Forecasting the Size Distribution of Financial Plants in Swedish Municipalities*

Kurt Brännäs
Department of Economics, Umeå University
S-90187 Umeå, Sweden
email: kurt.brannas@econ.umu.se

Umeå Economic Studies 478, October 1998

Abstract

The paper studies the forecasting of a future size distribution of plants. As a model we use an open Markov chain model for macro data. Estimation is by reparametrization instead of by inequality restrictions using single equation least squares. The estimator is studied in a small Monte Carlo experiment for short time series lengths and macro data. Well-known mobility indices and a new idea of using a truncated transition probability matrix are discussed and also studied in the Monte Carlo experiment. For the financial plants (1984-1993) we find evidence of mobility of a downsizing nature. In a one-step-ahead forecast evaluation we find some overprediction.

Key words: Open Markov chain, Mobility index, Reparametrization, Least squares estimation, Plant size, Municipality.

JEL: C3, C5, G2, L2

*Elisabet Berglund is thanked for her work on the data set used in the empirical illustration, Gunnar Rosenqvist is thanked for a useful discussion, and Xavier de Luna for comments. The financial support from the Swedish Council for Research in the Humanities and Social Sciences is acknowledged.

1. Introduction

The paper focuses on the forecasting of size distribution and on the mobility of plants[1] within the financial sector of Swedish municipalities. Plant size is measured annually along disjoint size intervals, such that for each interval the number of plants is recorded. This type of data is said to be at the macro level, since it is not possible to record individual plants' transitions between states (micro data).

Models of firm size and indices of mobility separately have long histories, for a review see Bartholomew (1982). Zepeda (1995) appears to be the only paper focusing in a related way on dynamic models for firm/plant size. The present framework differs from hers in the way we separately account for entry and exit rather than focusing on 'net-entry'. The model of Berglund and Brännäs (1998) for the stock of plants at a semi-aggregate level is based on assumptions of a common exit probability across plants and a single mean entry stream. The present modelling approach adds more structure and can therefore provide more detailed inferences. In addition, we discuss an idea of relating model based measures of mobility to explanatory variables in a regression like manner (cf. the related area of inefficiency studies, e.g., Battese and Coelli, 1995).

In Section 2, we introduce the open Markov chain model (e.g., Bartholomew, 1982, ch. 3) for the movements between different plant size intervals (states) and account for plants' entry and exit. Based on this model we briefly introduce some of the mobility indices given for closed Markov chain models by Shorrocks (1978), Conlisk (1985), Geweke et al. (1986) and others in Section 3. This section also contains an alternative, ad hoc, index for upward or downward mobility. A new idea of obtaining mobility indices for a subset of all transitions (exclusive of entry and exit states) is also introduced.

Section 4 defines the estimation problem and considers alternative ways of incorporating the natural restrictions that hold for transition probabilities. Both single equation and joint estimation approaches are considered. While there has been substantial research on estimation for closed Markov chains, nothing much is known about estimation in the open Markov chain model.

Section 5 presents the forecast expression for the model. Section 6

[1] A firm may have have one or several plants, possibly located in different municipalities. The majority of firms have one plant. From a municipality point of view, employment within the plants of its territory is therefore of a more direct interest than that of firms.

reports the results of a limited Monte Carlo study on the small sample properties of estimators and indices. Some empirical results for employment mobility and forecasts are reported in Section 7. A few summarizing remarks are made in the final section. In this section we also discuss an extended model specification in which transition probabilities depend directly on explanatory variables.

2. The Basic Model

Following Bartholomew (1982, ch. 3) we define the model as an open Markov chain in discrete time. Let $n_j(t)$ denote the number of plants in an arbitrary municipality in size interval $j = 1, \ldots, k$, with k finite, at time $t = 1, \ldots, T$, so that $N(t) = \sum_{j=1}^{k} n_j(t)$ is the stock of plants at time t. We think of the size intervals as arranged in an increasing order. Further, let $\bar{n}_j(t)$ denote the expectation of $n_j(t)$ and let there be $R(t+1)$ entrants in the period $(t, t+1)$. Denote by $p_{ij}, i, j = 1, 2, \ldots, k$, the time invariant transition probability of moving from a state (size interval) i at time t to a state j at time $t+1$. The p_{ij} are collected into a transition probability matrix $\mathbf{P}$.

Since we wish to allow for exits, i.e. transitions from any state i to a state $k+1$, we have as a consequence the row sum property $\sum_{j=1}^{k} p_{ij} \leq 1$, for every row i. The exit probability from state i in the time interval $(t, t+1)$ is denoted $p_{i,k+1}$ and is then

$$p_{i,k+1} = 1 - \sum_{j=1}^{k} p_{ij} \qquad (i = 1, 2, \ldots, k).$$

The $R(t+1)$ entrants are distributed over the k states by transition probabilities $\mathbf{p_0} = (p_{01}, p_{02}, \ldots, p_{0k})$ such that $\sum_{j=1}^{k} p_{0j} = 1$.

The relationship for the expected frequencies of plants in each state j is for given entry, $R(t+1)$, of the form

$$\bar{n}_j(t+1) = \sum_{i=1}^{k} p_{ij}\bar{n}_i(t) + R(t+1)p_{0j} \quad (j = 1, 2, \ldots, k,\ t = 1, 2, \ldots, T-1)$$

or

$$\bar{\mathbf{n}}(t+1) = \bar{\mathbf{n}}(t)\mathbf{P} + R(t+1)\mathbf{p_0} \qquad (t = 1, 2, \ldots, T),$$

where $\bar{\mathbf{n}}(t) = (\bar{n}_1(t), \ldots, \bar{n}_k(t))$. Note that in this model formulation the exit probabilities are only implicitely present. Therefore, if $R(t+1)$ is known no information about the number of exits would be required for the estimation of the $\mathbf{P}$ matrix.

Unfortunately, the $R(t+1)$ is not observed. With $E[R(t+1)] = \lambda$ we have the unconditional result $\bar{\mathbf{n}}(t+1) = \bar{\mathbf{n}}(t)\mathbf{P} + \lambda\mathbf{p_0}$. Assuming $\{R(t)\}$ to be a sequence of independent and constant mean (λ) Poisson distributed random variables, and $\mathbf{n}(1)$ fixed, Pollard (1967) shows that the elements $n_j(t+1)$ of $\mathbf{n}(t+1)$ are independently Poisson distributed for large t.

With $E[R(t+1)] = \lambda_t$ we indicate that entry in the period $(t, t+1)$ is governed by the parameter value at the initial time point in the interval. In a related way we may specify a time dependent transition probability matrix $\mathbf{P}(t)$. We may write a time varying model as

$$\bar{\mathbf{n}}(t+1) = \bar{\mathbf{n}}(t)\mathbf{P}(t) + \lambda_t\mathbf{p_0}. \tag{1}$$

In Section 8, below, we make some additional remarks on $\mathbf{P}(t)$ and on λ_t specifications. The time dependence would typically be taken to be generated by explanatory variables.

3. Mobility Indices

In studying the dynamic properties of (1) the essential part of the model is obviously the $\mathbf{P}$ matrix. For other dynamic models, e.g., the eigenvalues of corresponding matrices are used to provide summarizing measures of model properties. For a closed system (i.e. when $\lambda\mathbf{p_0} = \mathbf{0}$ and $p_{i,k+1} = 0$, for any i), $\mathbf{P}$ is the transition probability matrix of the Markov chain. In the present case of an open Markov chain model, there is, however, some ambiguity in advancing a transition probability matrix. The ambiguity is due to how and if entry and exit states are to be incorporated. For this reason the 'strong' criteria forming the basis for the mobility indices suggested by Shorrocks (1978), Geweke et al. (1986) and others seem to loose some of their strength.

The following mobility indices, taken from Shorrocks (1978) and Geweke et al. (1986), are based on a transition probability matrix $\mathbf{P}^*$:

$$\begin{array}{ll} M_1(\mathbf{P}^*) = [k - \text{tr}\ \mathbf{P}^*]/(k-1) & M_2(\mathbf{P}^*) = \exp(-h\tau) \\ M_3(\mathbf{P}^*) = (k - \sum_{i=1}^{k} |\rho_i|)/(k-1) & M_4(\mathbf{P}^*) = 1 - |\det \mathbf{P}^*| \\ M_5(\mathbf{P}^*) = 1 - |\rho_2|. & \end{array}$$

Here, $h = -\ln 2/\ln|\rho_2|$, with $\rho_i, i = 1, \ldots, k$, the eigenvalues of $\mathbf{P}^*$ arranged in decreasing order so that ρ_2 is the second largest one. The

τ is the time period length (here $\tau = 1$ year). Note that $M_1, \ldots, M_5$ make no distinction between upward or downward mobility, except very indirectly. The ρ_2 is a measure of half life span (in demography the generational length), so that a large ρ_2 indicates a reduction in the sizes of plants (and then small M_2 and M_5).

The below diagonal elements $p^*_{ij}, i < j$, should for upward mobility be smaller than the elements above the diagonal, when $i > j$, if states are arranged in an increasing order. We introduce an ad hoc index based on this idea as

$$M_6(\mathbf{P}^*) = \frac{1}{2}\{1 + (\sum_{i>j} p^*_{ij} - \sum_{i<j} p^*_{ij})/[k(k-1)]\} \in [0, 1],$$

where the leading term $\frac{1}{2}$ standardizes to the unit interval. A large M_6 ($> \frac{1}{2}$) indicates upward mobility, $M_6 < \frac{1}{2}$ indicates downward mobility, while for $M_6 = \frac{1}{2}$ either no mobility or equal upward and downward mobility are indicated.

Bartholomew (1982, ch. 3) proposes the following transition probability matrix

$$\mathbf{P}^* = \begin{pmatrix} \mathbf{p}_0 & 0 \\ \mathbf{P} & \mathbf{p}'_{k+1} \end{pmatrix},$$

where $\mathbf{p}_{k+1} = (p_{1,k+1}, \ldots, p_{k,k+1})$. Obviously, we may rearrange states without altering the model. Here, with more employees in higher states, it appears natural to let the exit state be placed before the smallest plant size. Hence, we have

$$\mathbf{P}^* = \begin{pmatrix} 0 & \mathbf{p}_0 \\ \mathbf{p}'_{k+1} & \mathbf{P} \end{pmatrix}. \tag{2}$$

Note that only M_4 remains unchanged by this column order change. Since this matrix accounts for the entry but not for the exit mechanisms we could extend $\mathbf{P}^*$ by an exit state and restrict entry to only arise from existing firms or plants:

$$\mathbf{P}^* = \begin{pmatrix} 0 & 0 & \mathbf{p}_0 \\ \mathbf{q}'_0 & \mathbf{p}'_{k+1} & \mathbf{P} \\ 0 & 1 & \mathbf{0} \end{pmatrix}. \tag{3}$$

Admittedly there is arbitrariness in all of these $\mathbf{P}^*$ matrix specifications and for (3) we will not be able to estimate $\mathbf{q}_0$ given the data type at hand.

A simpler solution that avoids the arbitrariness is to restrict focus only on the mobility between the size states $1, \ldots, k$, i.e. to avoid the entry and exit states altogether. We obtain a truncated transition probability by deviding the underlying probability by the probability of being in any of the k states:

$$p^*_{ij} = p_{ij} / \sum_{r=1}^{k} p_{ir} \quad (i, j = 1, \ldots, k). \tag{4}$$

For the resulting truncated transition probability matrix the row sum property is satisfied and inferences about mobility within the restricted class can easily be made. Hence, for this matrix all properties of indices remain true, but then for the restricted state space.

4. Estimation

In this section we focus on the estimation of parameters in a time invariant case. The estimation is performed for a single municipality. Should a model covering all municipalities with common $\mathbf{P}$ and $\mathbf{p_0}$ be the issue of interest, data on municipalities need to be stacked. The time dependent case is briefly discussed in Section 8.

The observations for state j for an arbitrary municipality are generated according to

$$n_j(t+1) = \sum_{i=1}^{k} n_{ij}(t) + \lambda p_{0j} + \epsilon_j(t+1),$$

where $\lambda p_{0j} + \epsilon_j(t+1) = n_{0j}(t+1)$ corresponds to the number of entrants. Replacing in addition $n_{ij}(t)$ with $E[n_{ij}(t)|n_i(t)] + \xi_{ij}(t) = p_{ij} n_i(t) + \xi_{ij}(t)$ we get

$$n_j(t+1) = \sum_{i=1}^{k} p_{ij} n_i(t) + \lambda p_{0j} + [\epsilon_j(t+1) + \sum_{i=1}^{k} \xi_{ij}(t)]. \tag{5}$$

For all observations $t = 1, \ldots, T-1$ we may write

$$\mathbf{n}_j = \bar{\mathbf{n}} \mathbf{P}_j + \mathbf{1}_{T-1} \lambda p_{0j} + \boldsymbol{\epsilon}_j + \boldsymbol{\xi}_j, \tag{6}$$

where $\mathbf{n}_j = (n_j(2), \ldots, n_j(T))'$, $\mathbf{1}_{T-1} = (1, \ldots, 1)'$,

$$\bar{\mathbf{n}} = \begin{pmatrix} n_1(1) & n_2(1) & \cdots & n_k(1) \\ n_1(2) & n_2(2) & \cdots & n_k(2) \\ \vdots & \vdots & & \vdots \\ n_1(T-1) & n_2(T-1) & \cdots & n_k(T-1) \end{pmatrix},$$

and where $\mathbf{P}_j$ is the jth column of $\mathbf{P}$.

By stacking the k equations, the entire system for all $k(T-1)$ observations can for an arbitrary municipality be written on the form

$$\mathbf{n} = (\bar{\mathbf{n}} \otimes \mathbf{I}_{T-1}) vec(\mathbf{P}) + (\mathbf{1}_{T-1} \otimes \mathbf{p}_0)\lambda + \boldsymbol{\epsilon} + \boldsymbol{\xi}, \tag{7}$$

where $\mathbf{n} = (\mathbf{n}_1', \cdots, \mathbf{n}_k')'$, $\mathbf{I}_{T-1}$ is the $T-1$ identity matrix and $\otimes$ is the Kronecker matrix product.

The covariance matrix structure of $\boldsymbol{\epsilon} + \boldsymbol{\xi}$ is quite complicated and we will abstain from using estimation methods building on this structure. Similarly, we will not present any covariance matrices for the parameter estimators at this stage.

4.1 Single Equation Estimation

To estimate the unknown parameters in $\mathbf{P}, \mathbf{p}_0$ as well as λ the simplest estimator is based on the least squares criterion function for each state j separately. The criterion function to be minimized for each municipality and state is then of the form

$$S_j = \sum_{t=1}^{T-1} (n_j(t+1) - \sum_{i=1}^{k} p_{ij} n_i(t) - \lambda p_{0j})^2 \qquad (j = 1, \ldots, k).$$

It is obvious from S_j that λ and p_{0j} cannot be separately estimated in general. Setting $\alpha_j = \lambda p_{0j}$, makes estimation of α_j feasible. Once all equations are estimated we may use the fact that $\lambda = \sum_{i=1}^{k} \alpha_i$ so that using estimates of α_j we can get an estimate of p_{0j} by the expression $p_{0j} = \alpha_j / \sum_{i=1}^{k} \alpha_i$.

The logical restrictions $p_{ij} \in [0,1]$ and $\sum_{j=1}^{k} p_{ij} \leq 1$ are not automatically satisfied by a least squares or for that matter by any other estimator. Estimation under inequality or band restrictions is computationally more demanding than unrestricted estimation (e.g., Fomby et al., 1984, ch. 6). Employing reparametrizations, e.g., of the forms $p_{ij} = 1/[1 + \exp(\theta_{ij})]$ and $\alpha_j = \exp(\gamma_j)$, and minimizing S_j with respect to the unrestricted $\theta_{ij}, \gamma_j \in (-\infty, \infty)$ yields in a very convenient way $\hat{p}_{ij} \in [0,1]$ and $\hat{\alpha}_j \geq 0$. In this way linear estimation subject to inequality restrictions is replaced by unrestricted nonlinear least squares estimation. Note, however, that the row sum property can not be enforced with a single equation estimation technique.

4.2 Joint Estimation

The restriction that row sums of $\mathbf{P}$ should be smaller than or equal to one can only be imposed when all k equations for a municipality are estimated jointly. In the least squares framework we then minimize $S = \sum_{j=1}^{k} S_j$, i.e.

$$S = (\mathbf{n} - \mathbf{W}\boldsymbol{\beta})'(\mathbf{n} - \mathbf{W}\boldsymbol{\beta}),$$

where $\mathbf{W}\boldsymbol{\beta}$ corresponds to the systematic part of (7), i.e. $\mathbf{W} = (\bar{\mathbf{n}} \otimes \mathbf{I}_{T-1}, \mathbf{1}_{T-1} \otimes \mathbf{I}_k)$ and $\boldsymbol{\beta} = (vec(\mathbf{P})', \boldsymbol{\alpha}')'$, with $\boldsymbol{\alpha} = \lambda \mathbf{p_0}$.

For this case a useful reparametrization is, for instance, the multinomial logistic probability, such that for $i, j = 1, \ldots, k$:

$$p_{ij} = \exp(\theta_{ij}) / \sum_{r=1}^{k+1} \exp(\theta_{ir}).$$

For $\boldsymbol{\alpha}$ we suggest using the exponential reparametrization of the previous subsection. The criterion function S is minimized with respect to the unrestricted $k^2 + 2k$ dimensional parameter vector $\boldsymbol{\theta} = (\theta_{11}, \ldots, \theta_{kk+1}, \alpha_1, \ldots, \alpha_k)'$. Using the reparametrization, the linear inequality restrictions are avoided and a nonlinear least squares estimator can be employed.

4.3 Regressing Mobility Indices

Given an estimated $\mathbf{P}^*$ matrix for each municipality m, the mobility indices $M_{im}(\hat{\mathbf{P}}^*)$ can be calculated. Any of these indices can then be used directly or after transformation as dependent variables in formulating, say, linear (ecological) regression models

$$f[M_{im}(\hat{\mathbf{P}}^*)] = \mathbf{z}_m \boldsymbol{\gamma} + v_m \qquad (m = 1, 2, \ldots, M) \tag{8}$$

in which $\boldsymbol{\gamma}$ can be estimated by least squares. The $\mathbf{z}_m$ vector would for the present purposes characterize the municipality and could contain variables that can be controlled by the municipality or some higher level of government. These estimates may help in isolating factors that are important for the mobility of plants within the municipality. The function $f[\cdot]$ could, for instance, be the inverse of the logistic transformation such that $y_{im} = \ln(1/M_{im} - 1) \in (-\infty, \infty)$.

It is important to remember that for the interpretation of the mobility indices the sets of logical constraints on transition probabilities are assumed to be satisfied.

5. Forecasting

The h-steps-ahead forecast for $\mathbf{n}_{T+h}$ based on the time invariant model is of the form

$$\hat{\mathbf{n}}_{T+h|T} = \mathbf{n}_T \mathbf{P}^h + \lambda \mathbf{p}_0 \sum_{i=0}^{h-1} \mathbf{P}^i \quad (h = 1, 2, \ldots) .$$

By substitution of estimated parameters, forecasts to throw interesting light of future employment in the different size classes of plants and for each municipality can be made. To obtain the aggregate number of employees within a municipality, the best forecast is $\hat{\mathbf{n}}_{T+h|T}\mathbf{1}_k$, i.e. by summing over the k states.

Since the covariance matrix of the forecast depends on the covariance structure of the model, we abstain from presenting such results.

6. Monte Carlo Study

In this section we provide some evidence on the small sample properties of single equation least squares estimators for the full transition probability matrix $\mathbf{P}^*$ as given in (2) and of the mobility indices formed for the truncated transition probability matrix as defined in (4). A more complete study that would include, for instance, the estimator of $\boldsymbol{\gamma}$ in (8) is beyond the scope of the present study.

We consider a single $k = 3$ state open Markov chain with

$$\mathbf{P} = \begin{pmatrix} 0.6 & 0.1 & 0.1 \\ 0.1 & 0.6 & 0.0 \\ 0.0 & 0.1 & 0.6 \end{pmatrix} \quad \text{and} \quad \mathbf{p}_0 = (0.6, 0.3, 0.1).$$

For this model the truncated transition probability matrix is

$$\mathbf{P}^* = \begin{pmatrix} 0.75 & 0.125 & 0.125 \\ 0.14 & 0.86 & 0.0 \\ 0.0 & 0.14 & 0.86 \end{pmatrix}$$

and $M_1(\mathbf{P}^*) = 0.268$, $M_2(\mathbf{P}^*) = 0.107$, $M_3(\mathbf{P}^*) = 0.266$, $M_4(\mathbf{P}^*) = 0.462$, $M_5(\mathbf{P}^*) = 0.266$ and $M_6(\mathbf{P}^*) = 0.497$. The indices indicate modest mobility and M_6 suggests that this is marginally downward oriented. The $R(t+1)$ is generated as independently Poisson distributed with $\lambda = 15$. The time series length is set to $T = 10(10)50, 100$. 1000 replications are generated and estimated by least squares (without and with

Table 1: Means of estimated transition probabilities, based on 1000 replications.

Para-	Sample size (T)					
meter	10	20	30	40	50	100
p_{11}	0.36	0.47	0.53	0.56	0.58	0.61
p_{12}	0.24	0.20	0.17	0.16	0.15	0.14
p_{13}	0.18	0.15	0.13	0.12	0.11	0.10
p_{21}	0.14	0.10	0.08	0.08	0.07	0.07
p_{22}	0.40	0.49	0.54	0.56	0.58	0.59
p_{23}	0.06	0.04	0.03	0.02	0.01	0.01
p_{31}	0.06	0.03	0.02	0.02	0.01	0.01
p_{32}	0.22	0.14	0.11	0.10	0.10	0.09
p_{33}	0.33	0.43	0.48	0.50	0.53	0.56
p_{01}	0.43	0.42	0.41	0.40	0.39	0.38
p_{02}	0.17	0.20	0.20	0.21	0.22	0.22
p_{03}	0.40	0.38	0.39	0.39	0.39	0.39
λ	15.3	14.5	13.7	13.2	12.9	12.7

reparametrization) for single equations. For the reparametrized case estimates are obtained by a simplex algorithm (AMOEBA, Press et al., 1987) minimizing S_j directly. We use $\hat{\lambda} = \sum_{i=1}^{k} \hat{\alpha}_i$ and $\hat{p}_{0j} = \hat{\alpha}_j/\hat{\lambda}$ to estimate λ and p_{0j}, $j = 1, 2, 3$. If $\hat{p}_{i,k+1} \notin [0, 1]$, it is set to its closest boundary. Since the row sum condition is not enforced in single equation estimation, the sum may be larger than one. In such a case $|\rho_2|$ may well exceed one, if this happens we set $|\rho_2| = 0.9999$. To study the forecasting performance we employ $T - 1$ observations for estimation and make a one-step-ahead forecast for the Tth observation.

6.1 Results

We find that the least squares estimator without reparametrization is the weaker alternative both in terms of frequent violations of the range restriction and in terms of both bias and mean square error. For this reason no detailed results are presented. For the reparametrized least squares estimator we get the mean over replications results presented in Tables 1 and 2.

A general impression from the means over the replications for the

Table 2: Means of mobility indices based on 1000 replications of the truncated transition probability matrix for the reparameterized case.

Index	True value	Sample size (T) 10	20	30	40	50	100
M_1	0.268	0.67	0.46	0.37	0.33	0.30	0.26
M_2	0.107	0.22	0.14	0.09	0.07	0.06	0.05
M_3	0.266	0.53	0.43	0.36	0.33	0.30	0.26
M_4	0.462	0.79	0.69	0.61	0.56	0.52	0.46
M_5	0.266	0.31	0.23	0.18	0.16	0.15	0.13
M_6	0.497	0.51	0.50	0.50	0.50	0.50	0.50

Table 3: Means of forecast errors and means of absolute precentage errors (MAPE) for one-step-ahead forecasts based on 1000 replications for the reparameterized case.

Sample size	Mean error (state)			MAPE (state)		
T	1	2	3	1	2	3
10	-0.149	-0.054	0.168	21.36	20.79	16.58
20	-0.283	0.004	0.239	21.19	19.71	15.16
30	0.019	-0.108	0.066	20.31	18.61	14.93
40	0.030	-0.070	0.002	19.17	18.52	14.74
50	-0.090	0.234	0.181	18.83	18.69	13.29
100	0.017	-0.205	0.255	19.00	18.88	14.37

reparametrized least squares estimator (cf. Table 1) is that the diagonal elements are underestimated for small T, while off-diagonal elements are overestimated. For the largest sample size of $T = 100$, biases are quite small. With respect to the estimation of p_{0j} and λ the results are not encouraging as the biases remain large for p_{0j} and for λ increasing with the sample size. Note that for these parameters estimates from the three equations are combined.

Table 2 gives a summary of the mean results on the mobility indices. It appears that the indices that are based on ρ_2 only (i.e. M_2 and M_5) do not perform very well. For the other mobility indices there are rather small biases for $T = 100$, while sizeable for $T = 10$. Only M_6 appears to be more or less unbiased for all T.

Table 3 presents the results for the one-step-ahead forecasting exer-

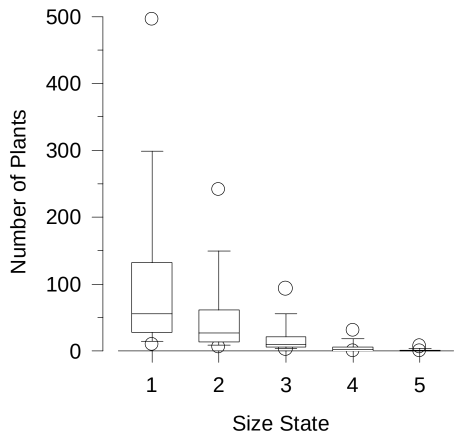

Figure 1: Box plot of the number of plants in municipalities by size state [state 1 (no employees), state 2 (1-4 employees), state 3 (5-19 employees), state 4 (19-100 employees) and state 5 (more than 100 employees)]. Circles indicate the 5 and 95 percent quantiles, respectively.

cise. We find that the mean errors (based on $n_j(T) - \hat{n}_j(T)$) as well as the mean absolute percentage errors (MAPE, based on $100|n_j(T) - \hat{n}_j(T)|/n_j(T)$) gets smaller at a slow rate as T increases. The mean errors are not extremely large as the process have averages of 15, 15 and 23 for the states for time T in the $T = 100$ case.

7. Financial Plants

Data on the number of financial sector (SNI 8) plants in different size intervals is obtained from registers at Statistics Sweden and covers the 10-year period 1984-1993. Figure 1 gives a box plot of the numbers of plants in each size interval. Apparently, most plants have no or only few employees. There is skewness within each of the size classes. Since for some municipalities there is no variation in the number of plants in the highest state, $n_5(t)$, single equation estimation is not feasible. Therefore, states 4 and 5 are added together to form a new state 4.

Due to too small or lacking variation in the data for 22 of the 283 municipalities results are based on the remaining 261 municipalities for which estimation was successful. The single equation estimation results for the **P** matrix are summarized in terms of the means of estimates over municipalities (with variances in parentheses):

$$\mathbf{P} = \begin{pmatrix} \mathbf{0.726} & \mathbf{0.151} & \mathbf{0.017} & \mathbf{0.002} \\ (0.035) & (0.014) & (0.001) & (0.000) \\ \mathbf{0.411} & \mathbf{0.407} & \mathbf{0.068} & \mathbf{0.004} \\ (0.129) & (0.066) & (0.010) & (0.000) \\ \mathbf{0.145} & \mathbf{0.326} & \mathbf{0.584} & \mathbf{0.027} \\ (0.069) & (0.109) & (0.010) & (0.010) \\ \mathbf{0.112} & \mathbf{0.304} & \mathbf{0.547} & \mathbf{0.847} \\ (0.049) & (0.117) & (0.186) & (0.108) \end{pmatrix}.$$

The $\mathbf{p}_0$ vector has mean estimates (with variances in parentheses) (**0.541** (0.139), **0.391** (0.134), **0.067** (0.034), **0.001** (0.000)) and $\hat{\lambda} = 29.6$ (3008).

Looking at the **P** matrix there appears to be a pattern of larger transition probabilities for downsizing than for growing. This corresponds to larger below diagonal than above diagonal elements. For states 2 and 3 the probabilities of remaining in the same state appear surprisingly small. In view of the presented Monte Carlo results for comparable sample sizes, we expect the diagonal elements to be too small on average, while off-diagonal elements are expected to be too large. Assuming constant transition probabilities across municipalities yields the estimated matrix

$$\hat{\mathbf{P}} = \begin{pmatrix} 0.85 & 0.06 & 0.00 & 0.00 \\ 0.28 & 0.75 & 0.00 & 0.00 \\ 0.49 & 0.51 & 1.00 & 0.09 \\ 0.00 & 0.00 & 0.04 & 0.81 \end{pmatrix}$$

with $\hat{p}_{01} = 0.92, \hat{p}_{03} = 0.08$ and $\hat{\lambda} = 4.9$. We note the obvious row sum violations for states 2 and 3. For state 1 the exit probability is estimated to 0.09, while for state 4 it is as large as 0.15.

Mobility indices based on the truncated transition probability matrix for each municipality are graphically displayed in Figure 2 for the 261 municipalities. Indices M_2, M_5 and M_6 all suggest that there is mobility of a downsizing nature. The impression from the index M_4 is that the mobility is quite strong. Note that M_6 does not appear very responsive,

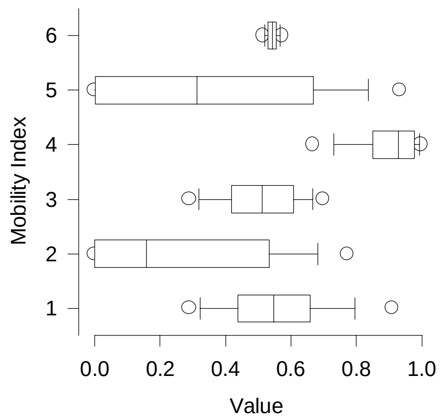

Figure 2: Box plot of mobility indices $M_1, \ldots, M_6$ based on 261 estimated truncated transition probability matrices. Circles indicate the 5th and 95th percent quantiles, respectively.

in the sense of its small variation, to different transition probability matrices. Nothing in the mean $\hat{\mathbf{P}}$ contradicts these conclusions.

Re-estimating the model using only the first nine observations makes it possible to compare the one-step-ahead forecasts with the tenth available and final observations. For the four states we get the mean errors (based on 257 municipalities) -2.7, -4.4 , -3.7 and -1.8, respectively. Hence the result is that we overpredict slightly. In this case the MAPEs are 8.8, 11.6, 18.2 and 29.6, respectively.

8. Discussion

The open Markov chain model has accounted for the distribution in plant size and is therefore richer in detail than a model focusing only on the total number of plants (cf. Berglund and Brännäs, 1998). Note that by summing over the states j we get from (3) the total at time $t+1$:

$$N(t+1) = \mathbf{n}(t+1)\mathbf{1} = \mathbf{n}(t)\boldsymbol{\pi} + \lambda_t + \epsilon(t+1) + \xi(t),$$

where $\mathbf{1} = (1,\ldots,1)'$ and $\boldsymbol{\pi} = (\sum_{j=1}^{k} p_{1j},\ldots,\sum_{j=1}^{k} p_{kj})'$. This expression resembles a multivariate integer-valued AR(1) model that has been subjected to expectation operators. A closed integer-valued AR(1) model coming close to the one considered here has been considered by Brännäs and Brännäs (1998) for the number of fish in connected tanks.

An alternative to the time invariant parameter approach is obvious and closely related to the reparametrization discussed in Sections 4 and 6. Let the transition probability matrix be time dependent such that $\mathbf{P}(t)$ has elements

$$p_{ij}(t) = 1/[1+\exp(\mathbf{z}_t\boldsymbol{\theta})].$$

Note that $\mathbf{z}_t$ and hence $p_{ij}(t)$ vary over municipalities while $\boldsymbol{\theta}$ remains constant, as does $\mathbf{p}_0$. The λ_t could be an exponential function in the same or other explanatory variables. In this case the mobility indices depend in a complicated manner on $\mathbf{z}_t$ and can be calculated for any $\mathbf{z}_t$ value at any time point t. In view of the rather disappointing small sample performance of the studied estimator, accounting for observed inter-municipality variation in transition probabilities appears a promising way to proceed.

Among the mobility indices $M_1,\ldots,M_6$, we observed poor bias-performance for M_2 and M_5, while the ad hoc measure M_6 does not appear to be very sensitive to the values of the transition probability matrices. In the simulation M_1 and M_3 are slightly better than M_4. This experiment is too small to draw far reaching conclusions, however.

Since micro-level data does not need to be available, studies based on highly aggregated data, which then should be more readily available, can be carried out. For instance, we could focus on wage mobility using official wage statistics. If on the other hand, the richer micro-level or combined micro/macro data is available improvements in estimator performance would result (e.g., Rosenqvist, 1985).

References

Bartholomew, D.J. (1982). *Stochastic Models for Social Processes.* 3rd edition, Wiley, New York.

Battese, G.E. and Coelli, T.J. (1995). A Model for Technical Ineffìcency Effects in a Stochastic Frontier Production Function for Panel Data. *Empirical Economics* **20**, 325-332.

Berglund, E. and Brännäs, K. (1998). Plants' Entry and Exit in Swedish Municipalities. Mimeo, Umeå University.

Brännäs, E. and Brännäs, K. (1998). A Model of Patch Visit Behaviour in Fish. To appear in *Biometrical Journal.*

Conlisk, J. (1985). Comparative Statics For Markov Chains. *Journal of Economic Dynamics and Control* **9**, 139-151.

Fomby, T.B., Hill, R.C. and Johnson, S.R. (1984). *Advanced Econometric Methods.* Springer-Verlag, New York.

Geweke, J., Marshall, R.C. and Zarkin, G.A. (1986). Mobility Indices in Continous Time Markov Chains. *Econometrica* **54**, 1407-1423.

Pollard, J.H. (1967). Hierarchical Population Models with Poisson Recruitment. *Journal of Applied Probability* **4**, 209-213.

Press, W.H., Flannery, B.P., Teukolsky, S.A. and Vetterling, W.T. (1987). *Numerical Recipes.* Cambridge University Press, Cambridge.

Rosenqvist, G. (1985). Micro and Macro Data in Statistical Inference on Markov Chains. PhD thesis, Swedish School of Business Administration and Economics, Helsinki.

Shorroks, A.F. (1978). The Measurement of Mobility. *Econometrica* **46**, 1013-1024.

Zepeda, L. (1995). Asymmetry and Nonstationarity in the Farm Size Distribution of Wisconsin Milk Producers: An Aggregate Analysis. *American Journal of Agricultural Economics* **77**, 837-852.