AE88010 October 1988

An Application of Markov Model for Market Share Analysis

by
Won W. Koo

AN APPLICATION OF MARKOV MODEL FOR MARKET SHARE ANALYSIS

Won W. Koo

Consumers' demand for a particular brand of a commodity cannot be measured as a function of the prices of the particular brand and other brands because prices are the same for all brands at the price level determined by the industry's demand and supply schedules. This is particularly true when all brands of a commodity are homogeneous. In this case, demand for a particular brand could be influenced more by consumers' loyalty to the brand rather than by prices of the brand. When products are differentiated from one another, yet are fully substitutable, consumers' loyalty to a brand is also a major factor affecting demand for or market share of the brand. The U.S. soft drink industry is a classical example related to consumers' loyalty. It could be generally assumed that Coca Cola and Pepsi Cola are homogeneous for most consumers. Because these two drinks are competitive, their prices are the same. In this case, consumers' choice between Coca Cola and Pepsi Cola does not depend on prices and tastes of the drinks but depends mainly upon consumers' loyalty to a particular drink.

This argument is not true if there is only one brand in the industry (monopoly market). In this case, the firm producing the brand faces a downward sloping demand curve. Demand for the commodity depends upon the price of the product rather than consumers' loyalty.

Demand for or market share of a particular brand of a commodity in a competitive market could be estimated on the basis of a stochastic process on which consumers' choices are made. Consumers' choice of a brand could be assumed to be generated via simultaneous, dynamic, and stochastic processes and, therefore, is a random variable. Consequently, it is possible to estimate the stochastic process using the Markov Model. The objective of this study is to introduce how to estimate a stochastic process on which demand for or market share of a particular brand is based using the Markov Model. This study includes empirical estimation of the constant transition probabilities for market share of different sizes of farm tractors, and also presents computer programs used to estimate the transition probabilities in the Appendix. The computer program can be used to estimate constant transition probabilities for the number of brands ranging from two to eight for any sample period.

Development of the Markov Probability Model

The Markov Model assumes that the probability distribution of an outcome of a given trial depends only on the outcome of the preceeding trial. This first order dependence is the same for all stages of a stochastic process and is as follows:

$$(1)\ Pr(X_t \mid X_{t-1}, X_{t-2} \ldots) = Pr(X_t \mid X_t{-1})$$

where $Pr(X_t \mid X_{t-1}, X_{t-2} \ldots)$ and $Pr(X_t \mid X_{t-1})$ are the conditional probabilities for an outcome X_t.

The joint probability for $X_0, X_1 \ldots X_T$ can be described on the basis of probability theory as

$$(2) \quad Pr(X_0, X_1, \ldots X_T) = Pr(X_0) \cdot Pr(X_1 \mid X_0) \cdot Pr(X_2 \mid X_0, X_1) \cdot Pr(X_3 \mid X_0, X_1, X_2) \ldots$$

This may be written for the Markov process using Equation 1 as

$$(3) \quad Pr(X_0, X_1, \ldots X_T) = Pr(X_0) \cdot Pr(X_1 \mid X_0) \cdot Pr(X_2 \mid X_1) \ldots$$

$$= Pr(X_0) \prod_{t-1}^{T} Pr(X_t \mid X_{t-1})$$

where $Pr\ (X_1 \mid X_{t-1})$ is the transition probability for X_t for the given X_{t-1}.

The stochastic process could be applied to develop a Markov Model with two states X_t and X_{t-1}. If $X_t = S_j$ and $X_{t-1} = S_i$, then the transition probability for S_j given S_i is as follows:

$$(4) \quad Pr\ (X_t = S_j \mid X_{t-1} = S_i) = P_{ij}$$

where P_{ij} is the constant transition probability associated with a change from state S_i to S_j. For every pair of states, S_i, S_j ($i, j = 1, 2, \ldots r$), P_{ij} must meet the following conditions:

(a) $0 < P_{ij} < 1.0 \qquad i, j = 1, 2, \cdots, r$

(b) $\sum_j P_{ij} = 1.0 \qquad i = 1, 2, \cdots, r$

The joint probability for $X_t = S_j$ and $X_{t-1} = S_i$ is defined as

$$(5) \quad Pr(X_{t-1} = S_i, X_t = S_j) = Pr(X_{t-1} = S_i) \cdot Pr(X_t = S_j \mid X_{t-1} = S_i)$$

Aggregating both sides of Equation 5 over S_i gives

$$(6) \quad \sum_i Pr(X_{t-1} = S_i, X_t = S_j) = \sum_i Pr(X_{t-1} = S_i) \cdot Pr(X_t = S_j \mid S_{t-1} = S_i)$$

Since $\sum_i Pr(X_{t-1} = S_i, X_t = S_j) = Pr(X_t = S_j)$, Equation 6 could be written as

$$(7) \quad Pr(X_t = S_j) = \sum_i Pr(X_{t-1} = S_i)\ Pr(X_t = S_j \mid X_{t-1} = S_i),$$

or

$$(8) \quad q_j\ (t) = \sum_i q_i\ (t-1)\ P_{ij}$$

where $q_j\ (t)$ and $q_i\ (t-1)$ represent the unconditional marginal probabilities $Pr\ (X_t = S_j)$ and $Pr\ (X_{t-1} = S_i)$, respectively. Equation 8 is known as the Markov Model. P_{ij} is the constant transition probability associated with a change from S_i (or q_i) to S_j (or q_j).

Estimation of the Transition Probability

When we estimate P_{ij} from actual observed proportions Y_j (t) and Y_i (t-1), Equation 8 should be rewritten including an error term as follows:

$$(9)\quad Y_j(t) = \sum_i Y_i(t-1)\, P_{ij} + e_j(t)$$

Equation 9 can be written in matrix form as

$$(10)\quad Y_j = X_j P_j + U_j \qquad j = 1, 2, \cdots, r$$

where Y_j = a(T x 1) vector of sample proportions, P_j is a(r x 1) vector of unknown transition parameters to be estimated, X_j is a(T x r) matrix of sample proportions, and U_j is a(T x 1) vector of random disturbances.

The set of equations for j = 1, 2, ⋯, r may be written as

$$(11)\quad \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_r \end{bmatrix} = \begin{bmatrix} X_1 & 0 & \cdots & 0 \\ 0 & X_2 & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & \cdots & X_r \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \\ \vdots \\ P_r \end{bmatrix} + \begin{bmatrix} U_1 \\ U_2 \\ \vdots \\ U_r \end{bmatrix}$$

or $Y = Xp + u$

where Y is (Tr x 1) vector of sample proportion, X is (Tr x r^2) matrix of sample proportions, P is a(r^2 x 1) vector of unknown transition parameters and U is a(Tr x 1) vector of random disturbance term. The stochastic assumption for U is

$$E(U) = 0$$
$$E(U'U) = \Sigma$$

The task is to estimate the transition probability, P_{ij}, from sample proportion data (Y_i (t-1), Y_j (t)). Telser used the conventional least squares estimator to estimate P_{ij} (1962). It, however, does not guarantee that the estimated transition probabilities satisfy the mathematical properties of the probability. Telser suggested a subjective adjustment procedure to correct the transition probability estimates falling outside of the zero-to-one interval. Based on Telser's work, Lee, Judge, and Takayama (1965) and Theil and Rey (1966) used a quadratic programming algorithm to estimate transition probabilities. The quadratic programming algorithm is to minimize the sum of squared errors as

$$(12)\quad e'e = (Y - XP)'\,(Y - XP)$$

subject to

$$(13)\quad GP = N_r$$
$$(14)\quad P > 0$$

where G is a(tr x r) matrix whose elements are 1.0 and N_r is a(Tr x 1) vector whose elements are all 1.0.

Using the conventional summation notation, the inequality restricted estimator based on a quadratic programming algorithm is to minimize the sum of squared errors as

$$(15)\quad SSE = \sum_{j=1}^{r} (Y_{i(t)} - \sum_{i=1}^{r} Y_{i(t-1)}\, P_{ij})^2$$

subject to

$$(16)\quad \sum_{j=1}^{r} P_{ij} = 1.0$$

$$(17)\quad P_{ij} > 0.0$$

Equation 12 (or 15) is in quadratic form in terms of transition probabilities, P_{ij}. The estimated P_{ij} obtained from the quadratic programming model are

$$P_{ij} = \begin{bmatrix} P_{11} & P_{12} & \cdots & P_{1r} \\ P_{21} & P_{22} & \cdots & P_{2r} \\ \vdots & & & \\ P_{r1} & P_{r2} & \cdots & P_{rr} \end{bmatrix}$$

where P_{ij} for j=i represents the transition probability of having the same state in time t as a state given in time t-1. For example, P_{11} represents the transition probability of having state 1 in time t when the same state was given in time t-1. Similar interpretation should be given to all diagonal elements of the matrix. Values of the diagonal elements, therefore, are known as repeat probabilities. P_{ij} for $i \neq j$ represents the transition probability of having state j in time t when state i was given in time t-1. P_{ij} for $i \neq j$, therefore, is the probability of switching from state i to state j and is known as a switching probability. For example, P_{12} represents the transition probability of having state 2 in time t when state 1 was given in time t-1. Similar interpretation should be given to all off-diagonal elements of the transition probability matrix. However, interpretation of P_{21} is different from that of P_{12}. P_{21} represents the transition probabilities of having state 1 in time t when state 2 was given in time t-1. In terms of market share analysis, P_{12} represents the probability for switching from brand 1 to brand 2, while P_{21} represents the probability for switching from brand 2 to brand 1. In general, each row of the transition probability (P_{1j}, P_{2j} ·· or P_{rj} for $j=1,2,\cdots,r$) represent the probabilities for switching to all other brands ($i=1,2,\cdots,r$) from a particular brand ($i=1,2,\cdots$ or r) while each column represents the probabilities for switching to a particular brand ($j=1,2,\cdots$ or r) from all other brands.

Estimation of Steady State Probability

Steady state probabilities are a long-run probability of being a particular state after the process has been operating long enough to wash out the initial condition. In market share analysis, steady state probability represents long-run market share for a particular state. The steady state probabilities can be calculated from the r x r transition probability matrix as follows:

$$[P_1, P_2, \ldots P_r] = [P_{i1}, P_{i2} \ldots P_{ir}] \begin{bmatrix} P_{11} & P_{12} \cdots & P_{1r} \\ P_{21} & P_{22} \cdots & P_{2r} \\ \cdot & & \\ \cdot & & \\ \cdot & & \\ P_{r1} & P_{r2} \cdots & P_{rr} \end{bmatrix}$$

where $[P_1, P_2 \ldots P_r]$ is steady state probabilities and $[P_{i1}, P_{i2} \ldots P_{ir}]$ is a low of the estimated transition probabilities. The steady state probabilities associated with a particular low are the same as those associated with other lows.

The steady state probabilities can be calculated more conveniently by solving a system of r-1 equations generated from the r x r transition probability matrix (Bierman, Bonini, and Hausman) as follows:

$$\begin{aligned} P_1 &= P_{11}P_1 + P_{21}P_2 + \ldots + P_{r1}\,(1-P_1 - P_2 - \ldots P_{r-1}) \\ P_2 &= P_{12}P_1 + P_{22}P_2 + \ldots + P_{r2}\,(1-P_1 - P_2 - \ldots P_{r-1}) \\ &\cdot \\ &\cdot \\ &\cdot \\ P_{r-1} &= P_{1r-1}P_1 + P_{2r-1}P_2 + \ldots + P_{rr-1}\,(1-P_1 - P_2 \ldots P_{r-1}) \end{aligned}$$

This is a system of r-1 equations. Since this system has r-1 unknown variables (P_1, P_2, ... P_{r-1}) and r-1 equations, we can solve this system for P_1, P_2 ... P_{r-1}. P_r is calculated as follows:

$$P_r = 1 - P_1 - P_2 - P_3 - \ldots - P_{r-1}$$

Quadratic Programming Algorithm

The constant transition probabilities can be estimated from Equations 12, 13, and 14 using a quadratic programming algorithm. The software package used for this study is MINOS developed by Murtagh and Saunders. Operation of MINOS is divided into two parts: one is related to the linear portion of a system of equations and the other is related to the nonlinear portion of the system of equations. The linear portion of the system requires the same input format as MPS/360 or MPS/X. The nonlinear portion of the system must be specified in a subroutine in the MINOS program.

The step-by-step procedures in applying the MINOS package to estimate the constant transition probabilities in the Markov Model are as follows:

1. Separate linear and nonlinear terms in P_{ij} in the objective function.

2. Develop a file with linear terms in the objective function and constraints based on the input format for MPS/X or MPS/360.

3. Develop a file with nonlinear terms in the objective function based on the input format specified in MINOS.

4. Specify the model based on the format specified in MINOS.

For example, the numerical illustration of the procedure in formulating the Markov Model with a simple case where a commodity has three different brands (r = 3) can be stated as follows:

$$(18)\quad q_j(t) = \sum_{i=1}^{3} q_i(t-1)\,P_{ij} + e_j(t) \qquad \text{for } j = 1, 2, 3$$

To estimate P_{ij}, Equation 18 can be written in quadratic form as follows. The objective function is to minimize the sum of squared errors as

$$(19)\quad SSE = \sum_{j=1}^{3} \left(q_j(t) - \sum_{i=1}^{3} q_i(t-1)\,P_{ij}\right)^2$$

This objective function is subject to the following constraints

$$(20)\quad \sum_{j=1}^{3} P_{ij} = 1.0$$

$$(21)\quad P_{ij} > 0.0$$

The solution procedures of the quadratic programming model are as follows:

1. Separate linear and nonlinear terms in the objective function by solving Equation 19 as follows:

$$(22)\ SSE = \sum_{j=1}^{3} \Big(\sum_{t=1}^{T} q_j^2 (t) - 2 \sum_{t=1}^{T} q_j (t)\, q_1 (t-1)\, P_{1j}$$

$$- 2 \sum_{t=1}^{T} q_j (t)\, q_2 (t-1)\, P_{2j} - 2 \sum_{t=1}^{T} q_j (t)\, q_3 (t-1)\, P_{3j}$$

$$+ \sum_{t=1}^{T} q_1^2 (t-1)\, P_{1j}^2 + 2 \sum_{t=1}^{T} q_1 (t-1)\, q_2 (t-1)\, P_{1j}\, P_{2j}$$

$$+ 2 \sum_{t=1}^{T} q_1 (t-1)\, q_3 (t-1)\, P_{1j}\, P_{3j} + \sum_{t=1}^{T} q_2^2 (t-1)\, P_{2j}^2$$

$$+ 2 \sum_{t=1}^{T} q_2 (t-1)\, q_3 (t-1)\, P_{2j}\, P_{3j} + \sum_{t=1}^{T} q_3^2 (t-1)\, P_{3j}^2 \Big)$$

We could rewrite the objective function as

$$(23)\ SSE = \sum_{j=1}^{3} (C (j) - B (j, 1)\, P_{1j} - B (j, 2)\, P_{2j}$$

$$- B (j, 3)\, P_{3j} + A (1, 1)\, P_{1j}^2 + 2A (1, 2)\, P_{1j}\, P_{2j}$$

$$+ 2A (1, 3)\, P_{1j}\, P_{3j} + 2A (2, 2,)\, P_{2j}^2 + 2A (2, 3)\, P_{2j}\, P_{3j},$$

$$+ 2A (3,3)\, P_{3j}^2)$$

where C is a constant term, Bs are coefficients for linear terms, and As are coefficients for nonlinear terms. There are 9 linear terms (P_{1j}, P_{2j}, P_{3j}, for $j = 1, 2, 3$) and 18 nonlinear terms (P_{1j}^2, P_{2j}^2, P_{3j}^2, $P_{1j}\ P_{2j}$, $P_{1j}\ P_{3j}$, $P_{2j}\ P_{3j}$ for $j = 1, 2, 3$). Equation (23) is in quadratic form in term of P_{ij}.

2. Develop a file with linear terms in the objective function [Equation (23)] and constraints specified in Equations 20 and 21 as follows:

```
NAME                       SAMPLE
ROWS
 N  OBJ.ROW
 E  P1
 E  P2
 E  P3
COLUMNS
    X1                     OBJ.ROW              B (1, 1)
    X1                     P1                   1.000000
    X2                     OBJ.ROW              B (1, 2)
    X2                     P2                   1.000000
    X3                     OBJ.ROW              B (1, 3)
    X3                     P3                   1.000000
    X4                     OBJ.ROW              B (2, 1)
    X4                     P1                   1.000000
    X5                     OBJ.ROW              B (2, 2)
    X5                     P2                   1.000000
    X6                     OBJ.ROW              B (2, 3)
    X6                     P3                   1.000000
    X7                     OBJ.ROW              B (3, 1)
    X7                     P1                   1.000000
    X8                     OBJ.ROW              B (3, 2)
    X8                     P2                   1.000000
    X9                     OBJ.ROW              B (3, 3)
    X9                     P3                   1.000000
RHS
    RHS                    P1                     1.0
    RHS                    P2                     1.0
    RHS                    P3                     1.0
ENDATA
```

3. Develop a subroutine with nonlinear terms in the objective function as follows:

```
      SUBROUTINE CALCFG (MODE, N, X, F, G, NSTATE, NPROB)
      IMPLICIT REAL*8 (A-H, O-Z)
      DIMENSION X (9), G (9)
      T1   = A(1,1)*X(1)**2+2*A(1,2)*X(1)*X(2)+2*A(1,3)*X(1)*X(3)
     1       +A(2,2)*X(2)**2+2*A(2,3)*X(2)*X(3)+A(3,3)*X(3)**2
      T2   = A(1,1)*X(4)**2+2*A(1,2)*X(4)*X(5)+2*A(1,3)*X(4)*X(6)
     1       +A(2,2)*X(5)**2+2*A(2,3)*X(5)*X(6)+A(3,3)*X(6)**2
      T3   = A(1,1)*X(7)**2+2*A(1,2)*X(7)*X(8)+2*A(1,3)*X(7)*X(9)
     1       +A(2,2)*X(8)**2+2*A(2,3)*X(8)*X(9)+A(3,3)*X(9)**2
      F    = T1 + T2 + T3
      G(1) = 2*A(1,1)*X(1)+2*A(1,2)*X(2)+2*A(1,3)*X(3)
      G(2) = 2*A(1,2)*X(1)+2*A(2,2)*X(2)+2*A(2,3)*X(3)
      G(3) = 2*A(1,3)*X(1)+2*A(2,3)*X(2)+2*A(3,3)*X(3)
      G(4) = 2*A(1,1)*X(4)+2*A(1,2)*X(5)+2*A(1,3)*X(6)
      G(5) = 2*A(1,2)*X(4)+2*A(2,2)*X(5)+2*A(2,3)*X(6)
      G(6) = 2*A(1,3)*X(4)+2*A(2,3)*X(5)+2*A(3,3)*X(6)
```

```
      G(7) = 2*A(1,1)*X(7)+2*A(1,2)*X(8)+2*A(1,3)*X(9)
      G(8) = 2*A(1,2)*X(7)+2*A(2,2)*X(8)+2*A(2,3)*X(9)
      G(9) = 2*A(1,3)*X(7)+2*A(2,3)*X(8)+2*A(3,3)*X(9)
      RETURN
      END
```

where T1, T2, and T3 are the nonlinear portions of the objective function for j=1, 2, and 3, respectively; G(1), G(2), . . ., G(9) are partial derivatives of the nonlinear portion of the objective function with respect to the transition probability associated with each brand and X(1), X(2), . . ., X(9) are the same as P_{11}, P_{12}, . . ., P_{33} , respectively. G(1), G(2), . . . , g(9) are specified as

$$G(1) = \frac{\partial T1}{\partial (X1)}$$

$$G(2) = \frac{\partial T1}{\partial (X2)}$$

$$G(3) = \frac{\partial T1}{\partial (X3)}$$

$$G(4) = \frac{\partial T2}{\partial (X4)}$$

$$G(5) = \frac{\partial T2}{\partial (X5)}$$

$$G(6) = \frac{\partial T2}{\partial (X6)}$$

$$G(7) = \frac{\partial T3}{\partial (X7)}$$

$$G(8) = \frac{\partial T3}{\partial (X8)}$$

$$G(9) = \frac{\partial T3}{\partial (X9)}$$

4. Specify the model based on the MINOS. This file for the example is as follows:

```
BEGIN SPECS
  MINIMIZE
  OBJECTIVE = OBJ.ROW
  RHS = RHS
  ROWS        50
  COLUMNS     50
  ELEMENTS   200
  OLD BASIS FILE    0
  NEW BASIS FILE   11
  CRASH OPTION   1
  PARTIAL PRICE  2
  ITERATIONS   100
  LOG FREQUENCY  1
```

```
    SAVE FREQUENCY  200
    SOLUTION    YES
    PROBLEM NUMBER 1
    NONLINEAR VARIABLES  75
    LOWER BOUND = NONE
    SUPERBASIC LIMIT 40
    HESSIAN DIMENSION  40
    LINSEARCH TOL   0.01
    REDUCED-GRADIENT TOL  0.9
    VERIFY GRADIENTS
    CALL CALCFG IF OPTIMAL
END
```

An Example of Markov Model for Market Share for Different Sizes of Farm Tractors

The productivity of labor and timeliness of field preparation have been enhanced by the use of larger, more sophisticated farm machines such as large horsepower farm tractors. Between 1965 to 1980, the trend was toward larger, more labor-efficient machines while small tractors declined in importance. These trends have moderated since 1980.

Farm tractors, in general, can be categorized as two-wheel drive and four-wheel drive. Two-wheel drive is subdivided into the following horsepower categories: 100-140, 140-170. The number of each size category sold are shown in Tables 1 and 2. It is generally assumed that these farm tractors are substitutable in operating farms. Since farm tractors are product inputs, the theoretical demand function for different sizes of farm tractors can be derived based on a criterion of farmers' profit maximization (Chow 1960, Fox 1966, Griliches 1960). Demand for farm tractors is generally specified as

$$(24)\quad Q_{jt} = f_j\,(P_t,\ r_{jmt},\ r_{imt},\ r_{ht})$$

where Q_t is aggregate demand for j^{th} size of farm tractors in time t, P_t is average price of all crops, r_{jmt} is an average price of the j^{th} size of farm tractor, r_{imt} is price of the i^{th} size of farm tractor and r_{ht} is price of other inputs used in farm operation.

There are two major problems in estimating the demand functions or market shares of these tractors by using an econometric technique. First, the prices of each size of farm tractor (r_{jmt}) are not available. Even if the price data are available for a particular tractor size, the data are not consistent over time because of continuous changes in the quality of tractors. Second, the prices of these tractors move in the same direction, leading to a high degree of correlation among these price variables. As an alternative, the Markov Model could be used to estimate the constant transition probability associated with different tractor sizes.

<u>Markov Model for Farm Tractors</u>

The Markov Model with four different brands of farm tractors is as follows:

$$(25)\quad Q_j(t) = \sum_{i=1}^{4} Q_i(t-1)\,P_{ij} + e_j(t) \qquad \text{for } j = 1, 2, 3, 4$$

where $Q_j(t)$ is market share of the j^{th} brand (tractor size) in time t, $Q_i(t-1)$ is market share of the i^{th} brand in time t-1, P_{ij} is the constant transition probability of having brand j in time t when brand i was given in time t-1, and $e_j(t)$ is the disturbance term.

P_{ij} could be estimated using quadratic programming algorithm. The objective of the model is to minimize the sum of squared errors as follows:

$$(26)\quad SSE = \sum_{j=1}^{4} (Y_j(t) - \sum_{j=1}^{4} Y_i(t-1)\, P_{ij})^2$$

The objective function is subject to the following constraints.

$$(27)\quad \sum_{j=1}^{4} P_{ij} = 1.0$$

$$(28)\quad P_{ij} > 0.0$$

Data used are new farm tractors sold by the four-horsepower classes from 1971 to 1984. Market shares of each class of farm tractors are calculated from the data and are used in the quadratic programming model. New farm tractors sold and estimated market shares are shown in Tables 1 and 2, respectively.

To execute the quadratic programming model (Equations 26, 27, and 28) two files are generated; one contains linear portions of the objective functions and constraints (linear file) which is formulated for the IBM MPS/360 or MPS/X, and the other is a subroutine which contains nonlinear portions of objective functions (nonlinear file) based on the format specified by MINOS. The computer programs used to generate objective values in the linear file and coefficients in the nonlinear file are shown in the Appendix.

Interpretation of the Estimated Transition Probabilities

Table 3 presents transition probabilities for tractor size categories based on time series data from 1972 to 1984. The repeat transition probabilities are the highest (0.98) for 140-170 horsepower tractors indicating that most farmers who own 140-170 horsepower tractors are happy with their tractor and tend to purchase the same tractor in the future. The repeat transition probabilities are 0.83 for four-wheel drive tractors and 0.92 for 100-140 horsepower tractors, which are considered to be high compared to the repeat probabilities for other durable commodities such as passenger cars. This implies that each size category has its own unique characteristics based on farm sizes and types.

Switching probabilities are small compared to repeat probabilities. Probability switching from four-wheel drive to 140-170 horsepower tractors is 0.11 while the reverse order is only 0.02 indicating that a 140-170 horsepower tractor is more commonly used than four-wheel drive.

TABLE 1. UNITS OF NEW TRACTORS SOLD, 1971-1984

Year	Four-Wheel Drive	140HP-170HP	100HP-140HP
1971	2,547	2,549	30,244
1972	3,856	6,191	39,765
1973	6,460	13,384	58,010
1974	8,287	16,951	52,816
1975	10,650	21,256	43,475
1976	10,511	18,221	43,082
1977	7,687	18,523	42,177
1978	8,744	22,178	43,349
1979	11,455	21,603	40,932
1980	10,887	18,718	31,610
1981	9,683	15,657	27,522
1982	6,763	10,536	18,711
1983	6,101	13,651	14,503
1984	3,975	14,653	9,843

SOURCE: Farm and Industrial Equipment Institute.

TABLE 2. CALCULATED MARKET SHARE OF EACH SIZE OF FARM TRACTORS, 1972-1984

Year	Four-Wheel Drive	140HP-170HP	100HP-140HP
1972	0.07207	0.07213	0.85580
1973	0.07741	0.12429	0.79830
1974	0.08298	0.17191	0.74511
1975	0.10617	0.21717	0.67666
1976	0.14096	0.28120	0.57785
1977	0.14636	0.25372	0.59991
1978	0.11240	0.27086	0.61674
1979	0.11773	0.29861	0.58366
1980	0.15482	0.29197	0.55321
1981	0.17785	0.30577	0.51638
1982	0.18318	0.29619	0.52064
1983	0.18781	0.29259	0.51961
1984	0.15339	0.41049	0.43611

TABLE 3. ESTIMATED TRANSITION PROBABILITIES

	4-WD	140-170 HP	100-140 HP
4-WD	0.83	0.10	0.07
140-170 HP	0.02	0.98	0.00
100-140 HP	0.04	0.04	0.92
Steady State	0.12	0.60	0.28

Steady state probabilities are also shown in Table 3. The probabilities 0.12 for four-wheel drive, 0.60 for 140-170 horsepower tractors, and 0.33 for are 100-140 horsepower tractors. These steady state probabilities which are interpreted as the long-run market shares indicate that market share for 140-170 horsepower tractors will be much larger than those in the last 15 years. This is mainly because 140-170 horsepower tractors are more affordable in terms of the prices of the tractor than four-wheel drive tractors and yet can do most work needed in operating farms.

Concluding Remarks

Transition probabilities associated with market share of a brand of a commodity could be estimated using the Markov Model. The estimation technique used is the inequality-restricted estimator based on a quadratic programming algorithm. The estimator was executed using a nonlinear software package developed by Murtagh and Sanders (MINOS). The solution procedure using MINOS was detailed in the text and was exemplified with the U.S. farm tractor industry. Computer programs used to estimate transition probabilities for different sizes of farm tractors on the basis of MINOS are presented in the Appendix.

The transition probabilities estimated in this paper are constant over the entire example period. There are many cases, however, for which the transition probabilities are not constant over the sample period. Further study should be focused on estimation of variable transition probabilities which could be functions of time or certain explanatory variables.

Appendix

This program, developed by Koo and Vreugdenhil, is designed to interact with a quadratic programming software developed by Murtagh and Sanders (MINOS). This program can estimate the constant transition probabilities for brands ranging from one to eight with unlimited number of cross-section and time-series data.

The program is divided into three parts: PROLP (program for the linear portion of the model), PROQP (program for the quadratic portion of the model) and PROSPEC (program for model specification). These three programs are executed by an interacting program called "CONTROL."

The data file can be generated as follows:

1. The first row has four elements.

   Columns 1-3: beginning point of the data
   Columns 4-6: ending point of the data
   Columns 7-9: number of brands
   Columns 10-12: number of observations

These elements are formatted as integer (4I3).

2. The rest of the rows is the quantities of each brand sold. The data are formatted as real (8F10.0). The data file can be constructed as:

Name: TDATA

```
  2  10   4   9
  2547.        2549.       30244.
  3856.        6191.       39765.
  6460.       13384.       58010.
  8287.       16951.       52816.
 10605.       21156.       43475.
 10511.       18221.       43082.
  7687.       18523.       42177.
  8744.       22178.       43349.
 11455.       21603.       40932.
 10887.       18718.       31610.
  9683.       15657.       27522.
  6763.       10536.       18711.
  5101.       13651.       14503.
  3975.       14653.        9843.
```

After completion of constructing the data file, execute the interacting program, CONTROL (this program could be rewritten depending upon the mainframe computer available). Execution of this program will be completed by typing the data file name (TDATA) when the name of the file is requested. The general procedure to use the quadratic programming on CMS is as follows:

1. Log on to CMS (IBM Operating System) and get into file listing by typing FL
2. Place the cursor next to the file called "control exec"
3. Type run and enter
4. Answer the query "enter name of data to be used"
5. The following will appear:

   DMSXGT564 EOF reached
   DMSXGT564 EOF reached
   DMSXGT565 EOF reached
   Submitting job NU038994G to NDSUJES2
   ----------Job Submitted----------
   Wait for output in RDRLIST

6. After this you must press the clear key

   On the PC, this is the plus key on the right side of the keyboard

7. When the program is finished the following message will appear:

   Prt file 8584 from RSCSSNA copy 001 NOHOLD
   File (8575) spooled to NU038994-ORG NDSUJES2 (NU038994) 8/07/87
   15:57:17 O.M.T.

8. This means that your output has been completed. Now we must get that output into your file list by typing RDRLIST. (You may have to press clear again now.)

The control and other programs are as follows:

Program Name: CONTROL

```
&Trace off
*&Print off
*&Type enter name of data to be used
*&Read 10
*File @#TEMP5
*File @#TEMP9
*Load &10
*Save @#TEMP5
*Run prolp
*Sub proqp specs @#TEMP9 @#TEMP5
*Pur @#TEMP5
*Pur @#TEMP9
*Clear
*STA
&Type enter name of data to be used
&Read string &F
Erase @#TEMP5 data A
Copyfile &F data A TEMP5 = A
Filedef FT05F001 disk TEMP5 data A
Filedef FT09F001 disk TEMP9 data A
S-WF
WATFOR77 PROLP
&IF &RC > 0 &GOTO -ERRO
*Submit PROQP job a SPECS1 minos a @#TEMP9 data A @#TEMP5 data A
XEDIT PROQP job a (PROFILE SUBJOB)
Erase TEMP5 data A
Erase TEMP9 data A
&Type----------Job Submitted----------
&Type wait for output in RDRLIST
&EXIT
-ERRO &type an error occurred in the fortran program
&EXIT
```

Program Name: PROLP

```
      REAL*8 M(22,5),TX(22),X(22,5),A(22,5),C(50),OB(50)
      DIMENSION A1(50),A2(50),A3(50),A4(50),A5(50)
      CALL OPSYS('ALLOC','@#TEMP5',5)
      CALL OPSYS('ALLOC','@#TEMP9',9)
      READ(5,76) IBEGIN,IEND,IBRAND,IOBS
   76 FORMAT(4I3)
      IBRD2=IBRAND*IBRAND
      WRITE(9,200)
  200 FORMAT('NAME',10X,'SAMPLE')
      WRITE(9,101)
  101 FORMAT ('ROWS')
      WRITE(9,102)
  102 FORMAT (1X,'N',2X,'OBJ.ROW')
C     SUPPLY AT SUPPLY ORIGINS
      DO 10 I1=1,IBRAND
      IF(I1.LT.10) WRITE(9,103)I1
      IF(I1.GE.10) WRITE(9,104)I1
  103 FORMAT(1X,'E',2X,'P',I1)
  104 FORMAT(1X,'E',2X,'P',I2)
   10 CONTINUE
      DO  50 I=1,IEND
   50 READ(5,77)(X(I,J),J=1,IBRAND)
   77 FORMAT (9F10.0)
      DO 55 I=1,IEND
   55 TX(I)=0.0
      DO 60 I=IBEGIN,IEND
      DO 60 J=1,IBRAND
   60 TX(I)=TX(I)+X(I,J)
      DO 70 I=IBEGIN,IEND
      DO 70 J=1,IBRAND
   70 M(I,J)=X(I,J)/TX(I)
      DO 80 I=1,IBRAND
      DO 80 J=1,IBRAND
   80 A(I,J)=0.0
      L=0
  110 L=L+1
      DO 100 J=1,IBRAND
      II=0
      L2=IBEGIN+1
      DO 100 I=L2,IEND
      II=I-1
  100 A(L,J)=A(L,J)+M(I,L)*M(II,J)
      IF (L.LE.5) GO TO 110
      DO 600 I=1,IBRAND
      DO 600 J=1,IBRAND
      IF (I.EQ.1) IM=J
      IF (I.EQ.2) IM=IBRAND+J
      IF (I.EQ.3) IM=(IBRAND*2)+J
      IF (I.EQ.4) IM=(IBRAND*3)+J
      IF (I.EQ.5) IM=(IBRAND*4)+J
      IF (I.EQ.6) IM=(IBRAND*5)+J
      IF (I.EQ.7) IM=(IBRAND*6)+J
      IF (I.EQ.8) IM=(IBRAND*7)+J
      IF (I.EQ.8) IM=(IBRAND*8)+J
      OB(IM)=A(I,J)*(-2.0)
```

```
 600 CONTINUE
     WRITE(9,121)
 121 FORMAT ('COLUMNS')
     NBRAND=0
     DO 22 I1=1,IBRAND
     DO 22 J1=1,IBRAND
     NBRAND=NBRAND+1
     IF (NBRAND.LT.10) GO TO 33
     IF (NBRAND.GE.10) GO TO 34
  33 WRITE(9,122)NBRAND,OB(NBRAND)
 122 FORMAT(4X,'X',I1,8X,'OBJ.ROW',6X,F9.6)
     WRITE(9,123) NBRAND,J1
 123 FORMAT(4X,'X',I1,8X,'P',I1,12X,'1.0')
     GO TO 22
  34 WRITE(9,124)NBRAND,OB(NBRAND)
 124 FORMAT(4X,'X',I2,7X,'OBJ.ROW',5X,F9.6)
     WRITE(9,125) NBRAND,J1
 125 FORMAT(4X,'X',I2,7X,'P',I1,12X,'1.0')
  22 CONTINUE
     WRITE(9,111)
 111 FORMAT ('RHS')
     DO 16 I1=1,IBRAND
     WRITE(9,113)I1
 113 FORMAT(4X,'RHS',7X,'P',I1,12X,'1.0')
  16 CONTINUE
     WRITE (9,184)
 184 FORMAT('ENDATA')
     END
```

Program Name: PROQP

```
//MINOS JOB 5031630,'K002',CLASS=A,TIME=(,10)
//A EXEC FORTHCLG,REGION=1024K,SYSO=Q,SYSA=Q
//FORT.SYSIN DD *
        SUBROUTINE CALCFG (MODE,N,X,F,G,NSTATE,NPROB)
         IMPLICIT REAL*8 (A-H,O-Z)
         DIMENSION X(81), G(81)
      REAL*8 TX(81),A(81,9),M(81,9),B(81,9),MS(81,9),T(9)
      IF (NSTATE.NE.1) GO TO 999
      READ(5,668) L1,N1,K,NO
C     L1=BEGING YEAR OF YOUR DATA
C     N1=ENDING YEAR OF YOUR DATA
C     K=NUMBER OF BRANDS
C     NO=NUMBER OF OBSERVATIONS
C     N=NUMBER OF GRADIANTS
      N=K*K
      N1=N1-1
      WRITE(6,668) L1,N1,K,NO
 668  FORMAT (4I3)
      DO  50 I=1,N1
      READ(5,76)(B(I,J),J=1,K)
C     B(I,J)=QUANTITIES OF EACH BRAND SOLD
  50  WRITE(6,76)(B(I,J),J=1,K)
  76  FORMAT(9F10.0)
      WRITE(6,123)
  123 FORMAT (' ','END OF INPUT DATA')
      DO 55 I=L1,N1
  55  TX(I)=0.0
      DO 60 I=L1,N1
      DO 60 J=1,K
  60  TX(I)=TX(I)+B(I,J)
      DO 70 I=L1,N1
      DO 70 J=1,K
C     CALCULATE MARKET SHARE OF EACH BRAND,B(I,J)
  70  M(I,J)=B(I,J)/TX(I)
      DO 99 I=1,N1
      WRITE(6,300) (M(I,J),J=1,K)
  99  CONTINUE
      WRITE (6,234)
 234  FORMAT(' ','END OF MARKET SHARES')
      DO 80 I=1,K
      DO 80 J=1,K
  80  A(I,J)=0.0
      L=0
 110  L=L+1
      K1=K-1
      DO 100 J=1,K
      DO 100 I=L1,N1
C     CALCULATE THE SUM OF CROSS MULT. OF MARKET SHARES OF THE BRANDS
C     OVER PERIOD WHICH IS EQUIVALENT TO X'X MATRIX IN LEAST SQUARES
C     ESTIMATOR
```

```
  100  A(L,J)=A(L,J)+M(I,L)*M(I,J)
       IF (L.LE.K1) GO TO 110
       DO 130 I=1,K
  130  WRITE(6,300) (A(I,J),J=1,4)
  300  FORMAT (9F20.5)
       WRITE(6,456)
  456  FORMAT (' ','END OF CROSS-MULTIPLICATION')
 C     SPECIFICATION OF THE NONLINEAR TERM IN THE OBJECTIVE FUNCTION;
 C     T(I) REPRESENTS  NONLINEAR TERMS ASSOCIATED WITH BRAND I
 C     AND G(1),G(2),,,G(KK) ARE PARTIAL DERIVATIVES OF T(I)
 C     WITH RESPECT TO X(J)(BRAND J).
  999  DO 104 I=1,K
  104  T(I)=0.0
       DO 105 I1=1,K
       DO 111 I=1,K
       DO 112 J=1,K
       IM=I+(K*(I1-1))
       JM=J+(K*(I1-1))
       T(I1)=T(I1)+A(I,J)*X(IM)*X(JM)
  112 CONTINUE
  111 CONTINUE
  105 CONTINUE
      F=0.0
      DO 115 I=1,K
  115 F=F+T(I)
      KK=K*K
      DO 116 I=1,KK
  116 G(I)=0.0
      II=0
      DO 120 I1=1,K
      DO 126 I=1,K
      II=II+1
      DO 125 J=1,K
      JM=J+(K*(I1-1))
      G(II)=G(II)+2*A(I,J)*X(JM)
  125 CONTINUE
  126 CONTINUE
  120 CONTINUE
      RETURN
      END
//LKED.XLIB DD DSN=ACAD.N5031508.MINOS.LOAD,DISP=SHR
//LKED.SYSIN DD *
 INCLUDE XLIB(MINOS)
 ENTRY MAIN
//GO.FT08F001 DD UNIT=SYSDA,SPACE=(TRK,(4,4)),DISP=(,PASS),
//  DCB=(RECFM=FB,LRECL=8,BLKSIZE=1600)
//GO.FT09F001 DD DUMMY
//GO.FT10F001 DD DUMMY
//GO.FT11F001 DD DUMMY
//GO.SYSIN DD *
```

Program Name: PROSPEC

```
BEGIN SPECS
   MINIMIZE
   OBJECTIVE = OBJ.ROW
   RHS = RHS
   ROWS         50
   COLUMNS      50
   ELEMENTS    200
   OLD BASIS FILE      0
   NEW BASIS FILE     11
   CRASH OPTION    1
   PARTIAL PRICE   2
   ITERATIONS 100
   LOG FREQUENCY   1
   SAVE FREQUENCY  200
   SOLUTION    YES
   PROBLEM NUMBER 1
   NONLINEAR VARIABLES  75
   LOWER BOUND = NONE
   SUPERBASIC LIMIT 40
   HESSIAN DIMENSION  40
   LINESEARCH TOL  0.01
   REDUCED-GRADIENT TOL  0.9
   VERIFY GRADIENTS
   CALL CALCFG IF OPTIMAL
END
```

References

Anderson, T.W., and L.A. Goodman. 1957. "Statistical Inference About Markov Chains." Annals of Mathematical Statistics 28:89-110.

Chow, G.C. 1960. "Statistical Demand Functions for Automobiles and Their Use for Forecasting." The Demand for Durable Goods, edited by A.C. Harbarger. Chicago, IL: University of Chicago Press, pp. 149-178.

Fox, Austin. 1966. Demand for Farm Tractors in the United States. Agr. Econ. Rpt. No. 103. Washington, D.C.: USDA.

Goodman, L.A. 1958. "A Further Note on Miller's Finite Markov Processes in Psychology." Psychometrika 18:245-48.

Griliches, Zvi. 1960. "The Demand for a Durable Input: Farm Tractors in the United States, 1921-57." The Demand for Durable Goods, edited by A.C. Harbarger. Chicago, IL: University of Chicago Press, pp. 181-207.

Judge, G.G., and E.R. Swanson. 1962. "Markov Chains: Basic Concepts and Suggested Uses in Agricultural Economics." Austrian Journal of Agricultural Economics 6:49-61.

Lee, T.C., G.G. Judge, and A. Zellner. 1970. Estimating the Parameters of the Markov Probability Model from Aggregate Time Series Data. Amsterdam: North-Holland Publishing Co..

Lee, T.C., G.G. Judge, and T. Takayama. 1965. "On Estimating the Transition Probabilities of a Markov Process." Journal of Farm Economics 47:742-62.

"Monthly Retail Sales and Inventory Report." 1964-1983. Farm and Industrial Equipment Institute, Chicago, IL.

Murtagh, Bruce A., and Michael A. Sanders. 1977. "MINOS Users' Guide." Technical Report 77-9. Stanford, CT: Stanford University, Systems Optimization Laboratory.

Telser, L.G. 1962. "The Demand for Branded Goods as Estimated from Consumer Panel Data." Review of Economic Statistics 44:300-24.

Telser, L.G. 1962. "Advertising and Cigarettes." Journal of Political Economy 70:471-99.

Theil, H., and Guido Rey. 1966. "A Quadratic Programming Approach to the Estimation of Transition Probabilities." Management Science 12:714-21.