International Journal of Operations Research

International Journal of Operations Research Vol. 9, No. 4, 209−219 (2012)

On the Transition Matrix of the Flow Mechanism in a Multi-Echelon Educational System

Virtue U. Ekhosuehi[1*] **and Augustine A. Osagiede**[2]

[1,2]Department of Mathematics, University of Benin, P.M.B. 1154, Benin City, Nigeria

Received April 2012; Revised September 2012; Accepted October 2012

Abstract — This paper is concerned with deriving, using logistic and Markov chain theoretic methodologies, a transition matrix for a multi-echelon educational system. The explanatory variables of the logistic model are the school differential variables, and the transition matrix of the Markov chain is the non-homogeneous empirical transition matrix (NHETM). We compare the NHETM with the periodically updated transition matrix suggested in literature using data in a university setting. The results indicate that the NHETM do not violate the flow mechanism of the academic programme and that the higher-order NHETM is not a sparse matrix.

Keywords — Binomial logistic model; Markov chain; multi-echelon educational system; transition matrix.

1. INTRODUCTION

The system we consider in this study is one with a set of states $\Re = \{0\} \cup S$, where the notation 0 denotes the state outside the educational system, and $S = \left\{s_1, \ldots, s_k\right\}$ is the set of levels in the educational system with the following features: there is a natural ordering of the states; the state $\{s_i\}$, $i = 1, \ldots, k$, is a singleton set; the states of the system are non-overlapping; the grades are finite and exhaustive; and a student enrolled in state $\left\{s_i\right\}$ cannot also be in state $\left\{s_j\right\}$, $i \neq j$, in any session t. Thereafter, we develop a transition matrix suitable for the non-homogeneous evolution of the educational system using the following notations: $\Psi\left(p_{i1}(t), \ldots, p_{ik}(t)\right)$ is the log-likelihood function of the distribution of the system. $\boldsymbol{\beta}_i = (\beta_{1i}, \cdots, \beta_{pi})'$ is an unknown parameter vector. $n_{ij}(t)$ is the number of students moving from level i to level j in period t. $n_{i0}(t)$ is the number of students leaving level i in period t. $p_{ij}(t)$ is the probability of students' flow from level i to level j in period t. $p_{i0}(t)$ is the probability of students leaving level i in period t. $p_{ij}(t)\big|_{\mathbf{x}_i'(t)}$ is the estimated probability of students' flow from level i to level j given $\mathbf{x}_i'(t)$ for period t. $p_{i0}(t)\big|_{\mathbf{x}_i'(t)}$ is the estimated probability of students leaving level i given $\mathbf{x}_i'(t)$ for period t. $p_{0j}(t)$ is the estimated admission probability into level j in period t. $\mathbf{x}_i'(t) = [1 \;\; x_{1i}(t) \;\cdots\; x_{(h-1)i}(t) \;\cdots\; x_{(p-1)i}(t)]$ with $x_{(h-1)i}(t)$, $h = 1, \ldots, p$, being an observation corresponding to the $(h-1)$th system's differential variables in level i in period t. $\mathbf{x}' = \begin{bmatrix} \mathbf{x}_1'(t) \\ \vdots \\ \mathbf{x}_k'(t) \end{bmatrix}$ is a $k \times p$ matrix of the set of differential variables of the system. $\mathbf{Q}(t)\big|_{\mathbf{x}'(t)}$ is the $k \times k$ non-homogeneous empirical transition matrix (NHETM) of the open system given $\mathbf{x}'(t)$ for each period t. $q_{ij}(t)\big|_{\mathbf{x}_i'(t)}$ is the entries in $\mathbf{Q}(t)\big|_{\mathbf{x}'(t)}$ such that $q_{ij}(t)\big|_{\mathbf{x}_i'(t)} = p_{ij}(t)\big|_{\mathbf{x}_i'(t)} + p_{i0}(t)\big|_{\mathbf{x}_i'(t)} p_{0j}(t)$ and $\sum_{j=1}^{k} q_{ij}(t)\big|_{\mathbf{x}_i'(t)} = 1$.

The states in the educational system may be either transient or absorbing. An absorbing state is a type of state in which upon entering the state, it is not possible to go to another state in the future. There are two absorbing states in the

∗ Corresponding author's email: virtue.ekhosuehi@uniben.edu

educational system, namely: dropout and completion (or graduation) states. Dropout state encapsulates students who withdraw from the programme before completion. The transient state is the state where upon entering the process may never return to the state again. The levels in the educational system are transient states because all movements will be towards absorption and away from the levels. The implication of this is that the outcome of the evolution of the system can be viewed as a binary variable. A computationally easy approach to model such binary variables is the binomial logistic model (Dobson, 2002). Jimenez and Salas-Velasco (2000) had earlier employed the binomial logistic model in the study of the demand for higher education in Spain. In this light, we formulate the following optimization problem:

OP: Maximize

$$\Psi\left(p_{i1}(t),\dots,p_{ik}(t) \right) = \sum_{j=1}^{k} n_{ij}(t)\log p_{ij}(t) \tag{1}$$

subject to:

i. the binomial logistic model constraint: $\log\left[p_{i0}\left(t\right)\left(\sum_{j=1}^{k} p_{ij}\left(t\right)\right)^{-1}\right] = \mathbf{x}_i'(t)\boldsymbol{\beta}_i$, (2a)

ii. the stochastic property: $\sum_{j=1}^{k} p_{ij}\left(t\right) + p_{i0}\left(t\right) = 1$, (2b)

iii. the non-negativity condition: $p_{ij}\left(t\right) \geq 0$, $p_{i0}\left(t\right) \geq 0$, $i, j \in S$. (2c)

What we intend to accomplish by the formulation of problem **OP** is to derive a NHETM wherein the school differential variables are inscribed into the transition probabilities of the multi-echelon educational system in a systematic manner. Uche (1978) earlier suggested that, for a better description of the educational system, the transition probabilities of the system may depend on the amount of money voted for education or on some other non-quantifiable factors, but he did not attempt to do this. From the literature available to us, no author has attempted to incorporate this suggestion of Uche (1978). By solving problem **OP**, we endeavour to bridge the gap identified by Uche (1978) using school fees as a proxy for money voted for education as well as other school's differential variables. The school's differential variables include: tuition fees and charges per year of study, promotion criteria, and environmental factors such as land use mix, traffic zone, etc. School fees and other differential variables are incorporated into the transition model by assuming a binomial logistic wastage rate for the multi-echelon educational system. The closest rivalry to our present study is the work of Ekhosuehi (2009). For this reason, we shall compare our proposed NHETM with the periodic updated transition matrix (PUTM) of Ekhosuehi (2009).

2. LITERATURE REVIEW

In modelling students' flows in the educational system, authors have commonly assumed a deterministic growth factor for new entrants (Osagiede and Omosigho, 2004; and Osagiede and Ekhosuehi, 2006). The growth factor was, hitherto, chosen arbitrarily by various authors until Osagiede and Ekhosuehi (2006) proffered a solution by providing a unique estimator for its computation. A common approach in modelling students' flows is the use of Markov chain. The classical Markov chain model for the multi-echelon educational system was developed by Gani (1963). Since then other similar models have been used in literature (Uche, 1978; Nicholls, 2009). The classical Markov chain model which Nicholls (2009) called the absorbing Markov chain model has a sub-stochastic transition matrix. The $N \times N$ absorbing Markov chain with r absorbing states, $\mathbf{A}$ say, has a block structure of the form (Ibe, 2009): $\mathbf{A} = \begin{bmatrix} \mathbf{P} & \vdots & \mathbf{w}' \\ \cdots & \cdots & \cdots \\ \mathbf{0} & \vdots & \mathbf{I} \end{bmatrix}$, where $\mathbf{w}$ is an $m \times r$ transition matrix from nonabsorbing to absorbing states, $m = N - r$ is the nonabsorbing (or transient) states, $\mathbf{P}$ is an $m \times m$ transition matrix among the transient states, $\mathbf{I}$ is an $r \times r$ identity matrix, $\mathbf{0}$ is an $r \times m$ matrix whose entries are all zero and the prime in $\mathbf{w}$ denotes transposition. The major limitation in the use of the absorbing Markov chain is that the entry probabilities are not accounted for. As a consequence, the process enters a state of absorption after some transition steps. This is shown as follows:

$$\lim_{t\to\infty} \mathbf{A}^t = \begin{bmatrix} \lim_{t\to\infty} \mathbf{P}^t & \vdots & \lim_{t\to\infty} \mathbf{w}_t^* \\ \cdots & \cdots & \cdots \\ \mathbf{0} & \vdots & \mathbf{I} \end{bmatrix},$$

where $\mathbf{w}_t^* = \mathbf{w}' + \mathbf{P}\mathbf{w}' + \mathbf{P}^2\mathbf{w}' + \cdots + \mathbf{P}^{t-1}\mathbf{w}'$. But, $\lim_{t\to\infty} \mathbf{P}^t = \mathbf{0}$ since $\left\|\mathbf{P}\right\| < 1$ as $\mathbf{P}$ is sub-stochastic. So,

$$\lim_{t\to\infty} \mathbf{A}^t = \begin{bmatrix} \mathbf{0} & \vdots & \lim_{t\to\infty} \mathbf{w}_t^* \\ \cdots & \cdots & \cdots \\ \mathbf{0} & \vdots & \mathbf{I} \end{bmatrix}.$$

The implication of this result is that no student would be in the system after a long period of time. This result is worrisome as it is expected that the educational system strives for long-term survival. Against this backdrop, Osagiede and Ekhosuehi (2006) proposed a homogeneous Markov chain, $\mathbf{Q}$, of the form: $\mathbf{Q} = \mathbf{P} + \mathbf{w'r}$, to describe the flow mechanism in the educational system, where $\mathbf{r}$ is the admission probability vector. We shall hereafter refer to the matrix $\mathbf{Q}$ as the homogeneous imbedded Markov chain (HIMC). The matrix $\mathbf{Q}$ is stochastic as $\mathbf{Qe'} = \mathbf{e'}$, where $\mathbf{e'}$ is a column vector of ones conformable with matrix $\mathbf{Q}$. Osagiede and Ekhosuehi (2006) however assumed that the HIMC is stationary. In practice, a transition matrix may not be stationary in the interim not until the process has evolved over time. This result can be shown using the chi-square test statistic on empirical data (Zanakis and Maret, 1980). In this light, Ekhosuehi (2009) suggested the periodic update of the transition matrix whenever it is non-stationary. Thus, the transition probabilities for each period t are computed as:

$$\hat{p}_{ij}(t) = \frac{\sum_{u=1}^{t} n_{ij}(u)}{\sum_{u=1}^{t}\sum_{j=1}^{k} n_{ij}(u)}, \quad t = 1,2,\cdots,T;\ i,j \in S. \tag{3}$$

In a similar manner, the wastage and admission probabilities are obtained. We refer to the transition matrix arising from the periodic update analogous to the HIMC as the periodically updated transition matrix (PUTM).

3. THE PROPOSED NON-HOMOGENEOUS EMPIRICAL TRANSITION MATRIX

The problem **OP** is formulated on the assumptions that:

i. the transition probabilities satisfy the multinomial distribution (Osagiede and Ekhosuehi, 2006) of the form:

$$P\left(n_{i0}(t), n_{i1}(t), \ldots, n_{ik}(t)\right) = \frac{\Gamma\left(\sum_{j=0}^{k} n_{ij}(t) + 1\right)}{\prod_{j=0}^{k} \Gamma\left(n_{ij}(t)+1\right)} \prod_{j=0}^{k} \left(p_{ij}(t)\right)^{n_{ij}(t)}.$$

ii. the wastage probabilities vary with the differential variables of the system in a binomial logistic manner. This is justifiable as the educational system is made up of two main states: transient and absorbing.

To solve **OP**, we let $S = \left\{1,2,\ldots,k\right\}$ since $\left\{s_i\right\}$ is a natural ordering singleton set, and consider the constraint (2a). We simplify constraint (2a) as follows:

$$p_{i0}\left(t\right)\left(\sum_{j=1}^{k} p_{ij}\left(t\right)\right)^{-1} = \exp(\mathbf{x}_i'(t)\boldsymbol{\beta}_i) \tag{4}$$

By adding one to both sides, we get: $1 + p_{i0}\left(t\right)\left(\sum_{j=1}^{k} p_{ij}\left(t\right)\right)^{-1} = 1 + \exp(\mathbf{x}_i'(t)\boldsymbol{\beta}_i)$. Using the constraint (2b), we have

$$\left(\sum_{j=1}^{k} p_{ij}\left(t\right)\right)^{-1} = 1 + \exp(\mathbf{x}_i'(t)\boldsymbol{\beta}_i). \tag{5}$$

Taking the natural logarithm of both sides, we obtain

$$\log\left(\sum_{j=1}^{k} p_{ij}\left(t\right)\right) = -\log[1 + \exp(\mathbf{x}_i'(t)\boldsymbol{\beta}_i)]. \tag{6}$$

Let $z_i^m\left(t\right)$ be a binary random variable defined as

$$z_i^m\left(t\right) = \begin{cases} 1 & \text{if a student } m \text{ leaves level } i \text{ in session } t \\ 0 & \text{if no student leaves level } i \text{ in session } t \end{cases}$$

with probability $\text{Prob}\left(z_i^m\left(t\right) = 1\right) = p_{i0}\left(t\right)$ and $\text{Prob}\left(z_i^m\left(t\right) = 0\right) = \sum_{j=1}^{k} p_{ij}\left(t\right)$, for all $i \in S$. Let $w_i\left(t\right)$ be a random variable which represents the total number of wastage in level i in session t. If $n_i\left(t\right)$ is the total number of students enrolled in level i in session t, then $w_i(t) = \sum_{m=1}^{n_i(t)} z_i^m(t)$. Since a student may either remain or leave the educational system, $z_i^m(t)$ is a binary random variable and the distribution of the random variable $w_i(t)$ is a binomial distribution (Lindgren, 1993) of the form

$$\text{Prob}\left(w_i(t) = n_{i0}(t)\right) = \binom{n_i(t)}{n_{i0}(t)} \left(p_{i0}(t)\right)^{n_{i0}(t)} \left(\sum_{j=1}^{k} p_{ij}(t)\right)^{n_i(t)-n_{i0}(t)}, \text{ for all } i \in S,$$

where $n_{i0}(t)$ is the wastage flow from level i in session t. Next, we express the distribution of the random variable $w_i(t)$ as a member of the exponential family using exponential and logarithmic operators as

$$\text{Prob}\left(w_i(t) = n_{i0}(t)\right) = \exp\left\{\log\left(\binom{n_i(t)}{n_{i0}(t)} \left(p_{i0}(t)\right)^{n_{i0}(t)} \left(\sum_{j=1}^{k} p_{ij}(t)\right)^{n_i(t)-n_{i0}(t)}\right)\right\}.$$

After some algebra, we obtain

$$\text{Prob}\left(w_i(t) = n_{i0}(t)\right) = \exp\left\{n_{i0}(t)\log\left[p_{i0}(t)\left(\sum_{j=1}^{k} p_{ij}(t)\right)^{-1}\right] + n_i(t)\log\left(\sum_{j=1}^{k} p_{ij}(t)\right) + \log\binom{n_i(t)}{n_{i0}(t)}\right\},$$

where

$$\log\binom{n_i(t)}{n_{i0}(t)} = \log\frac{n_i(t)!}{n_{i0}(t)!\left(n_i(t) - n_{i0}(t)\right)!}.$$

The log-likelihood function is

$$L_i = \sum_{t=1}^{T}\left\{n_{i0}(t)\log\left[p_{i0}(t)\left(\sum_{j=1}^{k} p_{ij}(t)\right)^{-1}\right] + n_i(t)\log\left(\sum_{j=1}^{k} p_{ij}(t)\right) + \log\binom{n_i(t)}{n_{i0}(t)}\right\}. \quad (7)$$

To estimate the values of $\boldsymbol{\beta}_i$, we substitute the right hand side terms of the constraint (2a) and equation (6) into equation (7), so that we have

$$L_i = \sum_{t=1}^{T}\left\{n_{i0}(t)\mathbf{x}_i'(t)\boldsymbol{\beta}_i - n_i(t)\log[1 + \exp(\mathbf{x}_i'(t)\boldsymbol{\beta}_i)] + \log\binom{n_i(t)}{n_{i0}(t)}\right\} \quad (8)$$

Taking the partial derivatives of (8) with respect to each element of $\boldsymbol{\beta}_i$, we have

$$\frac{\partial L_i}{\partial \beta_{1i}} = \sum_{t=1}^{T}\left(n_{i0}(t) - n_i(t)\frac{\exp(\mathbf{x}_i'(t)\boldsymbol{\beta}_i)}{[1 + \exp(\mathbf{x}_i'(t)\boldsymbol{\beta}_i)]}\right),$$

$$\frac{\partial L_i}{\partial \beta_{hi}} = \sum_{t=1}^{T}\left(n_{i0}(t)x_{(h-1)i}(t) - n_i(t)\frac{x_{(h-1)i}(t)\exp(\mathbf{x}_i'(t)\boldsymbol{\beta}_i)}{[1 + \exp(\mathbf{x}_i'(t)\boldsymbol{\beta}_i)]}\right), \; h = 2, \cdots, p.$$

Let $\mathbf{U}\left(\boldsymbol{\beta}_i\right) = \begin{pmatrix} \frac{\partial L_i}{\partial \beta_{1i}} \\ \vdots \\ \frac{\partial L_i}{\partial \beta_{pi}} \end{pmatrix}$.

Our interest is to obtain a solution to the problem $\mathbf{U}\left(\boldsymbol{\beta}_i\right) = \mathbf{0}$, which is nonlinear in $\boldsymbol{\beta}_i$. To do this, we employ the iteratively reweighted least squares algorithm (Hardle et al., 2004). The method involves solving repeatedly

$$\boldsymbol{\beta}_i^{(\gamma+1)} = \boldsymbol{\beta}_i^{(\gamma)} - [\Im(\boldsymbol{\beta}_i^{(\gamma)})]^{-1}\mathbf{U}(\boldsymbol{\beta}_i^{(\gamma)}), \qquad \gamma = 0, 1, 2, \ldots \quad (9)$$

provided $\det[\Im(\boldsymbol{\beta}_i^{(\gamma)})] \neq 0$, where $\Im(\boldsymbol{\beta}_i^{(\gamma)}) = \left[\left. \dfrac{\partial^2 L_i}{\partial \boldsymbol{\beta}_i \partial \boldsymbol{\beta}_i'} \right|_{\boldsymbol{\beta}_i = \boldsymbol{\beta}_i^{(\gamma)}} \right]$ is the $p \times p$ matrix of second-order partial derivatives of the log-likelihood function evaluated at the γ^{th} iteration around $\boldsymbol{\beta}_i$. The matrix $\Im(\boldsymbol{\beta}_i^{(\gamma)})$ is a symmetric matrix, and $[\Im(\boldsymbol{\beta}_i^{(\gamma)})]^{-1}$ is the dispersion matrix. The iteration is started from $\boldsymbol{\beta}_i^{(0)} = \mathbf{0}$ and it stops when the parameter estimates do not change significantly any more. Let $\boldsymbol{\beta}_i^*$ be the numerical solution to the system in equation (9). Then, from equations (4) and (5), we obtain $p_{i0}(t)$ and the sum $\sum_{j=1}^{k} p_{ij}(t)$ as:

$$p_{i0}(t) = [1 + \exp(\mathbf{x}_i'(t)\boldsymbol{\beta}_i^*)]^{-1}\exp(\mathbf{x}_i'(t)\boldsymbol{\beta}_i^*) \tag{10}$$

and

$$\sum_{j=1}^{k} p_{ij}(t) = [1 + \exp(\mathbf{x}_i'(t)\boldsymbol{\beta}_i^*)]^{-1}. \tag{11}$$

Using the result in equation (11), the objective function (1) is rewritten as

$$\Psi\left(p_{i1}(t),\ldots,p_{ik}(t)\right) = \sum_{j=1}^{k-1} n_{ij}(t)\log p_{ij}(t) + n_{ik}(t)\log\left([1 + \exp(\mathbf{x}_i'(t)\boldsymbol{\beta}_i^*)]^{-1} - \sum_{j=1}^{k-1} p_{ij}(t)\right).$$

Taking the partial derivatives of $\Psi\left(p_{i1}(t),\ldots,p_{ik}(t)\right)$ with respect to each $p_{ij}(t)$, $j \in S/\{k\}$, we obtain

$$\frac{\partial \Psi}{\partial p_{ij}(t)} = \frac{n_{ij}(t)}{p_{ij}(t)} - \frac{n_{ik}(t)}{[1 + \exp(\mathbf{x}_i'(t)\boldsymbol{\beta}_i^*)]^{-1} - \sum_{j=1}^{k-1} p_{ij}(t)}, \; j \in S/\{k\}.$$

By setting $\dfrac{\partial \Psi}{\partial p_{ij}(t)} = 0$ and taking $p_{ik}(t) = [1 + \exp(\mathbf{x}_i'(t)\boldsymbol{\beta}_i^*)]^{-1} - \sum_{j=1}^{k-1} p_{ij}(t)$, from equation (11), we have

$$n_{ij}(t)p_{ik}(t) = n_{ik}(t)p_{ij}(t). \tag{12}$$

Summing over all $j \in S$, we get $p_{ik}(t)\sum_{j=1}^{k} n_{ij}(t) = n_{ik}(t)\sum_{j=1}^{k} p_{ij}(t)$. Using the result in equation (11), we have

$$p_{ik}(t) = \frac{n_{ik}(t)}{\sum_{j=1}^{k} n_{ij}(t)}[1 + \exp(\mathbf{x}_i'(t)\boldsymbol{\beta}_i^*)]^{-1}, \qquad i \in S. \tag{13}$$

By substituting the result for $p_{ik}(t)$ in equation (13) into equation (12), we obtain

$$p_{ij}(t) = \frac{n_{ij}(t)}{\sum_{j=1}^{k} n_{ij}(t)}[1 + \exp(\mathbf{x}_i'(t)\boldsymbol{\beta}_i^*)]^{-1}, \qquad i \in S,\; j \in S/\{k\}. \tag{14}$$

Thus, we have for $i, j \in S$ that

$$p_{ij}(t) = \frac{n_{ij}(t)}{\sum_{j=1}^{k} n_{ij}(t)}[1 + \exp(\mathbf{x}_i'(t)\boldsymbol{\beta}_i^*)]^{-1}. \tag{15}$$

Since $\exp(\mathbf{x}_i'(t)\boldsymbol{\beta}_i^*), n_{ij}(t) \geq 0$, the non-negativity constraints (2c) are met automatically. Observe that the results in equation (10) and (15) are estimated from the differential variables and the enrolment stocks. Hence, we state our solution to **OP** as

$$p_{ij}(t)\Big|_{\mathbf{x}_i'(t)} = \begin{cases} [1 + \exp(\mathbf{x}_i'(t)\boldsymbol{\beta}_i^*)]^{-1}\exp(\mathbf{x}_i'(t)\boldsymbol{\beta}_i^*), & \text{for } i \in S,\; j = 0 \\ \dfrac{n_{ij}(t)}{\sum_{j=0}^{k} n_{ij}(t)}[1 + \exp(\mathbf{x}_i'(t)\boldsymbol{\beta}_i^*)]^{-1}, & \text{for } i \in S,\; j \in S \end{cases} \tag{16}$$

provided $\det[\Im(\boldsymbol{\beta}_i^{(\gamma)})] \neq 0$, for each $\gamma = 0,1,2,\ldots$. However, if $\det[\Im(\boldsymbol{\beta}_i^{(\gamma)})] = 0$, then no information is contained in the differential variables. Consequently, **OP** is maximized without the binomial logistic constraint. Subsequently, we construct the NHETM expressed as:

$$NHETM = \begin{pmatrix} \text{internal transition probabilities} \\ \text{in the system at time t} \end{pmatrix} + \begin{pmatrix} \text{probability of consequential recruitments} \\ \text{to replace losses at time t} \end{pmatrix}.$$

Let $p_{0j}(t)$ be the admission probability of new entrants into level j. Let:

$$\mathbf{P}(t)\Big|_{\mathbf{x}'(t)} = \left\{ \left(p_{ij}(t)\Big|_{\mathbf{x}_i'(t)} \right) : \sum_{j=1}^{k} p_{ij}(t)\Big|_{\mathbf{x}_i'(t)} \leq 1,\ p_{ij}(t)\Big|_{\mathbf{x}_i'(t)} \geq 0,\ i,j \in S \right\},$$

$$\mathbf{w}(t)\Big|_{\mathbf{x}'(t)} = \left\{ \left(p_{10}(t)\Big|_{\mathbf{x}_i'(t)} \quad \cdots \quad p_{k0}(t)\Big|_{\mathbf{x}_i'(t)} \right) : \sum_{i=1}^{k} p_{i0}(t)\Big|_{\mathbf{x}_i'(t)} \leq 1,\ p_{i0}(t)\Big|_{\mathbf{x}_i'(t)} \geq 0,\ i \in S \right\}, \text{ and}$$

$$\mathbf{P}_0(t) = \left\{ \left(p_{01}(t) \quad \cdots \quad p_{0k}(t) \right) : \sum_{j=1}^{k} p_{0j}(t) = 1,\ p_{0j}(t) \geq 0,\ j \in S \right\}.$$

Then the NHETM for session t is obtained in matrix form as: $\mathbf{Q}(t)\Big|_{\mathbf{x}'(t)} = \mathbf{P}(t)\Big|_{\mathbf{x}'(t)} + \mathbf{w}'(t)\Big|_{\mathbf{x}'(t)} \mathbf{P}_0(t)$.

4. APPLICATION

To demonstrate the utility of the NHETMs, we use enrolment data as contained in the results approved by Senate of the University of Benin, Nigeria for a part-time undergraduate programme for the period 2003/2004-2008/2009 sessions. The data are represented in the flow matrices: $\mathbf{F}1 - \mathbf{F}6$, $\mathbf{W}1 - \mathbf{W}6$, and $\mathbf{R}1 - \mathbf{R}6$.

$$\mathbf{F}1 = \begin{bmatrix} 0 & 112 & 0 & 0 & 0 & 0 \\ 0 & 0 & 53 & 0 & 0 & 0 \\ 0 & 0 & 0 & 56 & 0 & 0 \\ 0 & 0 & 0 & 0 & 30 & 0 \\ 0 & 0 & 0 & 0 & 0 & 35 \\ 0 & 0 & 0 & 0 & 0 & 8 \end{bmatrix},\ \mathbf{W}1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 10 \end{bmatrix},\ \mathbf{R}1 = \begin{bmatrix} 112 & 4 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\mathbf{F}2 = \begin{bmatrix} 0 & 106 & 0 & 0 & 0 & 0 \\ 0 & 0 & 90 & 0 & 0 & 0 \\ 0 & 0 & 0 & 45 & 0 & 0 \\ 0 & 0 & 0 & 0 & 48 & 0 \\ 0 & 0 & 0 & 0 & 0 & 26 \\ 0 & 0 & 0 & 0 & 0 & 8 \end{bmatrix},\ \mathbf{W}2 = \begin{bmatrix} 4 \\ 22 \\ 8 \\ 8 \\ 4 \\ 35 \end{bmatrix},\ \mathbf{R}2 = \begin{bmatrix} 110 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\mathbf{F}3 = \begin{bmatrix} 0 & 234 & 0 & 0 & 0 & 0 \\ 0 & 0 & 78 & 0 & 0 & 0 \\ 0 & 0 & 0 & 87 & 0 & 0 \\ 0 & 0 & 0 & 0 & 45 & 0 \\ 0 & 0 & 0 & 0 & 0 & 43 \\ 0 & 0 & 0 & 0 & 0 & 13 \end{bmatrix},\ \mathbf{W}3 = \begin{bmatrix} 2 \\ 28 \\ 3 \\ 0 \\ 5 \\ 21 \end{bmatrix},\ \mathbf{R}3 = \begin{bmatrix} 236 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\mathbf{F}4 = \begin{bmatrix} 0 & 346 & 0 & 0 & 0 & 0 \\ 0 & 0 & 226 & 0 & 0 & 0 \\ 0 & 0 & 0 & 78 & 0 & 0 \\ 0 & 0 & 0 & 0 & 87 & 0 \\ 0 & 0 & 0 & 0 & 0 & 43 \\ 0 & 0 & 0 & 0 & 0 & 20 \end{bmatrix},\ \mathbf{W}4 = \begin{bmatrix} 7 \\ 8 \\ 0 \\ 0 \\ 2 \\ 36 \end{bmatrix},\ \mathbf{R}4 = \begin{bmatrix} 353 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\mathbf{F5}=\begin{bmatrix}0&470&0&0&0&0\\0&0&404&0&0&0\\0&0&0&211&0&0\\0&0&0&0&78&0\\0&0&0&0&0&80\\0&0&0&0&0&35\end{bmatrix},\ \mathbf{W5}=\begin{bmatrix}1\\2\\15\\0\\7\\28\end{bmatrix},\ \mathbf{R5}=\begin{bmatrix}471&180&0&0&0&0\end{bmatrix},$$

$$\mathbf{F6}=\begin{bmatrix}0&179&0&0&0&0\\0&0&489&0&0&0\\0&0&0&397&0&0\\0&0&0&0&205&0\\0&0&0&0&0&67\\0&0&0&0&0&44\end{bmatrix},\ \mathbf{W6}=\begin{bmatrix}2\\3\\7\\6\\11\\71\end{bmatrix},\ \mathbf{R6}=\begin{bmatrix}181&22&0&0&0&0\end{bmatrix}.$$

Entries in the flow matrices satisfies the flow pattern: $\sum_{\forall j} n_{ij}\left(t\right)=n_{i,i+1}\left(t\right)$, $n_{i-1,i}\left(t-1\right)\geq n_{i,i+1}\left(t\right)$, for $i\in S\setminus\{2,6\}$, $\sum_{\forall j} n_{ji}\left(t\right)=0,\ j>i$, where $n_{ij}\left(t\right)$ for each $i,j\in S$ is the flow from level i to level j within the system during the time period $(t-1\ ,\ t)$, and S is the set of levels in the undergraduate programme.

Among the differential variables of the programme which include tuition fees and charges, environmental factors and promotion criteria, only tuition fees and charges varied during the period under consideration. Thus, tuition fees and charges for the programme are used as an explanatory variable for the variation in wastage. The data for tuition fees and charges are obtained from the Bursary Department of the University (see Table 1).

Table 1. Tuition fees and charges for B.Sc. in Statistics with Computer Science

Time in session	Tuition (in Naira)	Charges (in Naira)		Total (in Naira)	
		Returning students	New students	Returning students	New students
2003/2004	-	-	-	28,700	38,900
2004/2005	-	-	-	28,700	38,900
2005/2006	20,000	8,700	24,700	28,700	44,700
2006/2007	20,000	8,700	24,700	28,700	44,700
2007/2008	20,000	8,700	40,500	28,700	60,500
2008/2009	20,000	9,700	51,500	29,700	71,500

Source: Students Services Division, Bursary Department, University of Benin, Benin City, Nigeria.

By the method described in Osagiede and Ekhosuehi (2006), the HIMC is estimated from the flow matrices as:

$$\mathbf{Q}=\begin{bmatrix}0.0097&0.9903&0&0&0&0\\0.0398&0.0051&0.9551&0&0&0\\0.0323&0.0041&0&0.9636&0&0\\0.0245&0.0031&0&0&0.9724&0\\0.0797&0.0101&0&0&0&0.9102\\0.5420&0.0689&0&0&0&0.3891\end{bmatrix}.$$

In matrix $\mathbf{Q}$, the entries in columns 1 and 2 arise from the replacement matrix and the admission policy of the institution as new entrants are admitted either into Year 1 or Year 2. In particular, the $(1,2)$ entry is the probability that a student is promoted from Year 1 to Year 2 or a student is admitted into Year 2 to replace leavers in Year 1. Wherever zero entry occurs, it means no transition took place between the corresponding levels. The main diagonal elements of the transition matrices are either zero or relatively small, while the upper off-diagonal elements ('promotion' probabilities) are large. The main diagonal elements for columns 1 and 2 are small because the wastage probabilities are also very small, while the upper off-diagonal elements are large because a greater proportion of students are promoted. The diagonal entry in column 6 represents the probability of a student repeating Year 6. The block structure of the transition matrix therefore indicates that there is a normal progression to the next higher level and the few students who drop-out of the programme are being replaced by new entrants. In sum, the entries in matrix $\mathbf{Q}$ provide information on the direct transition between the year of study in the academic programme and the part of wastage replaced by new entrants into the programme in each period.

To verify the constancy of the HIMC, we use the chi-square test statistic as in Zanakis and Maret (1980). We obtain the calculated chi-square value as 393.4455. Since the number of time periods is six, the number of degrees of freedom for the test statistic is 150. This value (150 degrees of freedom) is large so the critical value for α percentile is computed using: $\chi^2_\alpha = \frac{1}{2}(z_\alpha + \sqrt{2k-1})^2$, $k > 30$, where k is the number of degrees of freedom and z_α is the corresponding percentile of the standard normal distribution (Lindgren, 1993). We obtain the critical value at the 5% significance level as $\chi^2_{0.95} = 179.2958$. The calculated chi-square value is greater than the critical value, so we conclude that the transition matrix is not stationary over the period of investigation at 5% significance level. Thus, we estimate the PUTMs from the flow matrices as:

$$\mathbf{Q}1 = \begin{bmatrix} 0 & 1.0000 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1.0000 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1.0000 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1.0000 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1.0000 \\ 0.5364 & 0.0192 & 0 & 0 & 0 & 0.4444 \end{bmatrix};$$

$$\mathbf{Q}2 = \begin{bmatrix} 0.0177 & 0.9823 & 0 & 0 & 0 & 0 \\ 0.1310 & 0.0024 & 0.8667 & 0 & 0 & 0 \\ 0.0721 & 0.0013 & 0 & 0.9266 & 0 & 0 \\ 0.0914 & 0.0016 & 0 & 0 & 0.9070 & 0 \\ 0.0604 & 0.0011 & 0 & 0 & 0 & 0.9385 \\ 0.7246 & 0.0131 & 0 & 0 & 0 & 0.2623 \end{bmatrix};$$

$$\mathbf{Q}3 = \begin{bmatrix} 0.0130 & 0.9870 & 0 & 0 & 0 & 0 \\ 0.1829 & 0.0016 & 0.8155 & 0 & 0 & 0 \\ 0.0548 & 0.0005 & 0 & 0.9447 & 0 & 0 \\ 0.0605 & 0.0005 & 0 & 0 & 0.9389 & 0 \\ 0.0790 & 0.0007 & 0 & 0 & 0 & 0.9204 \\ 0.6887 & 0.0060 & 0 & 0 & 0 & 0.3053 \end{bmatrix};$$

$$\mathbf{Q}4 = \begin{bmatrix} 0.0160 & 0.9840 & 0 & 0 & 0 & 0 \\ 0.1143 & 0.0006 & 0.8851 & 0 & 0 & 0 \\ 0.0395 & 0.0002 & 0 & 0.9603 & 0 & 0 \\ 0.0365 & 0.0002 & 0 & 0 & 0.9633 & 0 \\ 0.0693 & 0.0003 & 0 & 0 & 0 & 0.9304 \\ 0.6722 & 0.0033 & 0 & 0 & 0 & 0.3245 \end{bmatrix};$$

$$\mathbf{Q}5 = \begin{bmatrix} 0.0097 & 0.9903 & 0 & 0 & 0 & 0 \\ 0.0584 & 0.0075 & 0.9341 & 0 & 0 & 0 \\ 0.0458 & 0.0059 & 0 & 0.9483 & 0 & 0 \\ 0.0240 & 0.0031 & 0 & 0 & 0.9730 & 0 \\ 0.0651 & 0.0083 & 0 & 0 & 0 & 0.9265 \\ 0.5386 & 0.0689 & 0 & 0 & 0 & 0.3925 \end{bmatrix};$$

$$\mathbf{Q}6=\begin{bmatrix} 0.0097 & 0.9903 & 0 & 0 & 0 & 0 \\ 0.0398 & 0.0051 & 0.9551 & 0 & 0 & 0 \\ 0.0323 & 0.0041 & 0 & 0.9636 & 0 & 0 \\ 0.0245 & 0.0031 & 0 & 0 & 0.9724 & 0 \\ 0.0797 & 0.0101 & 0 & 0 & 0 & 0.9102 \\ 0.5420 & 0.0689 & 0 & 0 & 0 & 0.3891 \end{bmatrix}.$$

The PUTMs indicate that in the first period, admission was done to replace leavers in Year 6 only, while for the remaining periods, $t=2,\cdots,6$, admission was done to replace leavers at each year of study. Although, the PUTM provide information on the direct transition between the year of study in the programme and the part of wastage replaced by new entrants into the programme in each period by updating the existing data, yet its 'sparse' block structure limits its use for non-equidistant time epochs.

Consider the data in Table 1. To account for the fees of new entrants into Year 2, we compute the fees paid in Year 2 as the sum of the proportion of fees paid by new entrants and that paid by the returning students. Using the information in the flow matrices and Table 1, we estimate the parameters for the binomial logistic model as presented in Table 2.

Table 2. Parameter estimates for the binomial logistic model

Parameters	Year 1 (i =1)	Year 2 (i =2)	Year 3 (i =3)	Year 4 (i =4)	Year 5 (i =5)	Year 6 (i =6)
β_{1i}^*	-2.3876	10.6617	29.4068	-5.0816	-23.4219	0.4070
β_{2i}^*	0.0000	-0.0004	-0.0011	0.0001	0.0007	0.0000

Source: Authors computation

The results $\beta_{1i}^*\neq 0$ for $i=1,\cdots,6$, indicate that the link function for wastage rate is partly constant. For $\beta_{2i}^*<0$ ($i=2,3$) there is an inverse relationship between the link function and the tuition fees and charges; and there is a direct relationship for $\beta_{2i}^*>0$ ($i=4,5$). $\beta_{2i}^*=0$ ($i=1,6$) implies the absence of a relationship. Nonetheless, we compute the wastage probabilities using the estimator in equation (16). By coding the sessions $t=2003/2004,\ldots,2008/2009$ as $t=1,\ldots,6$, and using the parameter estimates of the logistic model, we estimate the wastage probabilities as shown in Table 3.

Table 3. Estimates of wastage probabilities for the NHETM

	1	2	3	4	5	6
Year 1	0.0185	0.0185	0.0158	0.0158	0.0102	0.0075
Year 2	0.0746	0.0920	0.0920	0.0920	0.0025	0.0415
Year 3	0.0518	0.0518	0.0518	0.0518	0.0518	0.0186
Year 4	0.0276	0.0276	0.0276	0.0276	0.0276	0.0289
Year 5	0.0735	0.0735	0.0735	0.0735	0.0735	0.1410
Year 6	0.5556	0.8140	0.6176	0.6429	0.4444	0.6174

Source: Authors computation.

Similarly, we obtain the transition probabilities. Accordingly, we estimate the NHETMs for each of the six-year period as:

$$\mathbf{Q}(1)\Big|_{\mathbf{x}'(1)}=\begin{bmatrix} 0.0179 & 0.9821 & 0. & 0 & 0 & 0 \\ 0.0720 & 0.0026 & 0.9254 & 0 & 0 & 0 \\ 0.0500 & 0.0018 & 0 & 0.9482 & 0 & 0 \\ 0.0266 & 0.0010 & 0 & 0 & 0.9724 & 0 \\ 0.0710 & 0.0025 & 0 & 0 & 0 & 0.9265 \\ 0.5866 & 0.0209 & 0 & 0 & 0 & 0.3925 \end{bmatrix},$$

$$\mathbf{Q}(2)\Big|_{\mathbf{x}'(2)} = \begin{bmatrix} 0.0185 & 0.9815 & 0 & 0 & 0 & 0 \\ 0.0920 & 0 & 0.9080 & 0 & 0 & 0 \\ 0.0518 & 0 & 0 & 0.9482 & 0 & 0 \\ 0.0276 & 0 & 0 & 0 & 0.9724 & 0 \\ 0.0735 & 0 & 0 & 0 & 0 & 0.9265 \\ 0.6075 & 0 & 0 & 0 & 0 & 0.3925 \end{bmatrix},$$

$$\mathbf{Q}(3)\Big|_{\mathbf{x}'(3)} = \begin{bmatrix} 0.0158 & 0.9842 & 0. & 0 & 0 & 0 \\ 0.0920 & 0 & 0.9080 & 0 & 0 & 0 \\ 0.0518 & 0 & 0 & 0.9482 & 0 & 0 \\ 0.0276 & 0 & 0 & 0 & 0.9724 & 0 \\ 0.0735 & 0 & 0 & 0 & 0 & 0.9265 \\ 0.6075 & 0 & 0 & 0 & 0 & 0.3925 \end{bmatrix},$$

$$\mathbf{Q}(4)\Big|_{\mathbf{x}'(4)} = \begin{bmatrix} 0.0185 & 0.9815 & 0 & 0 & 0 & 0 \\ 0.0920 & 0 & 0.9080 & 0 & 0 & 0 \\ 0.0518 & 0 & 0 & 0.9482 & 0 & 0 \\ 0.0276 & 0 & 0 & 0 & 0.9724 & 0 \\ 0.0735 & 0 & 0 & 0 & 0 & 0.9265 \\ 0.6075 & 0 & 0 & 0 & 0 & 0.3925 \end{bmatrix},$$

$$\mathbf{Q}(5)\Big|_{\mathbf{x}'(5)} = \begin{bmatrix} 0.0076 & 0.9924 & 0. & 0 & 0 & 0 \\ 0.0019 & 0.0006 & 0.9975 & 0 & 0 & 0 \\ 0.0387 & 0.0131 & 0 & 0.9482 & 0 & 0 \\ 0.0206 & 0.0070 & 0 & 0 & 0.9724 & 0 \\ 0.0549 & 0.0186 & 0 & 0 & 0 & 0.9265 \\ 0.4535 & 0.1540 & 0 & 0 & 0 & 0.3925 \end{bmatrix},$$

$$\mathbf{Q}(6)\Big|_{\mathbf{x}'(6)} = \begin{bmatrix} 0.0067 & 0.9933 & 0 & 0 & 0 & 0 \\ 0.0370 & 0.0045 & 0.9585 & 0 & 0 & 0 \\ 0.0166 & 0.0020 & 0 & 0.9814 & 0 & 0 \\ 0.0258 & 0.0031 & 0 & 0 & 0.9711 & 0 \\ 0.1257 & 0.0153 & 0 & 0 & 0 & 0.8590 \\ 0.5505 & 0.0669 & 0 & 0 & 0 & 0.3826 \end{bmatrix}.$$

The entries in column j, $j = 1,2$, of $\mathbf{Q}(\varsigma)\Big|_{\mathbf{x}'(\varsigma)}$, $\varsigma = 1,\cdots,6$, provide information on the probability that losses in the system would result to a consequential admission into Year j. The upper off-diagonal entries in column j, $2 \le j \le 6$, give information on the promotion probabilities, while the diagonal entry in column 6 provides information on the repetition rates. Unlike the PUTMs, the NHETMs reflect the period of no direct entry admission into the programme as in the flow matrices. More so, the six-step higher-order transition matrices obtained as

$$\prod_{\varsigma=1}^{6}\mathbf{Q}(\varsigma)\Big|_{\mathbf{x}'(\varsigma)} = \begin{bmatrix} 0.4262 & 0.1003 & 0.0474 & 0.0474 & 0.0746 & 0.3041 \\ 0.1905 & 0.3826 & 0.1854 & 0.0305 & 0.0407 & 0.1703 \\ 0.1024 & 0.1693 & 0.5495 & 0.0701 & 0.0224 & 0.0862 \\ 0.0535 & 0.0747 & 0.2418 & 0.5308 & 0.0593 & 0.0399 \\ 0.0429 & 0.0527 & 0.1530 & 0.2277 & 0.4644 & 0.0593 \\ 0.0865 & 0.0419 & 0.0951 & 0.1460 & 0.2067 & 0.4238 \end{bmatrix},$$

removes 'sparsity' from the higher-order form of the NHETM. Since the higher-order form of the PUTM is sparse, it cannot be imbedded in continuous-time (Singer and Spilermann, 1976), whereas the higher-order NHETM can be imbedded in continuous-time. Therefore, the NHETM can be used for long-term projection for any time horizons. In this regard, the trauma that would have emanated from modifying a transition model so as to accommodate unequal time epochs (arising from distortions to a proposed academic calendar of the educational system) are ameliorated as a stationary continuous-time transition matrix can now be obtained from the NHETM using any appropriate regularization techniques (Kreinin and Sidelnikova, 2001).

5. CONCLUSION

In this study, we propose the NHETM with a binomial logistic wastage rate for the multi-echelon educational system. The estimators for the parameters of the NHETM wherein the school's differential variables are inscribed are novel to the Markovian evaluation of the educational system. Using datasets from a university programme, we illustrate the estimation of the NHETMs vis-a-vis the HIMC and the PUTM. The merits of using the NHETM over other existing Markov chains for enrolment projection have been highlighted to include the imbedding potentials with a view to making probabilistic statements about the future structure of the system for any time horizons. The practical implication of estimating the NHETM is the collection of new data on the school's differential variables.

ACKNOWLEDGEMENTS

The authors are very grateful to the anonymous referees for their valuable comments and suggestions to improve the paper in the present form.

REFERENCES

1. Ekhosuehi, V.U. (2009). Entropy of a state-transition model for organizational systems. *M.Sc.Thesis* Department of Mathematics, University of Benin, Benin City, Nigeria.
2. Dobson, A.J. (2002). *An Introduction to Generalized Linear Models*. Chapman & Hall/CRC, Boca Raton, Florida.
3. Gani, J. (1963). Formulae for projecting enrolments and degrees awarded in universities. *Journal of the Royal Statistical Society, Series A (General)*, 126(3):400-409.
4. Hardle, W., Muller, M., Sperlich, S., and Werwatz, A. (2004). *Nonparametric and Semiparametric Models*. Springer-Verlag, Berlin Heidelberg.
5. Ibe, C.O. (2009). *Markov Processes for Stochastic Modelling*. Elsevier Academic Press, Burlington, USA.
6. Jimenez, J.D. and Salas-Velasco, M. (2000). Modeling educational choices: A binomial logit model applied to the demand for higher education. *Higher Education*, 40(3):293-311.
7. Kreinin, A. and Sidelnikova, M. (2001). Regularization algorithms for transition matrices. *Algo Research Quarterly*, 4(1/2): 23-40.
8. Lindgren, B.W. (1993). *Statistical theory* (4th ed.). Chapman & Hall, New York.
9. Nicholls, M.G. (2009). The use of Markov models as an aid to the evaluation, planning and benchmarking of doctoral programs. *Journal of the Operational Research Society*, 60:1183-1190.
10. Osagiede, A.A. and Ekhosuehi, V.U. (2006). Markovian approach to school enrolment projection process. *Global Journal of Mathematical Sciences*, **5**(1):1-7.
11. Osagiede, A.A. and Omosigho, S.E. (2004). School enrolment projection using spreadsheet under scanty data. *Nigerian Annals of Natural Sciences*, **5**(2):87-99.
12. Singer, B. and Spilerman, S. (1976). The representation of social processes by Markov models. *The American Journal of Sociology*, 82(1):1-54.
13. Uche, P.I. (1978). On stochastic models for educational planning. *International Journal of Mathematical Education in Science and Technology*, 9(3):333-342.
14. Zanakis, S.H. And Maret, M.W. (1980). A Markov chain application to manpower supply planning. *Journal of the Operational Research Society*, 31:1095-1102.