	.0523	.9337
	.09	0	0	.0001	.0010	.0075	.0578	.9336
$\alpha=-0.4$	0	.0258	.0362	.0566	.0886	.1359	.2145	.4424
	.09	.0210	.0310	.0514	.0857	.1395	.2345	.4369
$\alpha=-0.2$	0	.0893	.0941	.1101	.1294	.1489	.1762	.2520
	.09	.0869	.0922	.1099	.1316	.1538	.1857	.2399
$\alpha=0$	0	.1847	.1558	.1459	.1371	.1262	.1195	.1308
	.09	.1922	.1592	.1480	.1383	.1261	.1188	.1174
$\alpha=0.2$	0	.2857	.2008	.1566	.1227	.0941	.0743	.0658
	.09	.3030	.2057	.1568	.1200	.0898	.0692	.0555
$\alpha=0.4$	0	.3755	.2262	.1513	.1015	.0667	.0452	.0336
	.09	.3988	.2294	.1481	.0961	.0609	.0399	.0268
$\alpha=0.8$	0	.5107	.2393	.1244	.0650	.0332	.0175	.0099
	.09	.5378	.2370	.1172	.0583	.0283	.0142	.0072
$\alpha=1$	0	.5602	.2363	.1106	.0520	.0239	.0113	.0057
	.09	.5872	.2317	.1026	.0457	.0198	.0089	.0041
$\alpha=1.2$	0	.6011	.2305	.0981	.0419	.0175	.0075	.0034
	.09	.6275	.2241	.0898	.0362	.0142	.0058	.0024
$\alpha=1.4$	0	.6352	.2232	.0871	.0341	.0131	.0052	.0021
	.09	.6607	.2156	.0789	.0290	.0105	.0039	.0014
$\alpha=1.8$	0	.6887	.2074	.0694	.0233	.0077	.0026	.0009
	.09	.7123	.1981	.0618	.0194	.0059	.0019	.0006

Comments on Table 31

In obtaining the maintainable structure for the Kenyan primary education system we have used the average wastage vector between 1963-1980. This wastage vector is given for grades 1,2,...,7 as

$$\underline{w}' = (0.1564, 0.0640, 0.0603, 0.0795, 0.0528, 0.0674, 0.8522)$$

For purposes of illustration we have used two repeat rates for grades 1,2,...,6; one corresponding to no repeat so that $p_{ii} = 0$ and the other corresponding to a repeat rate of 0.09, that is $p_{ii} = 0.09$.

It is observed that if the system annual growth rate $\alpha \leq -1$ then no grade structure is maintainable. This is obviously because for a growth rate of less than or equal to negative one, all pupils are expected to leave the system after one year of schooling. If the system declines at a rate of 0.8 ($\alpha = -0.8$) then with no repeats we can maintain a structure which is very top heavy; for example no student will be expected in grade one in this structure while 0.9337 of the students will be expected in grade seven. For a system with 0.09 repeat rate we can still maintain a top heavy structure with no student in grade one while 0.9336 of the students will be expected in grade seven. If the growth rate is now $\alpha = -0.4$ the system which can be maintained becomes less top heavy than in the case of $\alpha = -0.8$. For example, for the no repeat case, we expect 0.0258 of the students to be in grade one while we now expect only 0.4424 of the students to be in grade seven. For the case of 0.09 repeat rate, we expect 0.0210 of the students to be in grade one while we expect 0.4369 of them to be in grade seven.

For a fixed size system, that is $\alpha = 0$, a nearly uniform grade structure is maintainable. For example, for the no repeats case, we expect 0.1847 of the pupils to be in grade one while we expect 0.1308 to be in grade seven. For the case of 0.09 repeat rate, we expect 0.1922 to be in grade one and 0.1174 to be in grade seven. For a system which doubles in size after each year, that is $\alpha = 1$, the maintainable structure is a bottom heavy type. For example, in the no repeat case, we expect 0.5602 of the students to be in grade one while we now expect only 0.0057 of the pupils to be in grade seven. For the case of a 0.09 repeat rate, we expect 0.5872 of the pupils to be in grade one and only 0.0041 to be in grade seven.

We note that as the growth rate 'α' increases from negative to positive, the maintainable grade structure changes from top heavy to bottom heavy. We also note that the case of no repeats is generally more top heavy than that of a 0.09 repeat rate.

6.4 ATTAINABILITY

Attainability is a more complex concept than maintainability. This is because, for example, a structure which is attainable from one starting point may not be attainable from another starting point. Furthermore the number of steps required for the change and the route followed may be subject to variations. We note again that there is little or no point in trying to attain a structure which is not maintainable. It follows that, provided we equate attainability of a structure with getting arbitrarily close to the structure, then any maintainable structure can be attained from any starting structure.

We illustrate the concept of attainability of a maintainable structure by means of recruitment control. The argument for promotion control is similar. It is sufficient to do this in the case where the total size of the system is fixed. Suppose we start from a structure $\underline{q}(0)$ and ask what would happen if we repeatedly used a recruitment vector $\underline{r}^*$ which maintains a structure $\underline{q}^*$. We know that the limiting structure in this case satisfies

$$\underline{q}' = \underline{q}'P + \underline{q}'\underline{w}\,\underline{r}^{*\prime} \qquad (6.35)$$

But by definition $\underline{r}^*$ also satisfies

$$\underline{q}^* = \underline{q}^{*\prime}P + \underline{q}^{*\prime}\underline{w}\,\underline{r}^{*\prime} \tag{6.36}$$

implying that $\underline{q} = \underline{q}^*$. Such an $\underline{r}^*$ only exists when $\underline{q}^*$ is in the set of maintainable structures. The above argument tells us to recruit in the same fixed proportions as if we were already at the target and wished to remain there. We may use some optimallity criterion for attainability which may be defined in terms of time and cost. We shall use two approaches to this problem:

(i) Free Time Attainability:- This aims at first determining the smallest number of steps in which the target can be attained and then choosing among the possible strategies on the basis of some objective function.

(ii) Fixed Time Attainability: Aims at getting as close as possible to the target in a prescribed number of steps.

FREE TIME ATTAINABILITY

We shall illustrate free time attainability of a maintainable structure by first controlling recruitment and then by controlling promotions.

Free Time Attainability by Recruitment

In formulating the problem in this case we shall work in terms of the enrollment numbers rather than proportions. We note first that the enrollment numbers at time (T + 1) are given by

$$\underline{n}'(T+1) = \underline{n}'(T)P + \underline{f}'(T+1) \tag{6.37}$$

where,

$$\underline{f}(T+1) = M(T+1)\ \underline{r}$$

is the recruitment vector at time (T+1) and M(T+1) is the new recruitment at time (T+1). We assume that the total system size N(T) is known or may be obtained via the growth rate of the system, 'α'.

The problem is now to find T* and a sequence of feasible recruitment vectors $\underline{f}(1), \underline{f}(2), \ldots, \underline{f}(T^*)$ such that T* is the smallest value of T for which $\underline{n}(T) = \underline{n}^*$, the target structure. The constraints imposed on the unknowns are

$$\underline{n}'(T+1) = \underline{n}'(T)P + \underline{f}'(T+1) \quad (T = 0,1,2,\ldots,T^*-2) \tag{i}$$

$$\underline{n}^{*\prime} = \underline{n}'(T^*-1)P + \underline{f}(T^*) \tag{ii}$$

$$\sum_1^s n_i(T) = N(T) \quad (T = 1,2,\ldots,T^*-1) \tag{iii}$$

$$\underline{n}(T) \geq \underline{0} \quad (T = 1,2,\ldots,T^*-1) \tag{iv}$$

$$\underline{f}(T) \geq \underline{0} \qquad (T = 1,2,\ldots,T^*) \qquad \text{(v)}$$

(6.38)

These equations are linear in $n_i(T)$'s and $f_i(T)$'s and we may therefore use the standard methods of mathematical programming to find a feasible solution.

Since T^* is unknown we put $T^* = 1,2,\ldots$ in turn until the first value is found for which a feasible solution exists. The solution to (6.38), will not only give the feasible recruitment vectors $\underline{f}(T)$'s which we require but also the intermediate structures $\{\underline{n}(T);\ T=1,2,\ldots,T-1\}$ through which the system passes. When T^* has been found there will normally be many solutions and it is therefore open to select one set by minimizing some function of economic or social interest. One such function may be, for example, the total expenditure. If the average expenditure per student in grade i is c_i then once T^* has been found, the problem then becomes:

minimize the total expenditure $\quad z = \sum_{T=1}^{T^*-1} \sum_{i=1}^{s} c_i\, n_i(T)$

subject to the conditions given in (6.38). When $T^*=1$ the objective is trivial since $\underline{n}(0)$ is already known. We therefore need only solve for the recruitment vector $\underline{f}(1)$ such that

$$\underline{n}^{*\prime} = \underline{n}'(0)P + \underline{f}'(1)$$

where

$$\underline{f}(1) \geq 0 \qquad (6.39)$$

The condition for one step attainability is therefore that for $\underline{f}(1)$ to be feasible, which is

$$\underline{n}^{*\prime} \geq \underline{n}'(0)P \qquad (6.40)$$

Free Time Attainability by Promotion

In this case we find a sequence of promotion matrices, P(T)'s, which when coupled with a given recruitment vector $\underline{r}$, takes the system from an initial grade structure $\underline{n}(0)$ to a target structure $\underline{n}^*$ in the shortest possible time. The basic equation of the system may now be written as

$$\underline{n}'(T+1) = \underline{n}'(T)P(T) + \underline{n}'(T)\underline{w}\,\underline{r}' + M(T+1)\underline{r}' \qquad (6.41)$$

where, M(T+1) is the new recruitment at time (T+1). The unknowns are the vectors $\underline{n}(T)$ and the matrices P(T), but the basic equations are not linear in these variables. If we express all flows as numbers rather than proportions, the restrictions in the basic equation become of linear type. Furthermore it is easier, for solvability of the problem, to temporarily abandon matrix notation.

Let,

$$n_{ij}(T) = n_i(T)\ p_{ij}(T)$$

be the number flowing from grade i to j in the time interval (T,T+1). We have to find T* and values of the variables $n_{ij}(T)$ and $n_i(T)$ $(i = 1,2,\ldots, s)$, such that

$$n_j(T+1) = \sum_{i=1}^{s} n_{ij}(T) + r_j\{\sum_1^s n_i(T)w_i + M(T+1)\} \qquad \text{(i)}$$

$$(j = 1,2,\ldots,s;T = 0,1,\ldots,T^*-2)$$

$$n_j^* = \sum_{i=1}^{s} n_{ij}(T^*-1) + r_j\{\sum_1^s n_i(T^*-1)w_i + M(T^*)\} \qquad \text{(ii)}$$

$$(j = 1,2,\ldots,s)$$

$$\sum_{j=1}^{s} n_{ij}(T) = n_i(T)(1-w_i);\ (i = 1,2,\ldots,s-1;\ T=0,1,\ldots,T^*-1) \qquad \text{(iii)}$$

$$\sum_1^s n_j(T) = N(T),\ (T = 1,2,\ldots,T^*-1) \qquad \text{(iv)}$$

$$\underline{n}(T) \geq \underline{0}\ ,\ (T = 1,2,\ldots,T^*-1) \qquad \text{(v)}$$

$$n_{ij}(T) \geq 0\ ,\ (i,j = 1,2,\ldots,s;\ T = 0,1,\ldots,T^*-1) \qquad (6.42)$$

Again as before once T* has been found there will normally be many solutions and we may, for example, choose the solution which minimizes the average expenditure. The problem then becomes minimize total expenditure, $z = \sum_{T=1}^{T^*-1} \sum_{i=1}^{s} c_i n_i(T)$

subject to the conditions given in (6.42).

Consider in particular a hierarchical organization of 's' grades with new enrollment or recruitment only into the lowest grade and promotion only to the next higher grade. Such a system may be for example, an education system. In this case, the condition for one step attainability, that is $T^* = 1$, is simply that of existence of a feasible solution to the system of the linear equations

$$\text{(s eqs)} \begin{cases} n_1^* = n_{11}(0) + n_1(0)w_1 + \ldots + n_s(0)w_s + M(1) \\ n_2^* = n_{12}(0) + n_{22}(0) \\ \cdots\cdots\cdots \\ n^*_s = n_{s-1,s}(0) + n_{ss}(0) \end{cases}$$

$$\text{(s-1 eqs)} \begin{cases} n_{11}(0) + n_{12}(0) = n_1(0)(1-w_1) \\ n_{22}(0) + n_{23}(0) = n_2(0)(1-w_2) \\ \cdots\cdots\cdots\cdots \\ n_{s-1,s-1}(0) + n_{s-1,s}(0) = n_{s-1}(0)(1-w_{s-1}) \end{cases} \tag{6.43}$$

These equations may be written in matrix notation as

$$A_1 \, \underline{n}_{ij}(0) = \underline{b}_1 \tag{6.44}$$

where,

$$\underset{(2s-1)\times(2s-1)}{A_1} = \left[\begin{array}{ccccccccc}
1 & 0 & 0 & & \cdot & \cdot & \cdot & \cdot & 0 \\
0 & 1 & 1 & 0 & \cdot & \cdot & \cdot & \cdot & 0 \\
0 & 0 & 0 & 1 & 1 & \cdot & \cdot & \cdot & 0 \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & 0 & & & \cdot & \cdot & 1 & 1 \\
\hline
1 & 1 & 0 & 0 & & \cdot & \cdot & 0 & 0 \\
0 & 0 & 1 & 1 & 0 & \cdot & \cdot & 0 & 0 \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & 0 & & & \cdot & 1 & 1 & 0
\end{array}\right]
\begin{array}{l} \left.\vphantom{\begin{array}{c}1\\1\\1\\1\\1\end{array}}\right\} s \\ \left.\vphantom{\begin{array}{c}1\\1\\1\\1\end{array}}\right\} s-1 \end{array}$$

$$\underset{(2s-1)\times 1}{\underline{n}_{ij}(0)} = \left(n_{11}(0), n_{12}(0), n_{22}(0), n_{23}(0), \ldots, n_{s-1,s}(0), n_{ss}(0)\right)'$$

and,

$$\underset{(2s-1)\times 1}{\underline{b}_1'} = \Big(\underbrace{n_1^* - M(1) - \sum_1^s n_j(0) w_j,\ n_2^*, n_3^*, \ldots, n_s^*}_{s} \;\vdots\; \underbrace{n_1(0)(1-w_1), n_2(0)(1-w_2), \ldots, n_{s-1}(0)(1-w_{s-1})}_{s-1}\Big)$$

If the solution obtained from (6.44) has $\underline{n}_{ij}(0) \geq \underline{0}$ then the structure $\underline{n}^* = (n_1^*, n_2^*, \ldots, n_s^*)'$ is attainable in one step.

The condition for two step attainability of a structure $\underline{n}^*$ (that is, $T^*=2$) assuming that the costs per grade are all equal, may be written in expanded form as

$$\text{minimize} \quad z = n_1(1) + n_2(1) + \ldots + n_s(1)$$

subject to the conditions

$$
\begin{array}{ll}
 & n_1(1) = n_{11}(0) + n_1(0)w_1 + \ldots + n_s(0)w_s + M(1) \\
 & n_2(1) = n_{12}(0) + n_{22}(0) \\
\text{s eqs} & n_3(1) = n_{23}(0) + n_{33}(0) \\
 & \ldots\ldots\ldots\ldots \\
 & n_s(1) = n_{s-1,s}(0) + n_{ss}(0) \\
\hline
 & n_{11}(0) + n_{12}(0) = n_1(0)(1-w_1) \\
\text{s-1 eqs} & n_{22}(0) + n_{23}(0) = n_2(0)(1-w_2) \\
 & \ldots\ldots\ldots\ldots \\
 & n_{s-1,s-1}(0) + n_{s-1,s}(0) = n_{s-1}(0)(1-w_{s-1}) \\
\hline
\text{1 eq} & n_1(1) + n_2(1) + \ldots + n_s(1) = N(1) \\
\hline
 & n_1^* = n_{11}(1) + n_1(1)w_1 + n_2(1)w_2 + \ldots + n_7(1)w_7 + M(2) \\
\text{s eqs} & n_2^* = n_{12}(1) + n_{22}(1) \\
 & \ldots\ldots\ldots\ldots \\
 & n_s^* = n_{s-1,s}(1) + n_{ss}(1) \\
\hline
 & n_{11}(1) + n_{12}(1) = n_1(1)(1-w_1) \\
\text{s-1 eqs} & n_{22}(1) + n_{23}(1) = n_2(1)(1-w_2) \\
 & \ldots\ldots\ldots\ldots \\
 & n_{s-1,s-1}(1) + n_{s-1,s}(1) = n_{s-1}(1)(1-w_{s-1})
\end{array}
\qquad (6.45)
$$

The problem may be formulated in matrix notation as follows:

$$\text{minimize} \quad z = \underline{n}'(1)\ \underline{1}$$

subject to,

$$A_2 \quad \times \quad \left[\begin{array}{c} \underline{n}_{ij}(0) \\ \hline \underline{n}(1) \\ \hline \underline{n}_{ij}(1) \end{array}\right] = \underline{b}_2 = \left[\begin{array}{c} \underline{b}_{21} \\ \hline N(1) \\ \hline \underline{b}_{22} \end{array}\right] \qquad (6.46)$$

where,

$$\underset{(4s-1)\times(5s-2)}{A_2} = \begin{array}{cc} \begin{array}{ccc} 2s-1 & \qquad s \qquad & 2s-1 \end{array} & \\ \left[\begin{array}{c|c|c} A_1 & \begin{matrix} -I_s \\ 0_{(s-1)\times s} \end{matrix} & 0 \\ \hline \underline{0}' & \underline{1}' & \underline{0}' \\ \hline 0 & W & A_1 \end{array}\right] & \begin{array}{c} 2s-1 \\ 1 \\ 2s-1 \end{array} \end{array}$$

A_1 is the matrix in (6.44); the vector $\underline{1} = (1,1\ldots, 1)'_{s\times 1}$; the vector $\underline{0} = (0,0,\ldots,0)'_{(2s-1)\times 1}$;

$$\underset{(2s-1)\times s}{W} = \begin{array}{c} \xleftrightarrow{\qquad\qquad s \qquad\qquad} \\ \left[\begin{array}{cccc} w_1 & w_2 & \cdots & w_s \\ 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 \\ \hline (w_1-1) & 0 & \cdots & 0 \\ 0 & (w_2-1) & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & (w_{s-1}-1) & 0 \end{array}\right] \end{array} \begin{array}{l} \\ \left.\begin{array}{c} \\ \\ \\ \\ \end{array}\right\} s \\ \left.\begin{array}{c} \\ \\ \\ \\ \end{array}\right\} s-1 \end{array}$$

and

0 is a matrix of zeros of the appropriate order

The vector of the unknown variables $\left[\underline{n}'_{ij}(0), \underline{n}'(1), \underline{n}'_{ij}(1)\right]$ is such that

$$\underset{(2s-1)\times 1}{\underline{n}_{ij}(0)} = \left(n_{11}(0), n_{12}(0), \ldots, n_{s-1,s}(0), n_{ss}(0)\right)'$$

$$\underset{(s\times 1)}{\underline{n}(1)} = \left(n_1(1),\ n_2(1), \ldots, n_s(1)\right)'$$

and

$$\underset{(2s-1)\times 1}{\underline{n}_{ij}(1)} = \left(n_{11}(1), n_{12}(1), \ldots, n_{s-1,s}(1), n_{ss}(1)\right)'$$

Lastly the vector $\underline{b}_2' = \left(\underline{b}_{21}', N(1), \underline{b}_{22}'\right)$ is such that

$$\underline{b}_{21}' = \Big[\underbrace{-M(1)-\sum_{1}^{s} n_j(0)w_j, 0, \ldots, 0}_{s} \;\Big\vert\; \underbrace{n_1(0)(1-w_1), n_2(0)(1-w_2), \ldots, n_{s-1}(0)(1-w_{s-1})}_{s-1}\Big]$$

$N(1)$ is the total system size at time $T = 1$, and

$$\underline{b}_{22}' = \Big[\underbrace{n_1^* - M(2), n_2^*, n_3^*, \ldots, n_s^*}_{s} \;\Big\vert\; \underbrace{0 \;\; 0, \ldots, 0}_{s-1}\Big]$$

The vector $\underline{n}^* = (n_1^*, n_2^*, \ldots, n_s^*)'$ is the target structure. The condition for two step attainability of the structure $\underline{n}^*$ is that of existence of a feasible solution to equations (6.46).

When $T^* = 3$, the condition for three step attainability of structure $\underline{n}^*$ is that of existence of a feasible solution to the problem

$$\text{minimize} \quad z = \sum_{T=1}^{2} \underline{n}'(T)\,\underline{1}$$

subject to,

$$A_3 \times \begin{bmatrix} \underline{n}_{ij}(0) \\ \hdashline \underline{n}(1) \\ \hdashline \underline{n}_{ij}(1) \\ \hdashline \underline{n}(2) \\ \hdashline \underline{n}_{ij}(2) \end{bmatrix} = \underline{b}_3 = \begin{bmatrix} \underline{b}_{31} \\ \hdashline N(1) \\ \hdashline \underline{b}_{32} \\ \hdashline N(2) \\ \hdashline \underline{b}_{33} \end{bmatrix} \qquad (6.47)$$

where,

$$\underset{(6s-1)\times(8s-3)}{A_3} = \begin{array}{cccc} 5s-2 & s & 2s-1 & \\ \left[\begin{array}{c|c|c} A_2 & \begin{matrix} 0 \\ -I_s \\ 0_{(s-1)\times s} \end{matrix} & 0 \\ \hline \underline{0}' & \underline{1}' & \underline{0}' \\ \hline 0 & W & A_1 \end{array}\right] & & & \begin{matrix} 4s-1 \\ \\ 1 \\ \\ 2s-1 \end{matrix} \end{array}$$

The entries of A_3 are defined as before, with A_2 as the matrix in (6.46). The vector of the unknown variables

$$\left[\underline{n}'_{ij}(0), \underline{n}'(1), \underline{n}'_{ij}(1), \underline{n}'(2), \underline{n}'_{ij}(2)\right]$$

is such that,

$$\underset{(2s-1)\times 1}{\underline{n}_{ij}(T)} = \left(n_{11}(T), n_{12}(T), \ldots, n_{s-1,s}(T), n_{ss}(T)\right)' \qquad (T = 0,1,2)$$

and

$$\underset{s \times 1}{\underline{n}(T)} = \left(n_1(T), n_2(T), \ldots, n_s(T)\right)'$$

$$(T = 1,2)$$

Lastly, the vector $\underline{b}_3' = \left(\underline{b}_{31}', N(1), \underline{b}_{32}', N(2), \underline{b}_{33}'\right)$ is such that

$$\underline{b}_{31} = \Big[\underbrace{-M(1)-\sum_{1}^{s} n_j(0)w_j, 0, \ldots, 0,}_{s} \vdots \underbrace{n_1(0)(1-w_1), n_2(0)(1-w_2), \ldots, n_{s-1}(0)(1-w_{s-1})}_{s-1}\Big]'$$

$$\underline{b}_{32} = \Big[\underbrace{-M(2), 0, 0, \ldots, 0}_{s} \vdots \underbrace{0, 0, \ldots, 0}_{s-1}\Big]'$$

$$\underline{b}_{33} = \Big[\underbrace{n_1^* - M(3), n_2^*, n_3^*, \ldots, n_s^*,}_{s} \vdots \underbrace{0, 0, \ldots, 0}_{s-1}\Big]'$$

and

$N(T)$ is the total system size at time $T = 1,2$.

Generally when $T^* = k \geq 2$, the condition for k-step attainability of a structure $\underline{n}^*$ is that of existence of a feasible solution to the problem

$$\text{minimize} \quad z = \sum_{T=1}^{k-1} \underline{n}'(T)\, \underline{1}$$

subject to,

$$A_k \quad \times \quad \begin{bmatrix} \underline{n}_{ij}(0) \\ \hdashline \underline{n}(1) \\ \hdashline \underline{n}_{ij}(1) \\ \hdashline \underline{n}(2) \\ \hdashline \vdots \\ \hdashline \underline{n}_{ij}(k-2) \\ \hdashline \underline{n}(k-1) \\ \hdashline \underline{n}_{ij}(k-1) \end{bmatrix} = \underline{b}_k = \begin{bmatrix} \underline{b}_{k1} \\ \hdashline N(1) \\ \hdashline \underline{b}_{k2} \\ \hdashline N(2) \\ \hdashline \vdots \\ \hdashline \underline{b}_{k,k-1} \\ \hdashline N(k-1) \\ \hdashline \underline{b}_{kk} \end{bmatrix} \tag{6.48}$$

where,

$$\underset{(2ks-1) \times [k(3s-1)-s]}{A_k} = \begin{array}{cc} \begin{array}{ccc} (k-1)(3s-1)-s & s & 2s-1 \end{array} & \\ \left[\begin{array}{c|c|c} A_{k-1} & \begin{matrix} 0 \\ -I_s \\ 0_{(s-1) \times s} \end{matrix} & 0 \\ \hline \underline{0}' & \underline{1}' & \underline{0}' \\ \hline 0 & W & A_1 \end{array}\right] & \begin{matrix} 2(k-1)s-1 \\ 1 \\ 2s-1 \end{matrix} \end{array}$$

As before A_{k-1} is the matrix of the condition of (k-1)-step attainability (i.e. when $T^* = k-1$). The matrices 0 and $\underline{0}'$ are zero matrices of appropriate order as indicated above. The vector of unknowns is such that

$$\underset{(2s-1)\times 1}{\underline{n}_{ij}(T)} = \left(n_{11}(T), n_{12}(T), \ldots, n_{s-1,s}(T), n_{ss}(T)\right)' \qquad (T = 0,1,2,\ldots,k-1)$$

and

$$\underline{n}(T) = \left(n_1(T), n_2(T), \ldots, n_s(T)\right)' \qquad (T = 1,2,\ldots,k-1)$$

Lastly, the vector $\underline{b}_k$ is such that

$$\underline{b}_{k1} = \Big[\underbrace{-M(1) - \sum_1^s n_j(0)w_j, 0, \ldots, 0,}_{s} \vdots \underbrace{n_1(0)(1-w_1), n_2(0)(1-w_2), \ldots, n_{s-1}(0)(1-w_{s-1})}_{s-1}\Big]'$$

$$\underline{b}_{kj} = \Big[\underbrace{-M(j), 0, \ldots, 0,}_{s} \vdots \underbrace{0, 0, \ldots, 0}_{s-1}\Big]' \qquad (j = 2,3,\ldots,k-1)$$

$$\underline{b}_{kk} = \Big[\underbrace{n_1^* - M(k), n_2^*, n_3^*, \ldots, n_s^*,}_{s} \vdots \underbrace{0, 0, \ldots, 0}_{s-1}\Big]$$

and

N(T) is the total system size at time $T = 1,2,\ldots,k-1$;

We note in particular that the order of the matrix A_k for the Kenyan primary education system where the number of grades is $s = 7$, is $(14k-1) \times (20k-7)$. In this case when $k = 2$, for instance, we are already dealing with a problem of order 27×33. When $k = 3$, the problem is of order 41×53 and so on. We can therefore see the magnitude of the problem as the steps for attainability of the structure increases. For illustrative purposes, we will give the promotion criteria for the Kenyan primary education system, if some of the maintainable structures given earlier in Table 32 are to be attained within one year. We shall additionally give any further effect on such implementation.

Tables 32(a) - 32(g) below give the promotion criteria for attaining various maintainable structures $\underline{q}^*$ in one step and the effective adjustments on the growth rate due to the action.

PROMOTION CRITERIA FOR ATTAINING SOME MAINTAINABLE STRUCTURES IN ONE YEAR

In these Tables α_0 is the possible initial adjustment in the growth rate which will enable the one step attainability of the maintainable structure $\underline{q}^*$

Table 32(a): Promotion criteria for attaining in one year, the maintainable structure

$\underline{q}^* = \left[.1476, .1340, .1366, .1397, .1399, .1445, .1577\right]'$, which corresponds to a growth rate $\alpha = -0.08$.

Promotion criteria	Adjustments in system growth rates leading to the single step attainability of $\underline{q}^*$			
	$\alpha_0 = -.15$	$\alpha_0 = -.16$	....	$\alpha_0 = 2$
p_{11}	.5608	"SINGLE STEP ATTAINABILITY NOT POSSIBLE"		
p_{12}	.2828			
p_{22}	.3154			
p_{23}	.6206			
p_{33}	.0825			
p_{34}	.8572			
p_{44}	.0715			
p_{45}	.8490			
p_{55}	.1845			
p_{56}	.7627			
p_{66}	.3892			
p_{67}	.5434			
p_{77}	.9310			
Average repeat rate	.3621			
Average promotion rate	.5692			

Table 32(b): Promotion criteria for attaining in one year, the maintainable structure

$\underline{q}^* = [.1697, .1470, .1429, .1396, .1333, .1313, .1362]'$, which corresponds to a growth rate $\alpha = -0.04$

Promotion criteria	Adjustments in system growth rate leading to the single step attainability of $\underline{q}^*$			
	$\alpha_0 = -.11$	$\alpha_0 = -.10$	$\alpha_0 = -.09$	$\alpha_0 = -.08, .., \alpha_0 = 2$
p_{11}	.5050	.4616	.4183	"SINGLE STEP ATTAINABILITY NOT POSSIBLE"
p_{12}	.3386	.3820	.4253	
p_{22}	.3424	.2880	.2338	
p_{23}	.5936	.6480	.7024	
p_{33}	.2006	.1289	.0572	
p_{34}	.7391	.8108	.8825	
p_{44}	.2447	.1655	.0863	
p_{45}	.6758	.7550	.8342	
p_{55}	.3663	.2803	.1943	
p_{56}	.5809	.6669	.7529	
p_{66}	.5145	.4265	.3386	
p_{67}	.4181	.5061	.5940	
p_{77}	.9300	.8182	.7065	
Average repeat rate	.4434	.3670	.2907	
Average promotion rate	.4880	.5644	.6407	

Table 32(c): Promotion criteria for attaining in one year, the maintainable structure

$\underline{q}^* = [.1922, .1592, .1480, .1383, .1261, .1188, .1174]'$, which corresponds to a growth rate $\alpha = 0$.

Promotion criteria	Adjustments in system growth rate leading to the single step attainability of $\underline{q}^*$				
	$\alpha_0 = -.07$	$\alpha_0 = -.06$	$\alpha_0 = -.05$	$\alpha_0 = -.04$	$\alpha_0 = -.03$
p_{11}	.4585	.4152	.3719	.3285	.2852
p_{12}	.3851	.4284	.4717	.5151	.5584
p_{22}	.3824	.3280	.2737	.2193	.1650
p_{23}	.5540	.6080	.6623	.7167	.7710
p_{33}	.3309	.2593	.1876	.1159	.0442
p_{34}	.6088	.6804	.7521	.8238	.8955
p_{44}	.4217	.3426	.2634	.1842	.1051
p_{45}	.4988	.5779	.6571	.7363	.8154
p_{55}	.5425	.4565	.3706	.2846	.1986
p_{56}	.4047	.4907	.5766	.6626	.7486
p_{66}	.6311	.4531	.4552	.3672	.2792
p_{67}	.3015	.3895	.4774	.5654	.6534
p_{77}	.9289	.8171	.7054	.5936	.4819
Average repeat rate	.5280	.4517	.3754	.2990	.2227
Average promotion rate	.4034	.4797	.5560	.6324	.7087

Table 32(d): Promotion criteria for attaining in one year, the maintainable structure

$\underline{q}^* = [.2375, .1808, .1545, .1326, .1112, .0963, .0871]'$, which

corresponds to a growth rate $\alpha = 0.08$.

Promotion criteria	Adjustments in system growth rates leading to the single step attainability of $\underline{q}^*$				
	$\alpha_0 = .01$	$\alpha_0 = .02, \ldots,$	$\alpha_0 = .05, \ldots,$	$\alpha_0 = .08$	$\alpha_0 = .09$
p_{11}	.3905	.3472	.2172	.0872	.0438
p_{12}	.4531	.4964	.6264	.7564	.7998
p_{22}	.4931	.4388	.2757	.1126	.0582
p_{23}	.4429	.4973	.6603	.8234	.8778
p_{33}	.6121	.5404	.3254	.1103	.0387
p_{34}	.3276	.3993	.6143	.8294	.9010
p_{44}	.7719	.6927	.4552	.2178	.1386
p_{45}	.1486	.2278	.4653	.7028	.7819
p_{55}	.8712	.7853	.5273	.2693	.1833
p_{56}	.0760	.1619	.4199	.6779	.7639
p_{66}	.8372	.7493	.4853	.2214	.1335
p_{67}	.0954	.1833	.4473	.7112	.7991
p_{77}	.9300	.8182	.4830	.1478	.0361
Average repeat rate	.7009	.6241	.3956	.1666	.0903
Average promotion rate	.2305	.3068	.5358	.7648	.8411

Table 32(e); Promotion criteria for attaining in one year, the maintainable structure

$\underline{q}^* = [.2817, .1984, .1568, .1245, .0967, .0774, .0645]'$, which

corresponds to a growth rate $\alpha = 0.16$.

Promotion criteria	Adjustment in the system growth rate leading to the single step attainability of $\underline{q}^*$				
	$\alpha_0 = .12$	$\alpha_0 = .13, .., \alpha_0 = .15$		$\alpha_0 = .16$	$\alpha_0 = .17$
p_{11}	.1283	.1750	.0883	.0450	.0017
p_{12}	.6253	.6686	.7553	.7986	.8419
p_{22}	.4668	.4124	.3037	.2493	.1949
p_{23}	.4692	.5236	.6323	.6867	.7411
p_{33}	.6851	.6134	.4700	.3984	.3267
p_{34}	.2546	.3263	.4697	.5413	.6130
p_{44}	.8621	.7829	.6246	.5454	.4663
p_{45}	.0584	.1376	.2959	.3751	.4542
p_{55}	.9011	.8151	.6431	.5571	.4711
p_{56}	.0461	.1321	.3041	.3901	.4761
p_{66}	.7426	.6547	.4787	.3908	.3028
p_{67}	.1900	.2779	.4539	.5418	.6298
p_{77}	.5948	.4830	.2595	.1478	.0361
Average repeat rate	.6258	.5624	.4097	.3334	.2571
Average promotion rate	.3056	.3690	.5217	.5980	.6743

Table 32(f): Promotion criteria for attaining in one year, the maintainable structure

$\underline{q}^* = [.3030, .2057, .1568, .1200, .0898, .0692, .0555]'$, which corresponds to a growth rate $\alpha = 0.20$

Promotion criteria	Adjustments in the system growth rate leading the single step attainability of $\underline{q}^*$			
	$\alpha_0 = .18$	$\alpha_0 = .19$	$\alpha_0 = .20$	$\alpha_0 = .21, \ldots, \alpha_0 = 2$
p_{11}	.1179	.0746	.0313	"SINGLE STEP ATTAINABILITY NOT POSSIBLE"
p_{12}	.7257	.7690	.8123	
p_{22}	.4316	.3772	.3228	
p_{23}	.5044	.5588	.6132	
p_{33}	.6836	.6119	.5403	
p_{34}	.2561	.3278	.3994	
p_{44}	.8571	.7780	.6988	
p_{45}	.0634	.1425	.2217	
p_{55}	.8578	.7718	.6858	
p_{56}	.0894	.1754	.2614	
p_{66}	.6401	.5521	.4641	
p_{67}	.2925	.3805	.4685	
p_{77}	.3726	.2609	.1491	
Average repeat rate	.5658	.4895	.4132	
Average promotion rate	.3656	.4419	.5182	

Table 32(g): *Promotion criteria for attaining in one year, the maintainable structure*

$\underline{q}^* = [.3236, .2120, .1560, .1153, .0832, .0620, .0479]'$, which corresponds to a growth rate $\alpha = 0.24$

Promotion criteria	Adjustments in the system growth rate leading to the single step attainability of $\underline{q}^*$		
	$\alpha_0 = .23$	$\alpha_0 = .24$	$\alpha_0 = .25, \ldots, \ldots, \ldots, \alpha_0 = 2$
p_{11}	.0645	.0211	"SINGLE STEP ATTAINABILITY NOT POSSIBLE"
p_{12}	.7791	.8225	
p_{22}	.4517	.3973	
p_{23}	.4843	.5387	
p_{33}	.7480	.6763	
p_{34}	.1917	.2634	
p_{44}	.9201	.8409	
p_{45}	.0004	.0796	
p_{55}	.8867	.8007	
p_{56}	.0605	.1465	
p_{66}	.6144	.5264	
p_{67}	.3182	.4062	
p_{77}	.2595	.1478	
Average repeat rate	.5636	.4872	
Average promotion rate	.3678	.4442	

Comments on Tables 32(a) - 32(g)

Here we start from an education structure $\underline{q}(0) = [0.2308, 0.1840, 0.1395, 0.1262, 0.1163, 0.1137, 0.0895]'$ at the initial time, $t = 0$. Single step attainability of a maintainable structure $\underline{q}^*$ means obtaining a feasible promotion matrix P or promotion criteria so as to move from structure $\underline{q}(0)$ to structure $\underline{q}^*$ in a single time step.

It is important to note here that not all the maintainable structures of the Kenyan primary education system are attainable in a single time step. The only possible structures that can be attained in a single step are those corresponding to system growth rates in the interval -0.08 to 0.24. Even though the structures corresponding to the mentioned growth rates are attainable in a single step, some initial alteration in the system growth rate has to be made. In addition, some of the penalties in trying to 'force' a single step attainability may be painful. For example in some cases it may mean reducing the total system size thereby denying those eligible to join the system an opportunity. It may even mean forcing a large number of students to be retained in the system during the following year. Once the structure has been attained, we can remain in the structure by following the system growth rate and promotion criteria, which corresponds to the maintainable structure of interest.

Let us consider, for example, the structure $\underline{q}^* = [0.1476, 0.1340, 0.1366, 0.1397, 0.1399, 0.1445, 0.1577]'$, which is maintainable when the system growth rate is $\alpha = -0.08$. In order to attain this structure in a single step, the only possibility is when the initial system growth rate is $\alpha_0 = -0.15$. This corresponds to a decrease in the system size of 15%. Attainability of $\underline{q}^*$ is then achieved by promoting as in the column corresponding to $\alpha_0 = -0.15$ given in Table 32(a). This implies that in order to attain the above structure the promotion criteria should be such that, the average repeat rate is 0.3621 while the average promotion rate is 0.5692. We particularly note the penalty of this action via the grade seven repeat rate which in this case is 0.9310. This implies that most of the grade seven pupils will be forced to repeat in the following year in order to achieve our goal in a single time step.

The above cited deductions can be made for any of the maintainable structures which are possible to attain in one time step. For example, it is possible to attain the structure, $\underline{q}^* = [0.1922, 0.1592,$

0.1480,0.1383,0.1261,0.1188,0.1174], see Table 32(c). This structure is maintainable when the growth rate is $\alpha = 0$, that is for a fixed size system. The structure can be attained in a single step when the initial system growth rate α_0 lies in the interval -0.07 to -0.03. Of all these possible adjustments in the growth rate α_0 the least painful is that corresponding to $\alpha_0 = -0.03$, since this case has the least negative change in the total system size. Furthermore corresponding to $\alpha_0 = -0.03$, is the promotion criteria with the least average repeat rate of 0.2227 and the highest average promotion rate of 0.7087, which leads to the single step attainability.

We note generally that the highest possible adjustment in the initial growth rate is the least painful action to be taken in order to attain the corresponding maintainable structure in a single step. This is mainly because this action allows the highest possible system expansion and at the same time corresponds to the maximum average promotion rate and the minimum average repeat rate.

FIXED TIME ATTAINABILITY

An alternative way of attaining a target structure is to aim at getting as close as possible to the target in a prescribed number of steps. Suppose T^* denotes the time available, then we now aim to get as close as possible to our target $\underline{n}^*$ in T^* steps. We shall define closeness between the actual value $\underline{n}(T^*)$ reached and the target value, by a distance function $D(\underline{n}^*, \underline{n}(T^*))$. We shall again illustrate fixed time attainability first by controlling recruitment and then by controlling promotion.

Fixed Time Attainability by Recruitment

In this case, following a similar procedure as in the case of free time attainability, the problem of attaining a structure $\underline{n}^*$ by recruitment in a fixed time T^* can be formulated as that of solving the problem

$$\text{minimize} \quad D(\underline{n}^*, \underline{n}(T^*))$$

subject to the constraints

$$\underline{n}(T+1) = \underline{n}(T)P + \underline{f}(T+1) \qquad (T=0,1,2,\ldots,T^*-1) \qquad \text{(i)}$$

$$\sum_{1}^{s} n_i(T) = N(T); \quad (T=1,2,\ldots,T^*) \qquad \text{(ii)}$$

$$\underline{n}(T) \geq \underline{0} \qquad (T=1,2,\ldots,T^*-1) \tag{iii}$$

$$\underline{f}(T+1) \geq \underline{0} \qquad (T=1,2,\ldots,T^*-1) \tag{iv}$$

(6.49)

Note that, here we need to find a sequence of feasible recruitment vectors $\underline{f}(1),\underline{f}(2),\ldots,\underline{f}(T^*)$ which will minimize the distance function $D(\underline{n}^*,n(T^*))$. The solution to (6.49) will not only give the feasible recruitment vectors $\underline{f}(T)$'s which we require, but also the structures $\{\underline{n}(T);\ T=1,2,\ldots,T^*\}$ which the system will follow.

<u>Fixed Time Control by Promotion</u>

Here we obtain a sequence of promotion matrices $P(T)$'s which when coupled with a given recruitment vector $\underline{r}$, will take the system from an initial grade structure $\underline{n}(0)$ to $\underline{n}(T^*)$ such that the distance function $D(\underline{n}^*,\underline{n}(T^*))$ is minimized. The fixed time in which we wish to be as close as possible to n^* is T^*. Following a procedure similar to that for free time control, the problem is now to solve for the stocks and flows, $n_{ij}(T)$'s and $n_i(T)$'s so as to

$$\text{minimize} \quad D(\underline{n}^*,\underline{n}(T^*))$$

subject to the constraints

$$n_j(T+1) = \sum_1^s n_{ij}(T) + r_j\{\sum_1^s n_i(T)w_i + M(T+1)\} \tag{i}$$

$$(j=1,2,\ldots,s;T=0,1,2,\ldots,T^*-1)$$

$$\sum_{j=1}^{s} n_{ij}(T) = n_i(T)(1-w_i); \quad (i=1,2,\ldots,s-1; T=0,1,2,\ldots,T^*-1)$$

(ii)

$$\sum_1^s n_i(T) = N(T); \quad (T=1,2,\ldots,T^*) \qquad \text{(iii)}$$

$$\underline{n}(T) \geq 0; \qquad (T=1,2,\ldots,T^*) \qquad \text{(iv)}$$

$$n_{ij}(T) \geq (i,j=1,2,\ldots,s;\ T=0,1,2,\ldots,T^*-1) \qquad (6.50)$$

The distance function, $D(\underline{n}^*,\underline{n}(T^*))$, ought to reflect the penalties attached to having various grades over or under the targeted strength. One example of such a distance function may be

$$D_a(\underline{n}^*,\underline{n}(T^*)) = \sum_1^s \omega_i \left| n_i^* - n_i(T^*) \right|^a; \quad (a > 0) \qquad (6.51)$$

where $\omega_1, \omega_2, \ldots, \omega_s$ are a set of non-negative weights chosen to reflect the importance attached to the correct manning of grades $1,2,\ldots,s$.

As before let us consider a heirarchical organization with new enrollment only into the lowest grade and with promotions only to the next higher grade. Under these conditions, for example, the problem of attainability of the structure $\underline{n}^*$ in a single step, i.e. $T^* = 1$(fixed), can be written in expanded form as

minimize $D(\underline{n}^*,\underline{n}(1))$

subject to the constraints

$$
\text{s eqs}\begin{cases}
n_1(1) = n_{11}(0)+n_1(0)w_1+\ldots+n_s(0)w_s + M(1) \\
n_2(1) = n_{12}(0)+n_{22}(0) \\
\cdots\cdots\cdots\cdots \\
n_s(1) = n_{s-1,s}(0) + n_{ss}(0)
\end{cases}
$$

$$
\text{s-1 eqs}\begin{cases}
n_{11}(0)+n_{12}(0) = n_1(0)(1-w_1) \\
n_{22}(0)+n_{23}(0) = n_2(0)(1-w_2) \\
\cdots\cdots\cdots\cdots \\
n_{s-1,s-1}(0)+n_{s-1,s}(0) = n_{s-1}(0)(1-w_{s-1})
\end{cases}
$$

$$
\text{1 eqn}\begin{cases}
n_1(1)+n_2(1)+\ldots+n_s(1) = N(1)
\end{cases}
\tag{6.52}
$$

This may be written in matrix notation as

$$\text{minimize}\quad D\left(\underline{n}^*,\underline{n}(1)\right)$$

subject to

$$
C_1 \times \begin{bmatrix} \underline{n}_{ij}(0) \\ \text{--------} \\ \underline{n}(1) \end{bmatrix} = \underline{d}_1 = \begin{bmatrix} \underline{d}_{11} \\ \text{------} \\ N(1) \end{bmatrix}
\tag{6.53}
$$

where,

$$\underset{2s \times (3s-1)}{C_1} = \begin{array}{c} \begin{array}{cc} 2s-1 & s \end{array} \\ \left[\begin{array}{c|c} A_1 & \begin{array}{c} -I_s \\ 0_{(s-1)xs} \end{array} \\ \hline \underline{0}' & \underline{1}' \end{array}\right] \end{array} \begin{array}{c} 2s-1 \\ \\ 1 \end{array}$$

with

A_1 as the matrix of one step free time attainability given in (6.44);

0 and $\underline{0}'$ as a zero matrix and vector, respectively of the indicated dimensions.

and

$\underline{1}'$ as a $1 \times s$ row vector of ones.

The vector of the unknown variables $\left[\underline{n}_{ij}'(0), \underline{n}'(1)\right]$ is as defined earlier in the case of free time attainability. Lastly the vector $\underline{d}_1 = (\underline{d}_{11}' | N(1))$ is such that

$$\underline{d}_{11}' = \Big[\underbrace{-M(1)-\sum_1^s n_j(0)w_j, 0, \ldots, 0,}_{s} \underbrace{n_1(0)(1-w_1), n_2(0)(1-w_2), \ldots, n_{s-1}(0)(1-w_{s-1})}_{s-1}\Big]$$

and $N(1)$ is the total size of the system at time $T = 1$. We note that $\underline{d}_{11} = \underline{b}_{21}$, where $\underline{b}_{21}$ is the vector in (6.46).

When $T^* = 2$(fixed), the problem of attaining the structure $\underline{n}^*$ in two time intervals may similarly be written in expanded form as

$$\text{minimize} \quad D(\underline{n}^*, \underline{n}(2))$$

subject to

s eqs

$$n_1(1) = n_{11}(0) + n_1(0)w_1 + n_2(0)w_2 \ldots + n_s(0)w_s + M(1)$$
$$n_2(1) = n_{12}(0) + n_{22}(0)$$
$$\cdots\cdots\cdots$$
$$n_s(1) = n_{s-1,s}(0) + n_{ss}(0)$$

(s-1)eqs

$$n_{11}(0) + n_{12}(0) = n_1(0)(1-w_1)$$
$$n_{22}(0) + n_{23}(0) = n_2(0)(1-w_2)$$
$$\cdots\cdots\cdots$$
$$n_{s-1,s-1}(0) + n_{s-1,s}(0) = n_{s-1}(0)(1-w_{s-1})$$

1 eqn

$$n_1(1) + n_2(1) + \ldots + n_s(1) = N(1)$$

s eqs

$$n_1(2) = n_{11}(1) + n_1(1)w_1 + n_2(1)w_2 + \ldots + n_s(1)w_s + M(2)$$
$$n_2(2) = n_{12}(1) + n_{22}(1)$$
$$\cdots\cdots\cdots$$
$$n_s(2) = n_{s-1,s}(1) + n_{ss}(1)$$

(s-1)eqs

$$n_{11}(1) + n_{12}(1) = n_1(1)(1-w_1)$$
$$n_{22}(1) + n_{23}(1) = n_2(1)(1-w_2)$$
$$\cdots\cdots\cdots$$
$$n_{s-1,s-1}(1) + n_{s-1,s}(1) = n_{s-1}(1)(1-w_{s-1})$$

1 eqn

$$n_1(2) + n_2(2) + \ldots + n_s(2) = N(2) \tag{6.54}$$

This problem may be written in matrix notation as

$$\text{minimize} \quad D(\underline{n}^*, \underline{n}(2))$$

subject to

$$C_2 \times \left[\begin{array}{c} \underline{n}_{ij}(0) \\ \hline \underline{n}(1) \\ \hline \underline{n}_{ij}(1) \\ \hline \underline{n}(2) \end{array}\right] = \underline{d}_2 = \left[\begin{array}{c} \underline{d}_{21} \\ \hline N(1) \\ \hline \underline{d}_{22} \\ \hline N(2) \end{array}\right] \tag{6.55}$$

where now

$$\underset{4s \times (6s-2)}{C_2} = \begin{array}{cc} \begin{array}{cc} 5s-1 & s \end{array} & \\ \left[\begin{array}{c|c} A_2 & \begin{array}{c} 0 \\ -I_s \\ 0_{(s-1)\times s} \end{array} \\ \hline \underline{0}' & \underline{1}' \end{array}\right] & \begin{array}{c} 4s-1 \\ \\ 1 \end{array} \end{array}$$

with,

A_2 as the matrix of two step free time attainability given in (6.46);

0 and $\underline{0}'$ as zero matrices of the indicated orders.

The vector of unknowns $\left[\underline{n}_{ij}'(0), \underline{n}'(1), \underline{n}_{ij}'(1), \underline{n}'(2)\right]$ has entries as defined earlier in the case of free time control. Lastly the vector $\underline{d}_2' = (\underline{d}_{21}', N(1), \underline{d}_{22}', N(2))$ is such that

$$\underline{d}_{22} = \underline{d}_{11} = \underline{b}_{21}$$

and

$$\underline{d}_{22}' = \Big[\underbrace{-M(2),0,\ldots,0}_{s}\,\vdots\,\underbrace{0,0,\ldots,0}_{(s-1)}\Big] = \underline{b}_{32}'$$

where $\underline{b}_{32}$ is also given in (6.47). The values $\{N(T);T=1,2\}$ are the total system sizes at time T.

Generally when $T^* = k \geq 1$, the problem of attaining the structure $\underline{n}^*$ by promotion in k time periods (fixed), is that of

$$\text{minimizing} \quad D(\underline{n}^*, \underline{n}(k))$$

subject to

$$C_k \times \begin{bmatrix} \underline{n}_{ij}(0) \\ \hdashline \underline{n}(1) \\ \hdashline \underline{n}_{ij}(1) \\ \hdashline \vdots \\ \hdashline \underline{n}_{ij}(k-1) \\ \hdashline \underline{n}(k) \end{bmatrix} = \underline{d}_k = \begin{bmatrix} \underline{d}_{k1} \\ \hdashline N(1) \\ \hdashline \underline{d}_{k2} \\ \hdashline \vdots \\ \hdashline \underline{d}_{kk} \\ \hdashline N(k) \end{bmatrix} \qquad (6.56)$$

where,

$$
\underset{2ks \times [k(3s-1)]}{C_k} =
\begin{array}{c c}
\begin{array}{c|c} \quad k(3s-1)-s \quad & \quad s \quad \end{array} & \\
\left[\begin{array}{c|c}
 & 0 \\
A_k & -I_s \\
 & 0_{(s-1)\times s} \\
\hline
\underline{0}' & \underline{1}'
\end{array}\right] & \begin{array}{c} \\ 2ks-1 \\ \\ 1 \end{array}
\end{array}
$$

A_k is the matrix for k step attainability in free time as given in (6.48);

0 and $\underline{0}'$ are zero matrices of the indicated orders

The vector of unknowns has entries $\{\underline{n}_{ij}(T);T=0,1,\ldots,k-1\}$ and $\{\underline{n}(T);t=1,2,\ldots,k\}$ all described as before.

Lastly the vector $\underline{d}_k$ is such that

$$\underline{d}_{k1} = \underline{d}_{11} = \underline{b}_{21}$$

and

$$\underline{d}_{kj} = [\underbrace{-M(j),0,\ldots,0}_{s} \mid \underbrace{0,0,\ldots,0}_{(s-1)}] \qquad (j=2,3,\ldots,k)$$

$\{N(T);T=1,2,\ldots,k\}$ are the total system sizes at time T.

Again we note that the matrix C_k for the Kenyan primary education system, where the number of grades is $s = 7$, is of order ($14k \times 20k$). When $k = 2$ we are already dealing with a problem of order (28×40). As k increases the problem is bound to be pretty large, for example, when $k = 3$ the problem is of order 42×60 and so on.

Step by Step Method

Rather than try to attain a structure '$\underline{n}^*$' at a fixed time 'T^*' all at one go, we may adopt a step by step procedure which aims at moving as close as possible to the target at each step. This procedure is of great mathematical convenience since at each step we shall be merely dealing with a problem of one step attainability until all the T^* steps are covered.

In the particular case of a heirarchical organization with s grades, having admissions only into the lowest grade and promotions only to the next higher grade, when we apply the step by step procedure to fixed time control by promotion, we obtain the results given below:

When $T^* = 2$, meaning that we wish to attain the structure $\underline{n}^*$ in a fixed time $T^* = 2$; we first move as close as possible to the structure $\underline{n}^*$ at time $T = 1$; that is we find the intermediate stocks and flows which will

$$\text{minimize} \quad D(\underline{n}^*, \underline{n}(1))$$

subject to

$$C_1 \times \begin{bmatrix} \underline{n}_{ij}(0) \\ \hline \underline{n}(1) \end{bmatrix} = \underline{d}_1^* = \begin{bmatrix} \underline{d}_{11}^* \\ \hline N(1) \end{bmatrix} \tag{6.57}$$

where,

$\underline{d}_{11}^* = \underline{d}_{11}$, as in (6.53),

and,

C_1 is the matrix for one step attainability as given in (6.53).

Suppose that,

$$\hat{\underline{n}}(1) = \left[\hat{n}_1(1), \hat{n}_2(1), \ldots, \hat{n}_s(1)\right]'$$

is the first intermediate stock vector which minimizes $D(\underline{n}^*, \underline{n}(1))$. Then in the second step we obtain the stocks and flow values which will

$$\text{minimizes} \quad D(\underline{n}^*, \underline{n}(2))$$

subject to

$$C_1 \times \begin{bmatrix} \underline{n}_{ij}(1) \\ \hline \underline{n}(2) \end{bmatrix} = \underline{d}_2^* = \begin{bmatrix} \underline{d}_{21}^* \\ \hline N(2) \end{bmatrix} \tag{6.58}$$

where the vector of unknowns $[\underline{n}_{ij}^{'}(1), \underline{n}'(2)]$ is such that

$\underline{n}_{ij}(1)$ is the flow vector at time $T = 1$

and

$\underline{n}(2)$ is the stocks vector at time $T = 2$.

The matrix C_1 is as before but now

$$\underline{d}^*_{21} = \Big[\underbrace{-M(2)-\sum_1^s \hat{n}_j(1)w_j, 0, \ldots, 0,}_{s} \; \underbrace{\hat{n}_1(1)(1-w_1), \hat{n}_2(1-w_2), \ldots, \hat{n}_{s-1}(1)(1-w_{s-1})}_{s-1}\Big]'$$

Generally when $T^* = k$ (fixed), the step by step procedure of attaining the structure $\underline{n}^*$ is that of solving the 'k' consecutive problems:

$$\text{minimize} \quad D(\underline{n}(T), \underline{n}^*)$$

subject to

$$C_1 \times \begin{bmatrix} \underline{n}_{ij}(T-1) \\ \hline \underline{n}(T) \end{bmatrix} = \underline{d}^*_T = \begin{bmatrix} \underline{d}^*_{T1} \\ \hline N(T) \end{bmatrix} \qquad (T = 1, 2, \ldots, k) \qquad (6.59)$$

where,

$\underline{n}_{ij}(T-1)$ and $\underline{n}(T)$ are the flows and stocks vectors at times $(T-1)$ and T respectively $(T = 1, 2, \ldots, k)$;

C_1 is the matrix of one step attainability in fixed time as before;

and

$$\underline{d}^*_{T1} = \Big[-M(T) - \sum_{1}^{s} \hat{n}_j(T-1)w_j, 0, \ldots, 0 \;\Big|\; \hat{n}_1(T-1)(1-w_1), \hat{n}_2(T-1)(1-w_2), \ldots, \hat{n}_{s-1}(T-1)(1-w_{s-1}) \Big]'$$

$$\xleftarrow{\quad s \quad}\Big|\xrightarrow{\quad s-1 \quad}$$

with, $\hat{n}_j(T-1) = n_j(0)$ when $T=1$.

Otherwise, $\hat{n}_j(T-1)$ is the j-th entry of the stock vector at the (T-1)-th step which minimizes the distance function

$$D\big(\underline{n}^*, \underline{n}(T-1)\big).$$

For illustrative purposes we will give the promotion criteria for the Kenyan primary education system so as to attain some maintainable structure in a fixed number of steps, using the step by step procedure. In this case we consider the distance function

$$D\big(\underline{n}^*, n(T^*)\big) = \sum_{1}^{s} |n_i^* - n_i(T^*)| \qquad (6.60)$$

This corresponds to a case when all the grades of the system are assumed to be of equal importance. At each step we shall therefore obtain the solution to the system of linear equations

$$\underset{2s \times (3s-1)}{C_1} \times \underset{(3s-1) \times 1}{\begin{bmatrix} \underline{n}_{ij}(T-1) \\ \hdashline \underline{n}(T) \end{bmatrix}} = \underset{2s \times 1}{\underline{d}_T^*} , \qquad (6.61)$$

which minimizes the distance function $D(\underline{n}^*, \underline{n}(T))$. Since there are 2s equations in (3s-1) unknowns, a method of solving the system would be to solve for 2s of the unknowns in terms of the other (s-1). As such, suppose

$$\underline{x}(T) = \left[\underline{n}'_{ij}(T-1); \underline{n}'(T)\right]$$

represents the vector of unknowns. We can partition the vector $\underline{x}(T)$ and correspondingly the matrix C_1 as,

$$\underline{x}(T) = \begin{array}{c} \begin{bmatrix} \underline{x}_1(T) \\ \hdashline \underline{x}_2(T) \end{bmatrix} \end{array} \begin{array}{l} 2s \\ \\ s-1 \end{array}$$

and

$$C_1 = \begin{array}{cc} 2s & s-1 \\ \left[C_{11} \right. ; & \left. C_{12} \right] \end{array}$$

The system of equations in (6.61) can now be written as

$$\left[C_{11}; C_{12}\right] \begin{bmatrix} \underline{X}_1(T) \\ \hdashline \underline{X}_2(T) \end{bmatrix} = d_T^* \qquad (6.62)$$

Since there is no redundancy in (6.61) it is possible to choose the components of $\underline{x}_1(T)$ so that the corresponding square matrix C_{11} is invertible. The solution to $\underline{x}_1(T)$ for varying $\underline{x}_2(T)$ is therefore,

$$\underline{x}_1(T) = C_{11}^{-1}\left[\underline{d}_T^* - C_{12}\underline{x}_2(T)\right]$$

The range of variation of $\underline{x}_2(T)$'s may be enormous but this may be reduced by dividing the vectors $\underline{x}(T)$ by a constant so that the variation is fractional. The constant may be, for example, the total system size, $N(T)$. In addition the solutions of interest should be such that the stochastic property of the flow matrix $P(T)$ is not violated, that is

$$\underline{n}_{ij}(T)'s \geq \underline{0} \;;\; \underline{n}(T)'s \geq \underline{0}$$

and

$$n_{ii}(T) + n_{i,i+1}(T) \leq n_i(T), \text{ for each } T.$$

A more elaborate solution to the problem would be for example to linearize the problem and solve it using the standard methods of linear programming. However this approach requires further investigation since the problem is not easily linearizable.

In tables 33(a) - 33(c) below we give, for illustration purposes, some promotion criteria for attaining a maintainable structure $\underline{q}^*$ in three steps (fixed), using the approximate step by step procedure described above. The maintainable structure to be attained is

$$\underline{q}' = [.1922, .1592, .1480, .1383, .1261, .1188, .1174].$$

This corresponds to the system growth rate $\alpha = 0$ [refer to Table 31].

Table 33(a): Promotion criteria for attaining the maintainable structure $\underline{q}^*$ in three steps (fixed), using the approximate step by step procedure, with a growth rate $\alpha_0 = 0$.

Promotion criteria	1st step	2nd step	3rd step
p_{11}	.8436	.8436	.8436
p_{12}	0	0	0
p_{22}	.9360	.9360	.9360
p_{23}	0	0	0
p_{33}	.9397	.9397	.9397
p_{34}	0	0	0
p_{44}	.9205	.9205	.9205
p_{45}	0	0	0
p_{55}	.9128	.9472	.9472
p_{56}	.0017	0	0
p_{66}	.7391	.9326	.9077
p_{67}	.0018	0	.0025
p_{77}	.3935	.1478	.1478
Average repeat rate	.8122	.8096	.8061
Average promotion rate	.0871	.1217	.1221

Table 33(b): *Promotion criteria for attaining the maintainable structure $\underline{q}^*$ in three steps (fixed), using the approximate step by step procedure, with a growth rate $\alpha_0 = 0.5$.*

Promotion criteria	1st step	2nd step	3rd step
p_{11}	.8436	.8436	.8436
p_{12}	0	0	0
p_{22}	.9360	.9360	.9360
p_{23}	0	0	0
p_{33}	.9397	.9397	.9397
p_{34}	0	0	0
p_{44}	.9205	.9205	.9205
p_{45}	0	0	0
p_{55}	.7666	.8967	.9472
p_{56}	.0052	.0050	0
p_{66}	.4048	.0067	.0089
p_{67}	0	0	.8431
p_{77}	.9856	.6579	.1478
Average repeat rate	.8281	.7430	.6777
Average promotion rate	.0003	.0496	.2422

Table 33(c): Promotion criteria for attaining the maintainable structure $\underline{q}^*$ in three steps (fixed), using the approximate step by step procedure, with a growth rate $\alpha_0 = 1.0$.

Promotion criteria	1st step	2nd step	3rd step
p_{11}	.8436	.8436	.8436
p_{12}	0	0	0
p_{22}	.9360	.9360	.9360
p_{23}	0	0	0
p_{33}	.9397	.9397	.9397
p_{34}	0	0	0
p_{44}	.9205	.9205	.9205
p_{45}	0	0	0
p_{55}	.9472	.9472	.9472
p_{56}	0	0	0
p_{66}	.8974	.9326	.9326
p_{67}	.0035	0	0
p_{77}	.1477	.1478	.1478
Average repeat rate	.8046	.8096	.8096
Average promotion rate	.1223	.1217	.1217

Comments on Tables 33(a) - 33(c)

In this case, starting from an initial education structure, $\underline{q}'(0) = [0.2308, 0.1840, 0.1395, 0.1263, 0.1163, 0.1137, 0.0895]'$, the problem is to attain the maintainable structure $\underline{q}^* = [0.1922, 0.1592, 0.1480, 0.1383, 0.1261, 0.1188, 0.1174]'$ in three time steps (fixed), using the approximate step by step procedure. The structure $\underline{q}^*$ corresponds to a structure which is maintainable for a fixed size system.

In the first case (see Table 33(a)) it is assumed that the system is of fixed size so that the system growth rate is $\alpha_0 = 0$. In order to move from structure $\underline{q}(0)$ to $\underline{q}^*$ in three steps (fixed), the promotion criteria that will minimize the distance function $D(\underline{n}^*, \underline{n}(1))$ during the first step is given in column one of the table. This corresponds to an average repeat rate of 0.8122 and an average promotion rate of 0.0871. The promotions are only from grades five to seven, implying that most of the pupils of grades one to five will be retained in those grades so as to minimize the distance function, $D(\underline{n}^*, \underline{n}(1))$. The promotion criteria that will minimize the distance function $D(\underline{n}^*, \underline{n}(2))$ during the second step is given in the second column of Table 33(a). It corresponds to an average repeat rate of 0.8096 and an average promotion rate of 0.1217. The promotion is only due to grade seven which is about 0.8522. Most pupils in the rest of the grades, that is grades one to six will be forced to repeat so as to minimize the distance function $D(\underline{n}^*,\underline{n}(2))$. Finally the promotion criteria that will minimize the distance function $D(\underline{n}^*,\underline{n}(3))$ during the third step is in the third column of Table 33(a). It corresponds to an average repeat rate of 0.8061 and an average promotion rate of 0.1221. The promotion is due to grades six and seven only. Generally this means that most of the pupils will be forced to repeat in grades one to six in order to minimize the distance function.

We can obtain similar results when the system growth rate is different from zero. For example, for a system which grows at a rate $\alpha_0 = 1$, the promotion criteria which minimize the distance function are given in Table 33(c). The first column of Table 33(c) gives the promotion criteria which will minimize the distance function $D(\underline{n}^*,\underline{n}(1))$ during the first step. It corresponds to an average repeat rate of 0.8046 and an average promotion rate of 0.1223. The promotions are only from grades six and seven. This implies that in order to minimize the distance function $D(\underline{n}^*,\underline{n}(1))$ most of the grades one to six pupils will be forced to repeat. The promotion criteria that will minimize the distance function $D(\underline{n}^*,\underline{n}(2))$ during the second time period is given in the second column of Table 33(c). It corresponds to an average repeat rate of 0.8096 and an average promotion rate of 0.1217. In this case it is again observed that most of the pupils of grades one to six will be forced to repeat in order to minimize the distance function $D(\underline{n}^*,\underline{n}(2))$. Finally the promotion criteria that will minimize the distance function $D(\underline{n}^*,\underline{n}(3))$ during the third step are similar to those that minimize $D(\underline{n}^*,\underline{n}(2))$ during the second step.

In general it is observed that in order to minimize the distance function $D(\underline{n}^*,\underline{n}(T))$ at the T-th step $(T = 1,2,3)$, most of the pupils in grades one to six are forced to repeat the respective grades. The action taken based on these promotion criteria may be more appropriate in a manpower planning system rather than in an educational set up where pupil flow needs to be evident for the smooth running of the system. As a matter of policy, however, for example when things are out of control and we would like to alter them suddenly, then such action may be taken but together with some alteration in the country's education policy. This may for example include increasing the pupil age at entrance.

6.5 GENERAL COMMENTS AND CONCLUSIONS

Before giving comments based on this chapter, we note that, in a single educational system as the one considered in the application, the new enrolment is only in the first grade and as such control in the system characteristics can mainly be brought about via promotions. In the case of attainability we have looked at pupils joining grade one for the first time as if they somehow replace all those who leave the system at the end of the previous year. This means that we are assuming that the facilities that would have been used by those who drop out may be reconverted for use by the new entrants . Of course an alternative way for considering the grade one enrolment may be, for example, to consider it as composed of repeaters and new recruits only. We will first give comments on the maintainable grade structures followed by comments on attainability of grade structures and finally give some general conclusions on the entire chapter.

Comments on the Maintainable Grade Structures

A condition for maintainability of grade structures is obtained in the form of a system of linear equations. For varying system growth rates, grade structures which are maintainable are obtained by solving the system of equations. We first observe that when the system growth rate is less than or equal to negative one ($\alpha \leq -1$), then no grade structure is maintainable. This is because, for a system with growth rate less than or equal to negative one, all pupils leave the system during the next time period. When the system decays at a rate of 80% ($\alpha = -0.8$), the structure that is maintainable is seen to be a very top heavy type. This is most likely a consequence of the decrease in the total size of the system due to the negative growth rate. When the system is of fixed size, that is for zero growth rate ($\alpha = 0$), the maintainable structure is of a fairly uniform type. However, when the system size doubles annually, that is for a 100% growth rate ($\alpha = 1$), the maintainable structure is seen to be a bottom heavy type. This is again a consequence of the high system growth rate and the restriction that entry is only through grade one.

In general we observe that, as growth rate increases from negative to positive, the maintainable grade structure changes from top heavy to bottom heavy. We note also that the higher the repeat rates the less top heavy the structure that can be maintained.

Comments on Attainability of Grade Structures

The problem of attaining a grade structure in an educational system is that of obtaining the promotion criteria which will move the system from the initial structure $\underline{q}(0)$ to the target structure $\underline{q}^*$. For both free time attainability and fixed time attainability a matrix generalization of the problem is obtained.

In the case of free time attainability the problem is generally a linear programming one. For illustration we have only given the steps taken for single step attainability of the structure $\underline{q}^*$ which is maintainable starting from the initial structure $\underline{q}(0)$. We observe that the only possible structures that can be attained in a single step are those $\underline{q}^*$ corresponding to growth rates in the interval -0.08 to 0.24. Furthermore in order to attain these maintainable structures, some initial adjustments in the system growth rate have to be made. Additionally, some of the penalties of trying to 'force' a single step attainability are painful. For example it may involve reducing the total system size thereby denying those eligible to join the system the opportunity. In other cases it may even involve forcing a large number of pupils to repeat certain grades during the following year. However, once the structure $\underline{q}^*$ has been attained, we can maintain the grade structure by following the system growth rate and promotion criteria which corresponds to the maintainable structure of interest.

Next, in fixed time attainability strategy the problem has been formulated in matrix notation which is related to that of free time attainability. The main difference in this case is that of minimization of a distance function, which is in the form of the absolute distances between enrolments in the various grades and the target enrolments. The entire problem is not easily linearizable even after opting for the more mathematically convenient step by step procedure. For illustrative purposes we have given an approximate iterative procedure. More specifically an example is given for the attainability of a structure $\underline{q}^*$ in three steps (fixed). The structure

$\underline{q}^*$ is that which is maintainable in a fixed size system. It is observed here that the promotion actions which lead to fixed time attainability of $\underline{q}^*$ in three steps would be more appropriate in a manpower system rather than in an educational set up. This is mainly because in an education set up pupils should be seen to be flowing through the system. As a matter of policy, however, some of these actions may be taken in order to stop things from getting out of control, but together with some alterations in the education policy. This may for example include increasing the pupil age at entrance.

GENERAL CONCLUSIONS

A condition for maintainability of grade structures is obtained in the form of a system of linear equations dependent on the system growth rates. For varying system growth rates, grade structures which are maintainable are obtained by solving the system of equations. We particularly note the following points on the maintainable grade structure:

(i) As the system growth rate increases from negative to positive, the maintainable grade structures change from top heavy to bottom heavy.

(ii) The higher the repeat rates the less top heavy the maintainable grade structure since more pupils are retained in the system together with the fact that new entry is only through the first grade.

After identifying the set of maintainable grade structures one would look for the path to follow so as to attain any of these structures. A matrix generalization of the problem of attainability of grade structures is given for the education system . The free time attainability problem can then be solved by the standard linear programming methods. For illustrative purposes we have considered the course of action to be taken in 'forcing' a single step attainability of some maintainable structures via the free time procedure. The fixed time attainability problem can be made more mathematically convenient by following the step by step procedure. However, the problem is not easily linearizable and as such an approximate itterative procedure has been suggested. On attainability we particularly note the following points:

(i) As a matter of policy we may wish to attain some grade structures in a certain number of steps so as to stop things from getting out of hand. Some of the penalties of trying to force this type of attainability may be painful. For example, this may involve reducing the total system size or even forcing some students to repeat certain grades. However once the structure is attained, it is maintained by following the system growth rate and promotion criteria which corresponds to the maintainable structure in question.

(ii) Step-wise fixed time attainability seems to be more appropriate in a manpower planning set up rather than in an educational set up.

CHAPTER VII

CONCLUDING REMARKS

In this thesis we have presented a detailed statistical analysis of the basic characteristics of an educational system. The numerical results given in the thesis can be extended to the entire education system of a country, provided the relevant data is available on the stocks and flows of students in the various states and over an appropriate time period. Due to limitation on data availability, the applications of the model are based mainly on the Kenyan primary education system. Indeed within the period of this study, the country's education system changed fundamentally. Instead of the former system where students spent seven years in primary school, four years in secondary school, two years in high school and three years at the university; the new system now stipulate eight years in primary school, four years in secondary and four years at the university. This new system is popularly known as the 8:4:4 system. This system is only five years old since its introduction.

The models developed in this study can however be easily applied to the new 8:4:4 system of education by simply re-dimensioning the models.

For illustrating the applications of the models developed, we have used data from the primary component of the previous education system. This is because there was no adequate data on the other sectors. Moreover, there is not yet suitable data on the 8:4:4 system.

The study begins by examining the homogeneity of the Kenyan primary education system via the educational characteristics, over three equal time periods: 1964-1969, 1969-1974 and 1975-1980. It is assumed that within each of these time periods, the transition process is homogeneous. This enables us to apply the theory of homogeneous Markov chains to study the flow process in each of these periods, by computing the corresponding educational characteristics. Any appreciable difference in these characteristics between the time periods will suggest departure from homogeneity of the process over the entire period 1964-1980.

Comparing the results in the three time periods, it is first observed that the period 1975-1980 had the least average promotion rate of approximately 82% as compared to the promotion rates of 1964-1969 and 1969-1974 which were respectively 89% and 88%. On the other hand, the period 1975-1980 had the highest average repeat rate of about 9% as compared to the periods 1964-1969 and 1969-1975 each of which had average repeat rates of about 5% and 6% respectively. As a consequence of these average promotion rates, the study has shown that a grade one pupil spent a comparatively longer time in primary school during the period 1964-1969 than during the later time periods, 1969-1974 and 1975-1980. Furthermore, on average, any pupil in primary school during the earlier periods 1964-1969 and 1969-1974 was expected to stay longer in primary school than those in primary school during the period 1975-1980. These results seem to suggest that the Kenyan primary education system has not remained homogeneous over the period from 1964 to 1980.

Having found out that the Kenyan primary education system appears to be inhomogeneous with respect to the educational characteristics, a statistical test for homogeneity of the transitional probabilities between the states of the system has been performed. A chi-square test shows that the transition process is inhomogeneous. This suggests that the system changes with time, which leads to the study of time dependent processes.

In modelling the time dependent system, it has been assumed that the transition rates are probabilities of occurrence of random events which characterise the transition process. Specifically, the transition probabilities $p_{ij}(t)$'s are considered as the probability of occurrence of a random process with a corresponding distribution function, F(t), defined on the time domain. The time dependent transition models based on the normal, the logistic and the exponential probability distributions are proposed. These probability transition models are found to be useful in describing the transition process of the Kenyan primary education system, on the basis of goodness of fit tests. In particular it is noted that the proposed models may quite adequately describe the transition process for time periods close to the base year of operation. These probability transition models are then used to obtain the educational characteristics, using the theory of the inhomogeneous Markov chains.

A further application of the more general Markov chain model is considered via the generalized cohort analysis in order to study the flow of a particular cohort of students through an education system. For the purpose of this study, the term cohort is used to denote a group of students regardless of age or socio-economic background, who enter the first grade in the same academic year.

The illustrations of the application of the generalised cohort model are given on the basis of a time dependent probability transition process. Most of the results on the retention properties obtained by the generalized cohort model coincide with those of the general Markov chain model, when we consider the new entrants into the first school grade at time zero (1980). It is observed that, in general, the average repeat rates for pupils joining the school grades for the first time at time zero is approximately 9%. On the other hand the average promotion rate for these pupils is about 59%. After joining school, a pupil has on average a 67% chance of enrolling in any of the remaining school grades for the first time at time zero. Finally it is observed that any pupil enrolled in the education system for the first time at time zero is expected to take approximately 3 years 5 months in primary school.

The study goes further to consider the inhomogeneous transition process due to the effect of some quantifiable variables which are assumed to change with time. Measures of academic retentions are obtained when these additional factors are taken into account. An attempt is made to control some of these factors in order to achieve transition

processes which are as close as possible to some desired future targets. Some consequences of the proposed analysis are demonstrated using data from the Kenyan primary education system. For this illustration the pupil transition is assumed to be affected by the following factors:-
pupil - teacher ratios, yearly cost of education and the sex ratios.

The main advantage of using the variable dependent approach is that it takes into account factors which affect the transition process of the education system. In addition, control of the factors implies that which ever assumption on the transition process we opt for, the method gives a plan of action to be taken on the controllable variable so as to achieve the targeted assumptions optimally.

The other type of problem considered in the present study is that of attainability and maintain-ability of some educational characteristics. Suppose that the desired educational characteristics are specified and the problem is to find the transition process to be followed so as to achieve them. In this study the desired characteristics are specified in terms of grade or enrollment structures. A condition for maintainability of the grade structures is obtained in the form of a system of linear equations dependent on the system's growth rate. For varying growth rate, grade structures which are maintainable are obtained by solving the system of equations mentioned above. It is observed that as the system growth rate increases from negative to positive, the maintainable grade structures change from top heavy to bottom heavy. Furthermore, the higher the repeat rates the less top heavy the maintainable grade structure.

After identification of the set of maintainable grade structures we look for the paths which lead to these structures: This is the problem of attainability. A matrix generalization of the problem of attainability of grade structures is given for the system. Free time attainability can be solved by the standard linear programming methods. The fixed time attainability problem on the other hand is illustrated by an approximate iterative method.

One important aspect of the model discussed here is that it provides a good account of the variate in consideration, the student, from entry to exit. Furthermore, the study is of great significance since it provides the educational characteristics while taking into account the changing nature of the system over time. Such models are more relevant to education systems in developing countries which are highly unstable due to rapid population increases and other socio-economic factors. Once the education characteristics together with the grade structures are obtained, they may easily be used to obtain further educational measures such as; cost of education upto completion; staffing and capital requirements; pupil performances and so on. Some of these measures for the Kenyan primary education system are illustrated in Owino (1982). Other studies involving resource allocations in graded systems include Bowles (1967) and Uche (1978b). The problem of control in the system as considered in this study is of great economic interest. This is because its implementation at an appropriate time would definately stop a bad situation in the system from getting out of hand.

As a consequence of the present study there arises a number of interesting problems which may require further investigation. Some of these problems are described below:

(a) The factor dependent transition process needs further investigation so as to identify the quantifiable variables which affect the educational transition process.

(b) The multiple period factor control problem needs to be investigated further, without fixed intermediate promotion criteria as is the case in this thesis.

(c) The solution of the fixed time attainability problem needs further investigation. In this thesis only an approximate solution to the step by step procedure is given. The solution of the problem can be simplified if we linearize the objective function, for example. Alternatively, one needs to know more about distance functions.

(d) Finally, there is the 'bottle neck' effect on the transition of pupils from one sector of the education system to another. This is as a result of the inavailability of places for all pupils in all sectors of the education system. See for example Armitage, Smith and Alper (1969). See also Moya-Angeler (1976) and Warren and Mike (1971). Subject to availability of relevant and adequate data, this effect requires further study.

REFERENCES

1. ARMITAGE, P.H., C. SMITH, and P. ALPER (1969): Decision models for education planning., Allen Lane the Penguin Press.
2. ARMITAGE, P.H., C.M. PHILIPS, and J. DAVIES (1970): 'Towards a model of the upper secondary school systems (with discussions).' J.R. Statistic. Soc.,A133, 166-205.
3. BAISUCK, A and WALLACE, W.A. (1970): 'A computer simulation approach to enrolment projection in higher education.' Socio-Econ. Plan. Sci., 4, 365-381.
4. BALINSKY, W and REISMAN, A. (1973): 'A Taxonomy of manpower - educational planning models.' Socio-Econ. Plan. Sci.,7, 13-18.
5. BARTHOLOMEW, D.J. (1982): Stochastic models for social processes., John Wiley and Sons. 3rd ed.
6. BARTHOLOMEW, D.J. (1973): 'A model of completed length of service.' Omega, 1. 235-240.
7. BARTHOLOMEW, D.J. (1975): 'Errors of prediction in Markov chain models.' J.R. Statist. Soc., B37, 444-456.
8. BARTHOLOMEW, D.J. and FORBES, A.F. (1979): statistical techniques for manpower

planning., John Wiley and Sons.

9. BARTHOLOMEW, D.J. and SMITH, A.R. (1971): Manpower and management sciences., English University Press.

10. BECKER, N.G. (1968): 'Spread of an epidemic to fixed groups within the population.' Biometrics, 24, 1007-1014.

11. BOWLES, S. (1967): 'The efficient allocation of resources in education.' Quarterly Journal of Economics, X, 1.

12. BUTLER, A.D. (1971): 'An analysis of flows in a manpower system.' Statistician, 20, 69-84.

13. CHATFIELD, C. and GOODHARDT, G.L. (1970): 'The beta-binomial model for consumer purchasing behaviour.' App. Stat., 19, 240-250.

14. CLOWES, G.A. (1972): 'A dynamic model for the analysis of Labour turnover.' J.R. Statistic. Soc., A135, 242-256.

15. DAVIES, G.S. (1973): 'Structural control in a graded manpower system.' Man. Sci., 20, 76-84.

16. DAVIES, G.S. (1975): 'Maintainability of structures in Markov chain models under

recruitment control.' J. Appl. Prob., 12, 376-382.

17. DALEY, D.J. and KENDALL, D.G. (1965): 'Stochastic rumours.' J. Inst. Math. Applns., 1, 42-55.

18. DRUI, A.B. (1963): 'The use of regression equations to predict manpower requirements.' Man. Sci., 8, 669-677.

19. GANI, J. (1963): 'Formulae for projecting enrolment and degrees awarded in universities.' J.R. Statist. Soc., A126, 400-409.

20. GLEN, J.J. (1977): 'Length of service distributions in Markov manpower models.' Operat. Res. Quart., 28, 975-982.

21. GRINOLD, R.C. and STANFORD, R.E. (1978): 'Optimal control of a graded manpower system.' Man. Sci., 20, 1201-1216.

22. HAGERSTAND, T. (1967): Innovation diffusion as a spatial process., The University of Chicago Press.

23. JOHNSTONE, J.N. (1974): 'Mathematical models developed for use in educational planning.' Rev. Educ. Res., 44, 177-201.

24. JOHNSTONE, J.N. and H. PHILIP (1973): 'Application of Markov chain in an educational planning.' Socio-Econ. Plan. Sci., 7, 283-294.

25. KEYFITZ, N. (1968): Introduction to the mathematics of population., Addison-Wisley, Reading Mass.

26. KHOGALI, A.K. (1982): 'A study of an education system using the cohort model.' MSc. project submitted to mathematics department, University of Nairobi.

27. LE'GARE, J. (1972): 'Methods for measuring school performance through cohort analysis.' Demography, No.4 , 9, 617-624.

28. Mc NAMARA, J.F. (1973): 'Mathematical programming applications in education planning.' Socio-Econ. Plan. Sci., 7, 19-35.

29. MARK BLAUG (June 1967): 'Approaches to educational planning.' Economic Journal., LXXVII, p306.

30. MARSHALL, K.T. (1973): 'A comparison of two personnel prediction models.' Operat. Res., Vol.21, 810-834.

31. MARSHALL, K.T. and OLIVER, R.M. (1970): 'A constant work model for student attendance

and enrollment.' Operat. Res., 18, 193-206.

32. MARSHALL, M.L. (1971): 'Some statistical methods for forecasting wastage.' Statistician , 20, 53-68.

33. MARSHALL, M.L. (1975): 'Equilibrium age distribution for graded systems.' J.R. Statistic. Soc., A138, 62-69.

34. MOORE, S.C. (1975): 'On Johnstone and Philip's Markov chain educational planning model.' Socio-Econ. Plan. Sci., 9, 283-294.

35. MOYA-ANGELER, J. (1976): 'A model with shortage of places for educational and manpower systems.' Omega, 4, 719-730.

36. ODONDO, H.H. (1985): 'Differential evaluation of educational parameters using the markovian model.' MSc. project submitted to mathematics department, University of Nairobi.

37. ODHIAMBO, J.W. and KHOGALI, A.K. (1986): 'A transition model for estimating academic survival through cohort analysis.' Int. J. Math. Educ. Sci. Technol., 17, No.3, 339-346.

38. ODHIAMBO, J.W. and OWINO, J. (1985): 'A stochastic model for estimating academic survival in an education system.' Kenya J. Sci. Technol., A6, 59-67.

39. OWINO,J. (1982): 'Markovian model of educational planning.' MSc. project submitted to mathematics department, Univeristy of Nairobi.

40. OWINO, J. and C.M. Philips (1989): 'A comparison of retention properties of the Kenyan education system before and after 1972.' Kenya J. Sci. Technol., A, vol.9. (In Press).

41. PICKFORD, M. (1974): 'A statistical analysis of university administration expenditure.' J.R. Statistic. Soc., A137, 35-47.

42. PIELOU, E.C. (1976): Mathematical Ecology., Wiley Interscience , John Wiley and Sons, New York.

43. PRESS, S.J. (1982): Applied multivariate analysis., University of Chicago, Holt Rinechart and Winston Incl.

44. PROPOI, A. (1978): 'Models for educational and manpower planning. A dynamic linear programming approach.' R.M. - 78-20(IIASA).,

A-2361 Laxenburg, Austria.

45. SALES, P. (1971): 'The validity of the Markov Chain model for a class of the civil service.' Statistician., 20, 85-110.

46. SMITH, A.R. (1971): Models for manpower systems., English University Press, London.

47. STONE, R. (1966): 'Input-output and demographic accounting. A tool for educational planning.' Minerva , vol. IV, No.3, pp 365-380 , Spring (1966).

48. THONSTAD, T. (1967); In OECD (1967): 'Mathematical models in education planning.' OECD, Paris, 125-258.

49. THONSTAD, T. (1969): Education and Manpower. Theoretical models and empirical applications., Oliver Boyd, Edinburgh and London.

50. UCHE, P.I. (1978a): 'On Stochastic models for educational planning.' Int. J. Math. Educ. Sci. Technol., 9, 333-342.

51. UCHE, P.I. (1978b): 'Statistical analysis of expenditure of universities in a developing country. A Nigerian example.' J.R. Statist. Soc., A141, 3, pp.394-400.

52. UCHE, P.I. (1980): 'A transition model of academic survival in a single channel system with empirical examination.' Int. J. Math. Educ. Sci. Technol., 11, 177-187.

53. UCHE, P.I. (1982): 'Statistical models of the length of stay in the education system.' Int. J. Math. Educ. Sci. Technol., 15, No.4, 479-483.

54. UCHE, P.I. (1984): 'Some control measures for markov systems.' Int. J. Math. Educ. Sci. Technol., 15, No.4, 479-483.

55. VASSILIOU, P.C.G. (1976): 'A Markov Chain model for wastage in manpower systems.' Operat. Res. Quart., 27, 57-70.

56. WARREN, R.H. and MIKE, T.C. (1971): 'Projection of enrolment distribution with enrolment ceilings by Markov processes.' Socio-Econ. Plan. Sci., 5, 467-473.

57. WHITE, H.C. (1970): Chains of opportunity., Havard University Press, Cambridge, Mass.

58. WILSON, N.A.B. (1969): Manpower Research., English University Press, London.

59. YOUNG, A. (1971): Demographic and ecological models for manpower planning., edited by D.J. Bartholomew and B.R. Morris, London, English University Press.

60. YOUNG, A and VISSILOIU, P.C.G. (1974): 'A non linear model on the promotion of staff.' J.R. Statist. Soc., A137, 584-595.

61. J.R. Statist. Soc. A(1985): 'General discussions on projections of student numbers in higher education.' Part 3, pp.175-213.