Original research paper

Homogeneity Versus Parsimony in Markov Manpower Models: A Hidden Markov Chain Approach

ABSTRACT

Homogeneity and parsimony are two important properties considered in most statistical models. Interestingly, in statistical manpower models, the more homogeneous a model is made to be the more it tends not to be parsimonious. In this paper, the comparative importance of these two properties are investigated for Markov manpower models. The mover-stayer principle and its extensions are employed to incorporate hidden classes in the model to achieve more homogeneity and this is compared with the model without the hidden classes, which is more parsimonious, using Likelihood ratio statistic, Akaike Information Criterion and Bayesian Information Criterion. The illustration shows a case of manpower data where, up to a certain level of hidden states, homogeneity is more important than parsimony.

Keywords: *Statistical manpower planning, hidden Markov model, homogeneity, parsimony*

1. INTRODUCTION

The use of Markov chain in functional modelling of manpower systems, which has its foundation on the work by Seal [1], is supported by the structural configuration of manpower systems dynamics. A manpower system is modelled in Markov framework as several aggregates of members of the system, where each aggregate of members is taken as a class or state of the Markov chain with interstate transitions of members and transitions to external environment governed by some probability laws. For a Markov manpower model to be adequately formulated, in terms of giving reliable estimates of the population parameters, members of each class have to be as homogeneous as possible with respect to state transitions [2-5]. If this is not the case then the problem of heterogeneity sets in. The general method of tackling the problem of heterogeneity to achieve homogeneity in the personnel classes of a manpower system is by the process of disaggregation [6,3]. In the case of observable heterogeneity [7], where personnel classes can be ascertained on the basis of observable data, disaggregation involves external subdivision of classes of the system to achieve the desired homogeneous personnel groupings. In the case of hidden heterogeneity [4,8], where personnel classes can no longer be completely ascertained on the basis of observable data, disaggregation involves internal subdivision of classes of the system to achieve the desired homogeneous personnel groupings.

In all cases, diaggregation involves the inclusion of more classes of members of manpower systems and therefore more parameters in the model. It is therefore opposed to the principle of parsimony, which advocates the inclusion of adequate minimum number of parameters in a model [9]. In other words, an attempt to achieve a better model with respect to class homogeneity in a manpower system may end up yielding a model less parsimonious than the original model. Yet, the need to ensure the consideration of these properties in building a manpower model has been emphasized by researchers in this area. For instance, Guerry and De Feyter [10] emphasize that dividing the personnel of a manpower system into more subgroups may result to a more homogeneous manpower model, but with additional problem in parameter estimation; Bartholomew et al. [2], De Feyter [3], Guerry [4] and Udom and Ebedoro [8], on the other hand, emphasize that lack of class homogeneity in manpower models may lead to unreliable parameter estimates. The side of the argument to be taken in manpower planning calls for validation of manpower models based on model performance with respect to class homogeneity and parsimony in parameter inclusion. This paper, therefore, focuses on investigating the comparative importance of the properties of homogeneity and parsimony in a manpower model under Markov framework.

Ugwuowo and McClean [7], Guerry [4] and Udom and Ebedoro [8] all identify two types of heterogeneity in manpower systems as observable and hidden heterogeneity. Observable heterogeneity occurs where an observable class of manpower personnel includes members that differ significantly with respect to probabilities of inter-class transitions. As stated above, the problem of observable heterogeneity can be tackled by splitting the class to include those with similar transition patterns in the same sub-class. Actually, in Markov manpower modelling the states of the Markov chain are from onset chosen, according to the observed manpower flow data for a given manpower system, to reflect homogeneity. On the other hand, the sources of hidden heterogeneity, such as innate traits of individual personnel, are not apparent from mere observation of manpower data. The problem of hidden heterogeneity is, therefore, handled differently. To tackle the problem of hidden heterogeneity in manpower models, Guerry [4] and Udom and Ebedoro [8] use the mover-stayer principle introduced by Blumen et al. [11] and its extensions by Spilerman [12] to incorporate sub-classes within the observable classes of the manpower system. Within each observable class each hidden sub-class holds manpower personnel homogeneous with respect to hidden sources of heterogeneity and hence probability of transiting to other observable classes. Guerry [4] includes two sub-classes for 'movers' and 'stayers' in each observable classes, where movers are identified by higher probability of transiting to other observable classes than the stayers. Following the work by Spilerman [12], who introduced the inclusion of more than two states in the mover-stayer principle, Udom and Ebedoro [8] include three sub-classes for 'movers', 'mediocres' and 'stayers' in each observable classes of a manpower system, where movers are identified by highest probability of transiting to other observable classes, followed by mediocres and then the stayers. In this paper we include four sub-classes for 'high movers', 'movers', 'mediocres' and 'stayers' in each observable classes, and five sub-classes for 'high movers', 'movers', 'above mediocres', 'mediocres' and 'stayers' in each observable classes where, in each of the two cases, high movers are identified by highest probability of transiting to other observable classes, followed in order by the other categories. In this way we establish up to five different hidden Markov model (HMM) types for the manpower system: HMM1, HMM2, HMM3, HMM4 and HMM5 corresponding to 1, 2, 3, 4 and 5 hidden sub-classes per observable class respectively, where a HMM is a bivariate stochastic process $\{(X_t, Y_t)\}$ with $\{X_t\}$ being an unobserved Markov chain in the states of the observable process $\{Y_t\}$, with the distribution of Y_t depending on X_t[8]. We note that HMM1 corresponds to the classical Markov manpower model where there is no observable class subdivision for latent heterogeneity. Also, by the foregoing discussion, HMM1 is more parsimonious and less homogeneous than the rest of the models; HMM2 is more parsimonious and less homogeneous than HMM3, HMM4 and HMM5; HMM3 is more parsimonious and less homogeneous than HMM4 and HMM5; and HMM4 is more parsimonious and less homogeneous than HMM5.

Hidden Markov models have been widely applied in other areas of research such as time series, econometrics, finance, biology and psychology [13-16]. Its application in the area of statistical manpower planning has, however, been scanty, just about those cited in this paper. Yet, it seems to be a promising model for unraveling some intricate features of manpower systems dynamics.

In this paper the manpower hidden Markov model of type HMMk' where k' can take values 1, 2, 3, 4 and 5, is formulated, as a multinomial hidden Markov model [4,8], on the basis of which the HMM′s described above are established for comparison. The parameters of the models are estimated using the Expectation-Maximization (EM) algorithm [17,18], and the performance of the models on the same manpower data compared using appropriate statistical tests. In this way the property that is more important in Markov manpower modelling can be inferred.

2. FORMULATION OF THE MANPOWER MODEL OF TYPE HMMk'

Let there be k observable classes of personnel, observable on the basis of available data, in a manpower system. All manpower personnel in each of the k classes are assumed to have transition patterns, across other classes, with the same probability distribution. Let these classes be denoted by $C_1, \ldots, C_k$. Additionally, let C_{k+1}represent the class of those who leave the system to the outside environment (wastage class). In other words, the rate at which all members of C_1, for example, move independently to any of the other k classes, $C_2, \ldots, C_{k+1}$, is assumed to be affected the same way by the observable sources of heterogeneity.

In a general simple Markov manpower model [19], the dynamics of the manpower flows are governed by a recruitment vector $\boldsymbol{R}$, a sub-stochastic transition matrix $\boldsymbol{P}$ and a wastage vector $\boldsymbol{W}$, such that the evolution of the manpower stock vector, $\boldsymbol{n}(t)$, at time t can be expressed in terms of the immediate past stock vector and these model parameters as

$$\boldsymbol{n}(t) = \boldsymbol{n}(t-1)\boldsymbol{P} + \boldsymbol{n}(t-1)\boldsymbol{W}'\boldsymbol{R}(t) \qquad (2.1)$$

In (2.1), $\boldsymbol{n}(t) = [n_1(t), \ldots, n_k(t)]$ and $n_i(t)$ is the number of workers in C_i at time t;

$$\boldsymbol{P} = \begin{matrix} & C_1 & \cdots & C_k \\ C_1 & p_{11} & \cdots & p_{1k} \\ \vdots & \vdots & \cdots & \vdots \\ C_k & p_{k1} & \cdots & p_{kk} \end{matrix}$$

and p_{ij} is the transition probability of moving from C_i to C_j;

$\boldsymbol{W} = [p_{1,k+1}, \ldots, p_{k,k+1}]$, $p_{i,k+1}$ is the wastage probability of moving from C_i to C_{k+1};

$\boldsymbol{R}(t) = [p_{01}(t), \ldots, p_{0k}(t)]$, $p_{0i}(t)$ is the probability of making a recruitment into C_i at t.

The model in (2.1) is required for analysis if all the goals of statistical manpower planning, which are description, prediction or forecasting and control of manpower structure, are to be achieved. Since the two types of heterogeneity in manpower models concern only the promotion and wastage flows, the points of focus in building the hidden Markov models of interest in the current work are the elements of $\boldsymbol{P}$ and $\boldsymbol{W}$.

Let a stochastic process whose states are the classes of the manpower system, $C_1, \ldots, C_k$, be defined as $\{Y_t\}$. Then $\{Y_t\}$ possesses the first order Markov property such that

$$p_{ij} = \mathrm{P}(Y_{t+1} = C_j | Y_t = C_i)$$
$$\text{for } i = 1, \ldots, k;\ j = 1, \ldots, k+1 \qquad (2.2)$$

$\{Y_t\}$ is actually a Markov chain with one absorbing state, C_{k+1}. This means that no personnel moves back to any of the k classes, $C_1, \ldots, C_k$, after leaving to the outside environment (wastage). For the hidden classes, let each of the k observable classes, C_i $(i = 1, \ldots, k)$, be subdivided into k' sub-classes, $H_l^i, l = 1, \ldots, k'$. The value of k' and what each sub-class stands for in the models, as discussed in section 1, are as follows:

For HMM5 : $k' = 5$, H_1^i = high movers class, H_2^i = movers class, H_3^i = above mediocres class, H_4^i = mediocres class, H_5^i = stayers class; for HMM4: $k' = 4$, H_1^i = high movers class, H_2^i = movers class, H_3^i = mediocres class, H_4^i = stayers class; for HMM3: $k' = 3$, H_1^i = movers class, H_2^i = mediocres class, H_3^i = stayers class; for HMM2: $k' = 2$, H_1^i = movers class, H_2^i = stayers class; for HMM1: $k' = 1$, H_1^i = observable class. Let the ordering relationships 'is less homogeneous than' be denoted by $<_h$ and 'is less parsimonious than' by $<_p$. Then, from the foregoing, as far as the models are valid, HMM1$<_h$ HMM2 $<_h$ HMM3 $<_h$ HMM4 $<_h$ HMM5 and HMM5 $<_p$ HMM4 $<_p$ HMM3 $<_p$ HMM2 $<_p$ HMM1

Let $\{X_t^i\}$ be the underlying Markov chain within each observable class C_i $(i = 1,\ldots,k)$ having its states as the $k^{\prime}$ sub-classes, $H_l^i, l = 1,\ldots,k^{\prime}$. The transition probabilities of the underlying Markov chain $\{X_t^i\}$ within C_i can be given as

$\eta_{lm}^i = \mathrm{P}(X_{t+1}^i = H_m^i | X_t^i = H_l^i); l, m = 1,\ldots,k^{\prime}$.

The transition probability matrix of $\{X_t^i\}$ is the $k^{\prime} \times k^{\prime}$ matrix

$$\eta^i = \begin{array}{c} H_1^i \\ \vdots \\ H_{k'}^i \end{array} \overset{\begin{array}{ccc} H_1^i & \cdots & H_{k'}^i \end{array}}{\begin{pmatrix} \eta_{11}^i & \cdots & \eta_{1k'}^i \\ \vdots & \cdots & \vdots \\ \eta_{k'1}^i & \cdots & \eta_{k'k'}^i \end{pmatrix}}$$

The dependence of the distribution of Y_t on X_t is expressed in the conditional probability of personnel transition from any state $H_l^i, l = 1,\ldots,k^{\prime}$, within C_i to another observable class C_j given by

$q_{lj}^i = \mathrm{P}(Y_{t+1} = C_j | Y_t = C_i, X_t^i = H_l^i) = \mathrm{P}(Y_{t+1} = C_j | X_t^i = H_l^i)$

Let the observed manpower flow from C_i to each of C_j $(j = 1,\ldots,k+1)$ from time period t *to* $t+1$ be denoted by $n_{ij}(t)$. When $j = 1,\ldots,k$ the manpower flow is promotion or demotion; when $j = k+1$ the manpower flow is wastage. For the observable class C_i and a given t, the random events of making the number of transitions $n_{i1}(t),\ldots,n_{ik}(t), n_{i,k+1}(t)$ are exhaustive (and exclusive) for all such transitions from C_i. Conditional on the hidden Markov chain being on any of the sub-classes $(X_t^i = H_l^i)$, the probability of each of these events corresponds to q_{lj}^i $(j = 1,\ldots,k+1)$ so that $\sum_{j=1}^{k+1} q_{lj}^i = 1$. In other words, at time t and conditional on $X_t^i = H_l^i$, the probability distribution of any random vector M_t^i whose realization is a vector of these observed and exhaustive manpower flow numbers $v_i(t) = (n_{i1}(t),\ldots,n_{ik}(t), n_{i,k+1}(t))$ is multinomial. In other words, given $\Sigma_i^t = \sum_{j=1}^{k+1} n_{ij}(t)$, the conditional probability $Q_{l,v_i(t)}^i = \mathrm{P}(M_t^i = v_i(t) | X_t^i = H_l^i)$ is

$$Q_{l,v_i(t)}^i = \binom{\Sigma_i^t}{n_{i1}(t),\ldots,n_{i,k+1)}(t)} (q_{l1}^i)^{n_{i1}(t)} \ldots (q_{lk}^i)^{n_{ik}(t)} . (q_{l,k+1}^i)^{n_{i,k+1}(t)}$$
$$l = 1,\ldots,k^{\prime}; t = 1,\ldots,T \qquad (2.3)$$

3. ESTIMATION OF THE MODEL PARAMETERS

One major procedure in hidden Markov models is the estimation of model parameters. In the case of HMM1, where the classes are the observable classes C_i $(i = 1,\ldots,k)$, and with $N_i(t) = \sum_{j=1}^{k+1} n_{ij}(t)$, maximum likelihood estimation (MLE) method gives the estimator of the model parameter p_{ij} as (see, for instance, [4])

$$\hat{p}_{ij} = \frac{\sum_{t=1}^{T} n_{ij}(t)}{\sum_{t=1}^{T} N_i(t)} \qquad (3.1)$$

Some other procedures are used for estimating the parameters of HMM's where latent classes exist in the models. One of such procedures is the EM algorithm. Both Guerry [4] and Udom and Ebedoro [8] have used this algorithm in Markov manpower models. Guerry [4] used it for the case of two hidden states while Udom and Ebedoro [8]extended it to the case of three hidden states but in a hierarchical manpower system. The use of the algorithm for the case of **HMM**$k^{\prime}$ is, therefore, a straightforward extension by including $k^{\prime}$ hidden states.

Consider the estimation of the parameters η_{lm}^i, $Q_{l,v_i(t)}^i$ and q_{lj}^i $(l, m = 1,\ldots,k^{\prime};\ t = 1,\ldots,T;\ j = 1,\ldots,k+1)$ of HMM$k^{\prime}$. These parameters can be assembled in matrices :

$\eta^i = (\eta^i_{lm})$, $Q^i = (Q^i_{l,v_i(t)})$ and $q^i = (q^i_{lj})$. The EM re-estimation algorithm is executed in two steps: the Expectation step and the Maximization step. The Expectation step utilizes the forward and the backward probabilities $u^i_l(t) = \mathrm{P}(M^i_1 = v_i(1),\ldots,M^i_t = v_i(t), X^i_t = H^i_l)$ and $d^i_l(t) = \mathrm{P}(M^i_{t+1} = v_i(t+1),\ldots,M^i_T = v_i(T) \,|\, X^i_t = H^i_l)$ satisfying, for the forward probabilities

$$u^i_l(t) = \begin{cases} \pi^i_l Q^i_{l,v_i(1);}\ l = 1,\ldots,k',\ t = 1 \\ \sum_{m=1}^{k'} u^i_m(t-1)\,\eta^i_{ml} Q^i_{l,v_i(t)}\,; l = 1,\ldots,k',\ \ t = 2,\ldots,T \end{cases} \quad (3.2)$$

and for the backward probabilities

$$d^i_l(t) = \begin{cases} 1;\ l = 1,\ldots,k';\ t = T \\ \sum_{m=1}^{k'} d^i_m(t+1)\,\eta^i_{lm} Q^i_{m,v_i(t+1)}\,;\ \ l = 1,\ldots,k',\ t = T-1,\ldots,1. \end{cases} \quad (3.3)$$

The term π^i_l in (3.2) will turn out to be the initial probability distribution of one of the two probability components of the re-estimation formulas. One of these two components expresses the probability of the hidden process X^i_t being in state H^i_l at t, given that the sequence of the observed data is $M^i_1 = v_i(1),\ldots,M^i_T = v_i(T)$; this is represented as $\beta^i_l(t) = \mathrm{P}(X^i_t = H^i_l \,|\, M^i_1 = v_i(1),\ldots,M^i_T = v_i(T))$. The second component expresses the probability of the hidden process X^i_t being in state H^i_l at t, and then moving to state H^i_m in the next transition given that the sequence of the observed data is $M^i_1 = v_i(1),\ldots,M^i_T = v_i(T)$; this is represented as $\gamma^i_l(t) = \mathrm{P}(X^i_t = H^i_l, X^i_{t+1} = H^i_m \,|\, M^i_1 = v_i(1),\ldots,M^i_T = v_i(T))$. The two components $\beta^i_l(t)$ and $\gamma^i_{lm}(t)$ are then expressible in terms of the forward and backward probabilities as

$$\beta^i_l(t) = \frac{u^i_l(t)d^i_l(t)}{\sum_{l=1}^{k'} u^i_l(t)d^i_l(t)} \quad (3.4)$$

and

$$\gamma^i_{lm}(t) = \frac{u^i_l(t)\eta^i_{lm}d^i_m(t+1)Q^i_{m,v_i(t+1)}}{\sum_{l=1}^{k'}\sum_{m=1}^{k'} u^i_l(t)\eta^i_{lm}d^i_m(t+1)Q^i_{m,v_i(t+1)}} \quad (3.5)$$

Next, we consider the likelihood of manpower flows from C_i $(i = 1,\ldots,k)$being a specified sequence of observations. For $t = 1,\ldots,T$, the joint likelihood of having $M^i_1 = v_i(1),\ldots,M^i_T = v_i(T)$ is given by $L^i_T = \mathrm{P}(M^i_1 = v_i(1),\ldots,M^i_T = v_i(T) \,|\, \pi^i,\eta^i,q^i)$, where π^i is the initial distribution vector of π^i_l, $l = 1,\ldots,k'$. Hence,

$L^i_T = \mathrm{P}(M^i_1 = v_i(1),\ldots,M^i_T = v_i(T) \,|\, \pi^i,\eta^i,q^i)$

Which gives

$$L^i_T = \sum_{l=1}^{k'}[\mathrm{P}(X^i_1 = H^i_l \,|\, \pi^i)\prod_{t=2}^{T}\mathrm{P}(X^i_t = H^i_m \,|\, X^i_{t-1} = H^i_l,\eta^i)\prod_{t=1}^{T}\mathrm{P}(M^i_t = v_i(t) \,|\, X^i_t = H^i_l,q^i)] \quad (3.6)$$

Equation (3.6) can be resolved to obtain the expected log-likelihood as

$$\mathrm{ELog}L^i_T = \sum_{l=1}^{k'}\beta^i_l(1)\log\pi^i_l + \sum_{t=2}^{T}\sum_{l=1}^{k'}\sum_{m=1}^{k'}\gamma^i_{lm}(t)\log\eta^i_{lm} + \sum_{t=1}^{T}\sum_{l=1}^{k'}\beta^i_l(t)\log Q^i_{l,v_i(t)} \quad (3.7)$$

The final step is to maximize (3.7) with respect to π_l^i, η_{lm}^i and $Q_{l,v_i(t)}^i$. By the method of Lagrange multipliers (3.7) is maximized taking the three ditinct parts saparetly with the respective constraints on the parameters to obtain the following formulas.

$$\pi_l^i = \beta_l^i(1), l = 1, \ldots, k' \quad (3.8)$$

$$\eta_{lm}^i = \frac{\sum_{t=2}^{T} \gamma_{lm}^i(t)}{\sum_{m=1}^{k'} \sum_{t=2}^{T} \gamma_{lm}^i(t)}, \quad l, m = 1, \ldots, k' \quad (3.9)$$

$$q_{lj}^i = \frac{\sum_{t=2}^{T} \beta_l^i(t)\, n_{ij}(t)}{\sum_{t=1}^{T} \beta_l^i(t)\, N_i(t)}, \quad l = 1, \ldots, k', \quad j = 1, \ldots, k+1 \quad (3.10)$$

The above re-estimation algorithm is an iterative procedure which, given any sequence of observations of a manpower system, $v_i(1), \ldots, v_i(T)$, begins by choosing initial values for the parameters π_l^i, η_{lm}^i and q_{ljr}^i. These, by (2.3), are used to realize the corresponding initial values for $Q_{l,v_i(t)}^i$. The estimation formulas are implemented during each iteration to obtain current estimates of the parameters. The process terminates at the convergence of the parameter estimates, when (3.7) is maximized.

4. COMPARISON OF MANPOWER MODELS OF TYPE HMMk'

In this section, the statistics that form the bases upon which manpower models of type HMMk' can be compared are presented. Three such statistics that have been used in comparing Markov manpower models are Likelihood ratio statistic (L_r), Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) [4,8]. The likelihood ratio test compares the adequacy of two models HMMl and HMMm on the basis of L_r distributed as

$$L_r = -2\log\left(\frac{L_{\mathrm{HMM}l}}{L_{\mathrm{HMM}m}}\right) \sim \chi_\alpha^2(\nu) \quad (4.1)$$

In (4.1), $L_{\mathrm{HMM}l}$ and $L_{\mathrm{HMM}m}$ are the likelihood of HMMl and HMMm respectively. ν is the degrees of freedom of the chi square distribution. The L_r test is used to test the hypothesis that the two models fit the data equally well; this is rejected if $L_r > \chi_\alpha^2(\nu)$. The rejection of the hypothesis of equality of good fit implies that HMMm fits the data better than HMMl. ν is computed as

$$\nu = (\text{number of free parameters of HMM}m - \text{number of free parameters of HMM}l).$$

The degrees of fredom, ν, is computed for each pair of HMM's compared. For example, to compare HMM1 and HMM2, each of the k classes of the HMM1 contributes $(k + 1 - 1) = k$ free parameters, which altogether, for the k classes, gives k^2 free parameters from HMM1. For the HMM2, each C_i $(i = 1, \ldots, k)$ has two hidden classes and $k + 1$osbserved destination classes for any transition from C_i giving $2(k + 1 - 1) = 2k$ free parameters of the transition probabilities. The transitions within the hidden classes and the initial states respectively give $2(2 - 1)$ and $(2 - 1)$ free parameters. Therefore each C_i $(i = 1, \ldots, k)$ contributes $2k + 3$ free parameters, giving altogether $k(2k + 3)$ free parameters from HMM2. Then $\nu = k(2k + 3) - k^2 = k^2 + 3k$. In this way ν is computed for all models compared in this paper.

The AIC and BIC are defined as follows [8].

$$\mathrm{AIC} = -2\log L_{\mathrm{HMM}l} + 2P \quad (4.2)$$

$$\mathrm{BIC} = -2\log L_{\mathrm{HMM}l} + P\log H \quad (4.3)$$

In (4.2) and (4.3) P and H are the number of parameters in the model, HMMl, fitted and the number of observations respectively. The model that has the smallest values of AIC and BIC is selected as the one with the best fit. AIC and BIC are used in this paper for further confirmation of the results from the L_r test.

4.1. NUMERICAL ILLUSTRATION

The foregoing developments are illustrated with a university senior academic manpower flow data presented in Table 1. In the data only the observed flow data of senior academic employees comprising senior lecturers, readers and professors with classes C_1, C_2 and C_3respectively are presented, which, from the nature of the manpower system, represent the cadres where hidden heterogeneity in manpower flows is assumed to be more pronounced. C_4 corresponds to C_{k+1}, the class of those who leave the system through wastage. The manpower data covers 8 time periods, $t = 1,...,8$. In Table 1, the first entry, 807, for example, corresponds to $n_{11}(1)$ which gives the number of senior lecturers who remained in the same rank after the first time period; $n_{14}(6) = 52$, $n_{33}(2) = 534$ and so on. The EM re-estimation algorithm is employed on the data, using depmixS4 R package, to estimate the transition probabilities for the five model cases. The transition probability matrices P_1, P_2, P_3, P_4 and P_5 are for $\text{HMM}_1, \text{HMM}_2, \text{HMM}_3, \text{HMM}_4$ and HMM_5 respectively. The results in Table 2 are calculated based on the results from the re-estimation algorithm and the statistical test formulas for L_r, AIC and BIC statistics.

Table 1: A university senior academic manpower flow data

	C_1	C_2	C_3	C_4
C_1: $t = 1$	807	50	40	25
$t = 2$	801	102	20	34
$t = 3$	788	81	18	20
$t = 4$	794	32	34	41
$t = 5$	820	72	27	37
$t = 6$	815	42	36	52
$t = 7$	826	61	30	45
$t = 8$	840	55	45	30
C_2: $t = 1$	0	182	31	10
$t = 2$	0	201	30	10
$t = 3$	0	194	28	12
$t = 4$	0	198	30	12
$t = 5$	0	211	35	21
$t = 6$	0	190	21	15
$t = 7$	0	185	42	26
$t = 8$	0	205	37	28
C_3: $t = 1$	0	0	560	56
$t = 2$	0	0	534	50
$t = 3$	0	0	521	54
$t = 4$	0	0	550	46
$t = 5$	0	0	578	32
$t = 6$	0	0	570	64
$t = 7$	0	0	540	93
$t = 8$	0	0	548	71

$$
P_5 = \begin{array}{c|cccc}
 & C_1 & C_2 & C_3 & C_4 \\
\hline
 & 0.859 & 0.063 & 0.031 & 0.047 \\
 & 0.858 & 0.075 & 0.028 & 0.039 \\
C_1 & 0.871 & 0.055 & 0.045 & 0.029 \\
 & 0.852 & 0.098 & 0.020 & 0.029 \\
 & 0.872 & 0.040 & 0.038 & 0.050 \\
\hdashline
 & 0 & 0.746 & 0.151 & 0.103 \\
 & 0 & 0.841 & 0.093 & 0.066 \\
C_2 & 0 & 0.828 & 0.123 & 0.049 \\
 & 0 & 0.790 & 0.131 & 0.079 \\
 & 0 & 0.824 & 0.131 & 0.045 \\
\hdashline
 & 0 & 0 & 0.853 & 0.147 \\
 & 0 & 0 & 0.892 & 0.108 \\
C_3 & 0 & 0 & 0.910 & 0.090 \\
 & 0 & 0 & 0.948 & 0.052 \\
 & 0 & 0 & 0.923 & 0.077
\end{array}
$$

$$
P_4 = \begin{array}{c|cccc}
 & C_1 & C_2 & C_3 & C_4 \\
\hline
 & 0.881 & 0.036 & 0.038 & 0.046 \\
C_1 & 0.852 & 0.098 & 0.020 & 0.029 \\
 & 0.864 & 0.062 & 0.037 & 0.036 \\
 & 0.862 & 0.044 & 0.038 & 0.055 \\
\hdashline
 & 0 & 0.826 & 0.127 & 0.047 \\
C_2 & 0 & 0.841 & 0.093 & 0.066 \\
 & 0 & 0.746 & 0.151 & 0.103 \\
 & 0 & 0.790 & 0.131 & 0.079 \\
\hdashline
 & 0 & 0 & 0.913 & 0.087 \\
C_3 & 0 & 0 & 0.948 & 0.052 \\
 & 0 & 0 & 0.853 & 0.147 \\
 & 0 & 0 & 0.892 & 0.108
\end{array}
$$

$$
P_3 = \begin{array}{c|cccc}
 & C_1 & C_2 & C_3 & C_4 \\
\hline
 & 0.854 & 0.090 & 0.023 & 0.032 \\
C_1 & 0.867 & 0.058 & 0.040 & 0.035 \\
 & 0.872 & 0.040 & 0.038 & 0.050 \\
\hdashline
 & 0 & 0.825 & 0.127 & 0.048 \\
C_2 & 0 & 0.746 & 0.151 & 0.103 \\
 & 0 & 0.817 & 0.115 & 0.068 \\
\hdashline
 & 0 & 0 & 0.928 & 0.072 \\
C_3 & 0 & 0 & 0.869 & 0.131 \\
 & 0 & 0 & 0.905 & 0.095
\end{array}
$$

$$
P_2 = \begin{array}{c|cccc}
 & C_1 & C_2 & C_3 & C_4 \\
\hline
C_1 & 0.868 & 0.051 & 0.039 & 0.041 \\
 & 0.854 & 0.090 & 0.023 & 0.032 \\
\hdashline
C_2 & 0 & 0.822 & 0.122 & 0.056 \\
 & 0 & 0.747 & 0.150 & 0.103 \\
\hdashline
C_3 & 0 & 0 & 0.874 & 0.126 \\
 & 0 & 0 & 0.918 & 0.082
\end{array}
$$

$$
P_1 = \begin{array}{c|cccc}
 & C_1 & C_2 & C_3 & C_4 \\
\hline
C_1 & 0.863 & 0.066 & 0.033 & 0.038 \\
C_2 & 0 & 0.801 & 0.130 & 0.069 \\
C_3 & 0 & 0 & 0.904 & 0.096
\end{array}
$$

Table 2: Model comparison test result

Compared models	Log likelihood	L_r	AIC	BIC	Model with better fit on the basis of L_r, AIC and BIC
HMM1 vs HMM2	-207.16303 (-167.61626)	79.0935	438.3261 (383.2325)	438.1133 (382.8070)	HMM2
HMM1 vs HMM3	-207.16303 (-157.02183)	100.2824	438.3261 (386.0437)	438.1133 (385.4054)	HMM3
HMM1 vs HMM4	-207.16303 (-148.68615)	116.9538	438.3261 (393.3723)	438.1133 (392.5213)	HMM4
HMM1 vs HMM5	-207.16303 (-138.29987)	137.7263	438.3261 (396.5997)	438.1133 (395.5360)	HMM5
HMM2 vs HMM3	-167.61626 (-157.02183)	21.1889	383.2325 (386.0437)	382.8070 (385.4054)	Equal
HMM3 vs HMM4	-157.02183 (-148.68615)	16.6714	386.0437 (393.3723)	385.4054 (392.5213)	Equal
HMM4 vs HMM5	-148.68615 (-138.29987)	20.7726	393.3723 (396.5997)	392.5213 (395.5360)	Equal

(Note: The Value in bracket is for the second model in each case)

4.2. RESULTS AND DISCUSSION

It can be observed in the estimated transition probabilities in the transition probability matrices that all the hidden Markov model types are realized from the data. For instance, in P_5 there can be distinguished five distinct probabilities of moving from each class to the other classes, which represent the transition probabilities from the five hidden classes within each of the observable classes. For example, in P_5 (HMM5) moving from C_1 (senior lecturer) to C_2 (reader) has values (arranged in descending order of magnitude) $(P_5)_{42}$ = 0.098, $(P_5)_{22}$ = 0.075, $(P_5)_{12}$ = 0.063, $(P_5)_{32}$ = 0.055 and $(P_5)_{52}$ = 0.040 corresponding respectively to the probabilities of high movers, movers, above mediocres, mediocres and stayers making this transition. In P_4 (HMM4) the same movement from C_1 to C_2has four values (arranged in descending order of magnitude) $(P_4)_{22}$ = 0.098, $(P_4)_{32}$ = 0.062, $(P_4)_{42}$ = 0.044and $(P_4)_{12}$ = 0.036 corresponding respectively to the probabilities of high movers, movers, mediocres and stayers making this transition. Similarly, in P_3 (HMM3) the same movement from C_1 to C_2has three values (arranged in descending order of magnitude) $(P_3)_{12}$ = 0.09, $(P_3)_{22}$ = 0.058 and$(P_3)_{32}$ = 0.040 corresponding respectively to the probabilities of movers, mediocres and stayers making this transition. In P_2 (HMM2) the same movement from C_1 to C_2 is made by movers and stayers with probabilities $(P_2)_{22}$ = 0.090 and $(P_2)_{12}$ = 0.051 respectively. In P_1(HMM1), however, the same movement from C_1 to C_2 is made with a single probability $(P_1)_{12}$ = 0.066.

In the model comparisons the three tests, based on L_r, AIC and BIC statistics, lead to the same conclusion in each case. For example, in comparing the performance of HMM1 versus HMM2 (HMM1 vs HMM2) $L_r = 79.0935 > \chi^2_{0.05}(18) = 28.869$, AIC and BIC for HMM1 have values 438.3261 and 438.1133 respectively which are correspondingly greater than 383.2325 and 382.8070, the respective values of AIC and BIC for HMM2. With this HMM2 is shown to have a significantly better fit to the data than HMM1. In comparing the performance of HMM2 versus HMM3 $L_r = 21.1889 < \chi^2_{0.05}(24) = 36.415$,AIC and BIC for HMM2 have values 383.2325 and 382.8070 respectively which are correspondingly less than 386.0437 and 385.4054, the respective values of AIC and BIC for HMM3. This shows that HMM2 and HMM3 perform equally in their fit to the data. Other comparisons are similarly made. Table 2 shows the results of the model comparisons. It can be seen in Table 2 that all the four models, HMM2, HMM3, HMM4 and HMM5, which are more homogeneous than HMM1, are significantly better than HMM1 which is more parsimonious but less homogeneous. However, after HMM2 all other models of higher homogeneity are only as good as the preceding model in the ladder of homogeneity.

5. CONCLUSION

In this paper, it has been shown how a seemingly simple Markov manpower system, with personnel classes arising from observable data only, can be transformed to a system with both observable and hidden personnel classes through hidden Markov model approach. This produces Markov manpower models that are more homogeneous with respect to personnel inter-class transitions.

It has also been demonstrated that homogeneity can be considered a more important property than parsimony in Markov manpower models. However, there is a point in the level of homogeneity (number of hidden classes allowed) beyond which there is no more gain in adding more hidden classes to achieve homogeneity. It may be possible to define this point in the level of homogeneity beyond which there is no more gain in adding more hidden classes. This may also be data dependent; this is left out for further research.

References

1. Seal HL. The mathematics of a population composed of *k* stationary strata each recruited from the stratum below and supported at the lower level by a uniform number of annual entrants. Biometrika. 1945; 33:226-230.
2. Bartholomew DJ, Forbes A F, McClean S I. Statistical techniques for manpower planning. 2nd ed. Chichester: Wiley; 1991.
3. De Feyter T. Modelling heterogeneity in manpower planning: dividing the personnel system into more homogeneous subgroups. Appl Stoch Models Bus Ind. 2006; 22(4):321–334.
4. Guerry M-A. Hidden heterogeneity in manpower systems: A Markov-switching model approach. European Journal of Operational Research. 2011; 210 (1):106-113. DOI: 10.1016/j.ejor.2010.10.039
5. Rombaut E, Guerry M-A. Decision trees as a classification technique in manpower planning. In: The 16th conference of the applied stochastic models and data analysis international society. 2015; 863–877.
6. Davies GS. A note on a continuous time Markov manpower model. Journal of Applied Probability. 1985;22:932-938. DOI: https://doi.org/10.2307/3213961
7. Ugwuowo FI, McClean S I. Modelling heterogeneity in a manpower system: A review. Applied Stochastic Models in Business and Industry. 2000;16 (2):99–110. DOI:10.1002/1526-4025(200004/06)16:2<99::AID-ASMB385>3.0.CO;2-3.

8. Udom, AU, Ebedoro UG. On multinomial hidden Markov model for hierarchical manpower systems. Communications in Statistics – Theory and Methods. 2019. https://doi.org/10.1080/03610926.2019.1650185.
9. Seasholtz MB, Kowalski B. The parsimony principle applied to multivariate calibration. Analytica Chimica Acta. 1993;277(2):165-177. https://doi.org/10.1016/0003-2670(93)80430-S.
10. Guerry MA, De Feyter T. Markovian approaches in modelling workforce systems. Journal of Current Issues in Finance, Business and Economics. 2009; 2 (4):1–20.
11. Blumen I, Kogan M, McCarthy PJ. The industrial mobility of labour as a probability process. Ithaca, New York: Cornell University Press; 1955.
12. Spilerman S. Extensions of the Mover-Stayer model. American Journal of Sociology. 1972;78 (3):599–626. DOI:10.1086/225366.
13. Visser I, Raijmakers ME, Molenaar PC. Fitting hidden Markov models to psychological data. Scientific Programming. 2002;10(3):185–99. DOI:10.1155/2002/874560.
14. Shirley KE, Small DS, Lynch KG, Maisto SA, Oslin DW. Hidden Markov models for alcoholism treatment trial data. The Annals of Applied Statistics. 2010; 4 (1):366–95. DOI: 10.1214/09-AOAS282.
15. Maruotti A, Punzo A, Bagnato L. Hidden Markov and semi-Markov models with multivariate leptokurtic-normal components for robust modelling of daily returns series. Journal of Financial Econometrics. 2019;17(1):91–117. DOI:10.1093/jjfinec/nby019.
16. Can CE, Ergun G, Soyer R. Bayesian analysis of proportions via a hidden Markov model. Methodology and Computing in Applied Probability. 2022. https://doi.org/10.1007/s11009-022-09971-0.
17. MacDonald IL, Zucchini W. Hidden Markov and other models for discrete-valued time series. London:Chapman & Hall; 1997.
18. Rabiner LR. A tutorial on hidden Markov model and selected applications in speech recognition. Proceedings of the IEEE. 1989;77(2):257–86. Doi:10.1109/5.18626.
19. Bartholomew DJ. Stochastic Models for Social Processes. 3rd ed. Chichester: Wiley; 1982.