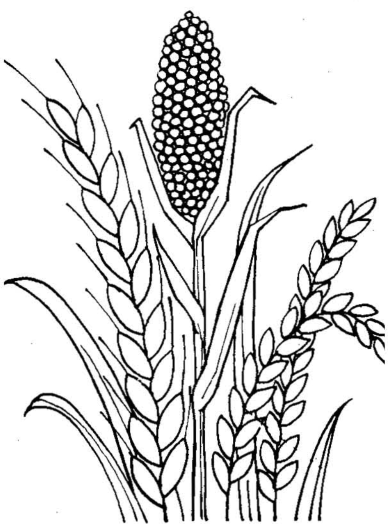
Vol XXVI
No. 2

APRIL-
JUNE
1971

ISSN 0019-5014

INDIAN JOURNAL OF AGRICULTURAL ECONOMICS



INDIAN SOCIETY OF
AGRICULTURAL ECONOMICS,
BOMBAY

APPENDIX

Output : Measured in current value terms. It is gross of all inputs and includes by-products.

Land input : Measured by the net sown area in acres and thus excludes any part of the farm area which remained uncultivated throughout the year. It also excludes area under structures that may be present in a field.

Bullock services : Measured in days of 8 hours of actual work. Since bullocks are worked in pairs with one man for each pair, the variable 'bullock team services' is defined as a unit comprising one bullock-day plus half man-day. This definition is necessary in order to allow for the substitutability between factors in the production function. Correspondingly the value of bullock team services is defined to be the sum of the cost of a day's bullock labour and the value of labour for half a day.

Labour input : Measured also in days of 8 hours and excludes labour used along with bullocks only. Value of labour input is measured in rupees at the going wage rate.

Seeds and manures : Measured in current value terms.

Capital services : The current market value of different kinds of farm assets, *viz.*, wells, farm buildings and implements and machinery were converted into flow of services through "annuity" approach separately for the three categories using a length of life of 50, 25 and 8 years respectively based on the information in the published source and a rate of discount of 3, 5 and 7 per cent. Thus, three measures of capital services were constructed corresponding to the three discount rates for each of the three types of assets and finally aggregated to get the total flow of capital services. This quantity represents the potential flow through the year. To arrive at the actual utilization the above quantity was multiplied by the fraction of the year for which land was utilized (cropped) on the average.*

Irrigation input : Measured in gross area irrigated.

Prices : Price of output refers to the harvest period; price of land services, labour and bullock services evaluated at the going rate in the market; price of land and other items of capital stock also evaluated at the going market rate.

Quality measures : One measure of bullock quality is the value of the animal itself. Another measure of the same is the cost of maintenance of an animal per year. Still another measure is the ratio of value of concentrates fed to the animal to the value of other feeds. The bullock quality measure used for the regressions reported is the first one, *i.e.*, value of the animal.

Intensity of land utilization : Measured as the quotient of gross sown area (*i.e.*, net sown area plus area cropped more than once) to net sown area.

APPLICATION OF MARKOV CHAIN ANALYSIS IN SHORT-TERM DEMAND FORECASTING FOR AGRICULTURAL INPUTS†

In recent years many studies have attempted to estimate the demand for agricultural inputs like fertilizers, pesticides, and seeds, based on the area under different crops and the rate of application of the inputs. From the point of view of economic theory, the demand for an input is a derived demand and is dependent on the price of the input, price of the output for which the particular input is used, and the income level of the producers. However, in order to estimate such demand relationships it is necessary to have fairly long time series data on these variables. In the absence of such data, attempts were made to collect cross-section data on the per acre use of the inputs and the proportion of area in which the input

* If the intensity of land utilization is one then it is assumed that capital is utilized at 5/12 of the potential. In so far as the intensity of land utilization varied between observations, the fraction by which capital is utilized also varies since the rate of capital utilization=intensity of land utilization × 5/12.

† The author would like to acknowledge the useful comments received from Dr. D. K. Desai and the referees of the Journal on an earlier draft of this paper.

was used. Here, the estimate of demand for an input, say fertilizer, was derived from the aggregate of cropwise use of fertilizers obtained as a product of the area under the crop and the per acre application rate. In order to use such estimates for forecasting demand, predicted values of the two variables, *viz.*, area under the crop and the rate of application are generally used. The major difficulty with respect to this approach lies in the fact that it is difficult to identify the variables influencing changes in area and adoption of practices, and even in such cases where variables can be identified they may be associated with a number of random variables whose probability distributions are unknown. The method of forecasting demand using trend projection also may suffer from such disadvantages. In this paper an attempt is made to develop an alternative methodology, using the transition probabilities associated with the Markov Chain Analysis,[1] to forecast short-term changes in the demand for seeds of high-yielding variety (HYV) paddy.

Methodology

The demand for high-yielding variety paddy seed is a function of the rate of adoption of high-yielding varieties by the farmers and the seed rate per acre. Since the seed rate is specified by agronomic factors, most of the adopters use more or less the prescribed seed rate, and therefore, the percentage change in the demand for seed can be assumed to be proportional to the change in the rate of adoption of high-yielding variety paddy. Suppose there are N° farmers in a given region belonging to different stages of adoption. At a given time period 'O', the N° farmers may belong to k adoption groups containing $n^\circ_{1.}$, $n^\circ_{2.}$, ..., $n^\circ_{k.}$ farmers such that $n^\circ_{1.}+n^\circ_{2.}+$ ...., $n^\circ_{k.} = N^\circ$. The adoption groups can be defined in terms of the proportion of paddy area under high-yielding varieties. Suppose that data are available on the behaviour of these N° farmers for another time period 't'. During the interval between the two periods, the $n^\circ_{1.}$ farmers belonging to the first adoption group during the first period might have moved to higher groups of adoption or remained in the same group. Table I describes the pattern of change during the intervals between the two time period 'O' and 't'.

TABLE I—NUMBER OF FARMERS IN DIFFERENT ADOPTION GROUPS

Adoption groups: Period 'O'	Adoption groups: Period 't'					
	1	2	3..........	j	k	Total
1	n_{11}	n_{12}	n_{13}		n_{1k}	$n_{1.}$
2	n_{21}	n_{22}	n_{23}		n_{2k}	$n_{2.}$
3						
i	n_{i1}	....	....	n_{ij}..	n_{ik}	$n_{i.}$
k	n_{k1}	n_{k2}	n_{k3}	n_{kj}..	n_{kk}	$n_{k.}$
Total ..	$n_{.1}$	$n_{.2}$	$n_{.3}$	$n_{.j}$	$n_{.k}$	N°

1. The Markov Chain Analysis is described in most of the standard textbooks of statistics. For a brief description of this method, see C. R. Rao : Advanced Statistical Methods in Biometric Research, John Wiley & Sons, New York, U.S.A., 1962, pp. 27-28.

The sum of the j^{th} column, *viz.*, $n_{.j}$ represents the number of farmers belonging to the j^{th} adoption group during the period t $(j=1, 2, \ldots, k)$. The probability that any farmer in the first group moves to the j^{th} group during the second period is given by $P_{1j} = \frac{n_{1j}}{n_{1.}}$, $(j=1, 2, \ldots, k)$. In general, the probability that a farmer belonging to the i^{th} adoption group during the initial period 'O' moves to the j^{th} adoption group during the period 't' is given by $P_{ij} = \frac{n_{ij}}{n_{i.}}$; $(i, j = 1, 2, \ldots, k)$. These probabilities are known as transition probabilities and they can be represented as in Table II.

TABLE II—TRANSITION PROBABILITY MATRIX

Adoption group : Period 'O'	Adoption group : Period 't'			
	1	2	j	.... k
1	P_{11}	P_{12}	P_{1j}...............	P_{1k}
2	P_{21}	P_{22}	P_{2j}...............	P_{2k}
3				
i	P_{i1}	P_{i2}	P_{ij}...............	P_{ik}
k	P_{k1}	P_{k2}	P_{kj}...............	P_{kk}

The transition probabilities, defined above, can be used to predict the distribution of farmers in different adoption groups for subsequent periods. The only assumption required for obtaining a reasonable forecast is that the structure of the population remains essentially the same for the period for which a forecast is made, as in the period which forms the basis for deriving the transition probability matrix. From Table I, it may be observed that the number of farmers in different adoption categories during the period 't' is given by $n_{.1}, n_{.2}, \ldots, n_{.k}$, and also the transition probability matrix P gives the probabilities associated with all possible levels of movements. Consider the product of the vector $(n_{.1}, n_{.2}, \ldots, n_{.k})$ and the first column of the transition probability matrix,

$$(n_{.1}, n_{.2}, \ldots, n_{.k}) \begin{pmatrix} P_{11} \\ P_{21} \\ P_{k1} \end{pmatrix} = n_{.1} P_{11} + n_{.2} P_{21} + \ldots + n_{.k} P_{k1} = \sum_{j=1}^{k} n_{.j} P_{j1}$$

The sum $\sum_{j=1}^{k} n_{.j} P_{j1}$ will give the total number of farmers in the first adoption group during the next period t_1. Similarly, the product of $(n_{.1}, n_{.2}, \ldots, n_{.k})$ and the i^{th} column in the transition matrix gives the number of farmers in the i^{th} adoption group during the period t_1. In other words, the k elements in the product

$$(n_{\cdot 1}, \ldots\ldots, n_{\cdot k}) \begin{pmatrix} P_{11}, \ldots\ldots\ldots\ldots, P_{1k} \\ P_{k1} \ldots\ldots\ldots\ldots, P_{kk} \end{pmatrix}$$

will give the number of persons in each adoption group during the period t_1. If we denote the vector $n^\circ = (n_{\cdot 1}, n_{\cdot 2}, \ldots\ldots, n_{\cdot k})$ as the initial distribution and the transition matrix P, then $n^\circ P$ will give the distribution of farmers in different adoption groups during the next period. Now during the subsequent time period $n^\circ P$ becomes the initial distribution, and, following the same argument as above, it can be shown that the distribution of farmers in different adoption groups during this period is given by $(n^\circ P) P = n^\circ P^2$. The distribution of farmers in different adoption groups during m periods beyond the last observed period will be given by $n^\circ P^m$. Thus, the initial distribution of farmers n° and the transition matrix P completely define a Markov Chain process, and the distribution at the end of any period m can be completely determined by the product $n^\circ P^m$.

The next step is to show how the predicted distribution of farmers in different adoption groups can be used to forecast the increase in demand for high-yielding variety paddy seed. During the initial period the number of farmers in different adoption groups was known, and therefore, the proportion of farmers in each adoption group can be calculated from the distribution data. These proportions multiplied by the respective mid-points of the adoption groups will give us the expected rate of adoption during the initial period. If $M_1, M_2, \ldots\ldots\ldots, M_k$ correspond to the middle points of the k adoption groups, E°, the expected rate of adoption during the initial period is given by $E^\circ = \sum_{j=1}^{k} \frac{n_{\cdot j}}{N} M_j$. We shall call this expected rate of adoption as the adoption index. Similarly, using the projected distribution during the period t_m, $n^\circ P^m$, we can calculate the adoption index E^m for that period. The rate of change of the adoption index during the interval between period t_o and t_m is given by $\frac{E^m - E^\circ}{E^\circ} \times 100 = \left(\frac{E^m}{E^\circ} - 1\right) 100$. Since we have assumed that the seed rate is independent of the rate of adoption, the increase in demand for high-yielding variety seed will come mainly from the increase in the area under high-yielding variety paddy. The changes in the adoption index essentially reflects the changes in the area under high-yielding variety, and therefore $100\left(\frac{E^m}{E^\circ} - 1\right)$ can be taken as the percentage increase in seed requirement during the period t_m.

Empirical Example

We have collected data on the area under the local and the high-yielding variety paddy from 37 farmers in the West Godavari district of Andhra Pradesh for the *rabi* cropping season during 1968-69. The same farmers were contacted after one year and data were collected on the area under both these varieties during the *rabi* season 1969-70. From the area under high-yielding variety, the percentage area under this variety during both the years was calculated and the farmers were grouped into seven adoption categories based on the area under HYV. Table III gives the distribution of farmers in different adoption groups.

TABLE III—DISTRIBUTION OF FARMERS IN DIFFERENT ADOPTION GROUPS DURING *Rabi* 1968-69 AND 1969-70

Rabi 1968-69 Percentage of area under HYV	*Rabi* 1969-70 Percentage of area under HYV							Total
	0-10	10-20	20-35	35-50	50-60	60-90	90-100	
0—10	6	1		2	1		4	14
10—20	1		1					2
20—35							1	1
35—50							1	1
50—60							2	2
60—90	1						3	4
90—100	1		1	1			10	13
Total	9	1	2	3	1	0	21	37

From Table III, it can be observed that the distribution of 37 farmers in the seven adoption groups during *rabi* 1969-70 was 9, 1, 2, 3, 1, 0, and 21 respectively. This forms the vector of initial distribution for the purpose of further analysis. Based on the movement of farmers in different adoption groups during the two crop season, transition probabilities were calculated. Table IV gives these probabilities.

TABLE IV—TRANSITION PROBABILITIES

Percentage area under HYV in 1968-69	*Rabi* 1969-70 Percentage of area under HYV						
	0-10	10-20	20-35	35-50	50-60	60-90	90-100
0—10	0.429	0.071	0	0.143	0.071	0	0.286
10—20	0.500	0	0.500	0	0	0	0
20—35	0	0	0	0	0	0	1.000
35—50	0	0	0	0	0	0	1.000
50—60	0	0	0	0	0	0	1.000
60—90	0.250	0	0	0	0	0	0.750
90—100	0.077	0	0.077	0.077	0	0	0.769

Using the initial distribution of farmers in different adoption groups as the distribution during *rabi* 1969-70 and the transition probabilities, the distribution of farmers in the seven groups was predicted for the subsequent years. These predicted values were used to calculate the adoption index and they formed the basis for estimating the anticipated increase in the demand for high-yielding variety paddy seed. Table V gives the predicted distribution of farmers and Table VI gives the anticipated increase in demand.

TABLE V—PREDICTED DISTRIBUTION OF FARMERS

Adoption group	Actuals		Predicted				
	1968-69	1969-70	1970-71	1971-72	1972-73	1973-74	1974-75
0—10	14	9	6	5	4	4	4
10—20	2	1	(.5)	(.5)	(.5)	(.5)	(.5)
20—35	1	2	2	2	2	2	2
35—50	1	3	3	3	3	3	3
50—60	2	1	(.5)	(.5)	(.5)	(.5)	(.5)
60—90	4	—	—	—	—	—	—
90—100	13	21	25	26	27	27	27
Total ..	37	37	37	37	37	37	37

TABLE VI—ADOPTION INDICES AND GROWTH RATE

Year	Adoption index	Percentage increase with 1969-70 as base
1969-70	61.98	—
1970-71	70.88	11.44
1971-72	73.31	11.83
1972-73	75.74	12.22
1973-74	75.74	12.22

The predicted distribution of farmers in the different adoption groups shows interesting results. If the present technology continues, the adoption rate will be stabilized by 1972-73. Even at that point there will be movement from one group to another, but the number of persons in different adoption groups will not change. (The probabilities of transition from the stage in 1969-70 to another stage in the years 1971-72, 1972-73, and 1973-74 are given in the Appendix.) The number of persons in the adoption groups (10-20) and (50-60) comes out to be the same (in both the cases the number is around .5) and therefore in Table V it is reported as .5 without rounding it of. The stabilization tendency shown in Table V is again reflected in Table VI. Though there is a predicted increase in the demand for high-yielding variety seed by 11.44 per cent during the year 1970-71 over the demand in 1969-70, in the subsequent years, the increase in demand is not substantially high. This is expected as the adoption level of seed achieves a stabilized position within the first few years. At this stage, if new varieties of seed are not available, the chances of further adoption may be very small. Here it may be noted that the results are point estimates, and these estimates are likely to have large standard errors associated with them. Since the calculations of these errors might be a tedious job, here no attempt was made to obtain the possible range in which the growth rates would fall. Also the transition probabilities are derived from only two cross-sections. The validity of the transition probability matrix would have been enhanced if data were available for more time periods. A large number of observations would have reduced the investigator's bias in reporting the adoption rate.

Summary and Conclusions

The demand for many agricultural inputs is determined by a number of stochastic variables that cannot be taken into account in demand forecasts based on standard regression analysis or cross-section analysis. The present paper provides an alternative approach to project the demand for an input based on Markov Chain Analysis. The basic assumption behind this approach is that the farmer's adoption or use rate of an input can be determined using the initial level of adoption and the probability associated with moving to another level of adoption. The distribution of farmers in the different adoption groups could be first determined and this distribution could be used to forecast the demand for the input. This approach was illustrated using cross-section data on high-yielding varieties of paddy by 37 farmers in the West Godavari district during *rabi* 1968-69 and 1969-70. The results indicate that the level of adoption will tend to stabilize around 75 per cent of the area under the *rabi* crop if new varieties of paddy are not introduced in subsequent periods.

While using this approach, it is important to remember some of the limitations of this approach. The Markov Chain Analysis assumes the inter-temporal constancy of the transition probability matrix, and therefore, it does not incorporate any structural changes over the years. The only possible solution in this case may be to make allowance in the final forecast to take structural changes into account.

Another problem associated with this approach is regarding the difficulties in obtaining the standard errors associated with the estimates. Since the estimation of standard errors is a difficult job, often it will be possible to obtain only a point estimate. The reliability of these point estimates could be increased if the transition probability matrix is derived using data from a number of time periods. Also the forecasts will be more appropriate to periods immediately after the last period for which observations are available as compared to periods of longer intervals.

The basic model presented here can be modified to incorporate long-term projections of demand for the inputs. Here the deciding factor is the availability of data to generate the transition probability matrix. Suppose there exist data for three time periods so that the probability matrix may indicate the probability associated with a change during two-year period. The product of the initial distribution and the transition matrix will give the distribution at the end of the next two years. The subsequent periods will be four years, six years, and so on. In general, if the transition matrix is based on $(k+1)$ periods, the projections will cover k, 2k, 3k and so on.

In the case of certain inputs it may be possible to isolate the contribution of some known factors on the observed behaviour. In this case it is desirable to remove the effects on these known factors and run the Markov Analysis based on the unexplained part. Essentially here we are separating out the observed data into two components, one corresponding to known factors and the other corresponding to stochastic variables and the Markov process is applied to the stochastic components. In the seed example used here we have assumed the entire variations to be influenced by stochastic variables alone.

In spite of all these limitations, the present approach could be used to obtain an alternate estimate of demand. This approach would prove to be useful especially in cases where long time series data are not available on the use of agricultural inputs in a given region or when the adoption rate is subject to a lot of random fluctuations. When other methods are used, the estimates obtained from this approach could be used to compare the other estimates.

P. S. GEORGE*

APPENDIX

PROBABILITIES OF TRANSITION FROM 1969-70 STAGE IN THE YEARS 1971-72, 1972-73 AND 1973-74

A =	9	1	2	3	1	0	21
(a) 1971-72	.242	.030	.057	.083	.030	0	.556
	.214	.035	0	.071	.035	0	.643
	.077	0	.077	.077	0	0	.769
	.077	0	.077	.077	0	0	.769
	.077	0	.077	.077	0	0	.769
	.165	.017	.057	.093	.017	0	.648
	.092	.005	.059	.070	.005	0	.767
(b) 1972-73	.161	.017	.058	.077	.017	0	.668
	.159	.015	.067	.080	.015	0	.663
	.092	.005	.059	.070	.005	0	.767
	.092	.005	.059	.070	.005	0	.767
	.092	.005	.059	.070	.005	0	.767
	.129	.011	.059	.073	.011	0	.715
	.101	.006	.061	.072	.006	0	.751
(c) 1973-74	.129	.011	.060	.074	.011	0	.712
	.127	.011	.058	.073	.011	0	.718
	.101	.006	.061	.072	.006	0	.751
	.101	.006	.061	.072	.006	0	.751
	.101	.006	.061	.072	.006	0	.751
	.116	.009	.060	.073	.009	0	.730
	.104	.007	.061	.072	.007	0	.747

* Indian Institute of Management, Vastrapur, Ahmedabad-15.