*The Annals of Applied Probability*
1992, Vol. 2, No. 4, 1019–1023

# MARKET SHARE PARADOX AND HETEROGENEOUS CHAINS

By Yoav Benjamini and Abba M. Krieger

*Tel Aviv University and University of Pennsylvania*

This resolves a paradox in an application of a Markov chain model to market share estimation.

**1. Introduction.** The paradox addressed here concerns data from a study of a subpopulation of IBM mainframe users. At the time of the study, there were three user categories of concern:

1. Customers with full-service IBM contracts, about 94% of the subpopulation.
2. Customers paying IBM for service on a per call basis, about 4%.
3. Customers using non-IBM service, about 2%.

A uniform random sample of customers was drawn from the population and these customers were interviewed concerning their intentions for the coming year. Estimates $a_i$ and $p_{ij}$ were obtained for the probability of a customer being in category $i$ (at the time of the survey) and the probability of a transition from category $i$ to category $j$ during the next year. When the $p_{ij}$ were used as elements of a transition matrix and a stationary distribution $(\pi_1, \pi_2, \pi_3)$ was computed, it was found that $\pi_1$ was approximately 3%. The implication that the fraction of full-service users would shrink from 94% to 3% seemed inconceivable, so the fault of the model had to be explored.

The results obtained here show that the possibility of heterogeneity within the population is all one needs to dispel the initial alarm. This possibility is one that had been considered in the marketing literature [see Morrison, Massy and Silverman (1971)], but which had been dismissed as not making enough difference. Thus, this note also serves to adjust some conventional thinking.

Our main result, as well as the formalities needed for its careful expression, are laid out in the next section.

**2. Main results.** We are concerned with two different models. First, we have a Markov chain model where each member of a population undergoes transitions according to a matrix $P = \{p_{ij}\}$. A second model consists of a heterogeneous population with individuals of $S$ different types. The individuals of the $s$th type are assumed to undergo transitions according to a matrix $P(s) = \{p_{ij}(s)\}$.

Received January 1990; revised February 1992.
*AMS* 1980 *subject classifications*. Primary 60J10; secondary 62M05.
*Key words and phrases*. Markov chain, market share, heterogeneity, stationary distribution.



Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Applied Probability*. www.jstor.org

Naturally, we insist that the models be compatible in the sense that if we generate interview data consistent with the second model and we use a maximum likelihood estimator, we are led back to $\{p_{ij}\}$. To formalize this we first let $\alpha(s)$ denote the actual fraction of individuals from the complete population in the $s$th subpopulation. Next, we let $a_i(s)$ denote the observed fraction of individuals in the $s$th subpopulation that are in state $i$ at the time of survey. Recalling that $a_i$ is the original estimate of a customer being in state $i$, we see the compatibility conditions can be given formally as

$$\sum_s \alpha(s) = 1,$$

$$\sum_s \alpha(s) a_i(s) = a_i$$

and

$$\sum_s \alpha(s) a_i(s) p_{ij}(s) = a_i p_{ij}.$$

A small technical point worth recording now is that we will assume that the estimates $a_i$ and $p_{ij}$ are all strictly positive. The message is unaffected by this assumption and some messy trivialities are also avoided.

Finally, we let $\lambda(M)$ be the proportion of individuals who are in state 1 when the heterogeneous model $M = \{\alpha(s), a(s), P(s);\ s = 1, \dots, S\}$ reaches steady state. If we let $\lambda^*(P)$ and $\lambda_*(P)$ denote the supremum and infimum of $\lambda(M)$ over all $M$ compatible with $P$, then the key concern boils down to seeing how different $\lambda^*(P)$ and $\lambda_*(P)$ can be.

Our main result tells us heterogeneity can indeed have a very substantial impact on the steady state distributions.

THEOREM. *For all $P$, we have $\lambda^*(P) - \lambda_*(P) \geq 1/2$.*

**3. Proof of the main result.** First, we consider the case of a two state problem. The idea of the proof is that for any $P$ we can define heterogeneous models compatible with $P$ as follows:

*Model $M_1$.*

$$\alpha(1) = a_2 + a_1 p_{11}, \qquad a(1) = \begin{pmatrix} \dfrac{a_1 p_{11}}{\alpha(1)} \\ \dfrac{a_2}{\alpha(1)} \end{pmatrix}, \qquad p(1) = \begin{pmatrix} 1-\delta & \delta \\ p_{21} & p_{22} \end{pmatrix},$$

$$\alpha(2) = a_1 p_{12}, \qquad a(2) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad p(2) = \begin{pmatrix} \dfrac{p_{11}\delta}{p_{12}} & 1 - \dfrac{p_{11}\delta}{p_{12}} \\ 1 & 0 \end{pmatrix}.$$

*Model $M_2$.*

$$\alpha(1) = a_1 + a_2 p_{22}, \qquad a(1) = \begin{pmatrix} \dfrac{a_1}{\alpha(1)} \\ \dfrac{a_2 p_{22}}{\alpha(1)} \end{pmatrix}, \qquad p(1) = \begin{pmatrix} p_{11} & p_{12} \\ \delta & 1-\delta \end{pmatrix},$$

$$\alpha(2) = a_2 p_{21}, \qquad a(2) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad p(2) = \begin{pmatrix} 0 & 1 \\ 1 - \dfrac{p_{22}\delta}{p_{21}} & \dfrac{p_{22}\delta}{p_{21}} \end{pmatrix}.$$

Since $\lambda(M_1) \to 1 - a_1 p_{12}/2$ as $\delta \to 0$ and $\lambda(M_2) \to a_2 p_{21}/2$ as $\delta \to 0$, then $\lambda^*(P) - \lambda_*(P) \geq 1 - a_1 p_{12}/2 - a_2 p_{21}/2 \geq 1/2$.

When the number of states $T \geq 3$, we again prove the theorem by describing two heterogeneous models $M_1$ and $M_2$ each with two subpopulations. Once $M_1$ and $M_2$ are presented for $T = 3$, it is easy to see how to construct $M_1$ and $M_2$ for $T > 3$.

*Model $M_1$.*

$$\alpha(1) = 1 - a_1(1 - p_{11}), \qquad a(1) = \begin{pmatrix} \dfrac{a_1 p_{11}}{\alpha(1)} \\ \dfrac{a_2}{\alpha(1)} \\ \dfrac{a_3}{\alpha(1)} \end{pmatrix}, \qquad p(1) = \begin{pmatrix} 1-\delta & \delta & 0 \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{pmatrix},$$

$$\alpha(2) = a_1(1 - p_{11}), \qquad a(2) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},$$

$$p(2) = \begin{pmatrix} \dfrac{p_{11}\delta}{1 - p_{11}} & \dfrac{p_{12} - p_{11}\delta}{1 - p_{11}} & \dfrac{p_{13}}{1 - p_{11}} \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

*Model $M_2$.*

$$\alpha(1) = \frac{[a_2 p_{22} + a_3(1 - p_{33})]}{(a_2 + a_3)}, \qquad a(1) = \begin{pmatrix} a_1 \\ \dfrac{a_2 p_{22}}{\alpha(1)} \\ \dfrac{a_3(1 - p_{33})}{\alpha(1)} \end{pmatrix},$$

$$p(1) = \begin{pmatrix} p_{11} & p_{12} & p_{13} \\ 0 & 1 & 0 \\ \dfrac{p_{31}}{1 - p_{33}} & \dfrac{p_{32}}{1 - p_{33}} & 0 \end{pmatrix},$$

$$\alpha(2) = \frac{[a_2(1 - p_{22}) + a_3 p_{33}]}{(a_2 + a_3)}, \qquad a(2) = \begin{pmatrix} a_1 \\ \dfrac{a_2(1 - p_{22})}{\alpha(2)} \\ \dfrac{a_3 p_{33}}{\alpha(2)} \end{pmatrix},$$

$$p(2) = \begin{pmatrix} p_{11} & p_{12} & p_{13} \\ \dfrac{p_{21}}{1 - p_{22}} & 0 & \dfrac{p_{23}}{1 - p_{22}} \\ 0 & 0 & 1 \end{pmatrix}.$$

Since $\lambda(M_1) \to 1 - a_1(1 - p_{11})/2$ as $\delta \to 0$ and $\lambda(M_2) = 0$, then $\lambda^*(P) - \lambda_*(P) \geq 1 - a_1(1 - p_{11})/2 \geq 1/2$.

**4. More precise result.** More specifically, we can show that $\lambda^*(P) = 1 - a_1(1 - p_{11})/2$ and

$$\lambda_*(P) = \begin{cases} a_2 p_{21}/2, & \text{if } T = 2, \\ 0, & \text{if } T \geq 3. \end{cases}$$

The fact that $\lambda_*(P) = 0$ if $T \geq 3$ follows from $M_2$ described above. The result for $\lambda_*(P)$ when $T = 2$ is the same as the result for $\lambda^*(P)$ when $T = 2$, since maximizing the proportion of individuals in state 2 is equivalent to minimizing the proportion in state 1.

For any $P$, consider the transition matrix $\tau(P)$ whose first row is the same as the first row in $P$ and all other rows are $(1, 0, \dots, 0)$. Clearly, $\lambda^*(\tau(P)) \geq \lambda^*(P) \geq 1 - a_1(1 - p_{11})/2$. The second inequality follows from $M_1$ described above. Let $\Omega$ be the class of transition matrices with $(1, 0, \dots, 0)$ in all but row 1. It suffices to show that for any $P \in \Omega$, $\lambda^*(P) = 1 - a_1(1 - p_{11})/2$. If $P \in \Omega$ and $M$ is compatible with $P$, then $P(s) \in \Omega$ for every $s$. However, the transition matrix for the two state problem with first row $(p_{11}, 1 - p_{11})$ and second row $(1, 0)$ has the same steady state probability for state 1 as any $P \in \Omega$ with transition from state 1 to state 1 of $p_{11}$. Hence, we only need to consider $T = 2$.

Let $M$ be a heterogeneous model compatible with $P$ for the two state problem. Then

$$\begin{aligned}\lambda(M) &= \sum_s \alpha(s) p_{21}(s)/(p_{21}(s) + p_{12}(s)) \\ &= \sum_s \alpha(s)[1 - (1 - p_{11}(s))/(p_{21}(s) + p_{12}(s))] \\ &\leq \sum_s \alpha(s)[1 - a_1(s)(1 - p_{11}(s))/2] = 1 - a_1(1 - p_{11})/2.\end{aligned}$$

The inequality follows from $a_1(s) \leq 1$ and $p_{21}(s) + p_{12}(s) \leq 2$, and the last equality follows from the compatibility conditions.

**5. Conclusion.** Since steady state probabilities are generally affected by heterogeneity, it is important to collect data so that potential heterogeneity is estimable from the data. This suggests that we collect longitudinal data (i.e., history of states for each respondent). We can then use the EM algorithm [cf. Dempster, Laird and Rubin (1977)] to estimate the parameters assuming that we know the number of subpopulations. This approach has been developed by van de Pol and Langeheine (1989). Our results suggest that parameter values near the extreme cases in this paper be considered, particularly since it was shown that the likelihood function for this problem potentially has many local maxima.

**Acknowledgments.** The authors are extremely grateful to Mike Steele and two anonymous referees for many helpful suggestions.

## REFERENCES

DEMPSTER, A. P., LAIRD, N. M. and RUBIN, D. R. (1977). Maximum likelihood from incomplete data. *J. Roy. Statist. Soc. Ser. B* **37** 1–38.

MORRISON, D. G., MASSY, W. F. and SILVERMAN, F. N. (1971). The effect of nonhomogeneous populations on Markov steady-state probabilities. *J. Amer. Statist. Assoc.* **66** 268–274.

VAN DE POL, F. and LANGEHEINE, R. (1989). Mixed Markov models, mover–stayer models and the EM algorithm. In *Multiway Data Analysis* (R. Coppi and S. Bolasco, eds.) 485–495. North-Holland, Amsterdam.

DEPARTMENT OF STATISTICS
TEL AVIV UNIVERSITY
TEL AVIV 69978
ISRAEL

DEPARTMENT OF STATISTICS
UNIVERSITY OF PENNSYLVANIA
THE WHARTON SCHOOL
3011 STEINBERG-DIETRICH HALL
PHILADELPHIA, PENNSYLVANIA 19104