A NEW PRODUCT GROWTH FOR MODEL CONSUMER DURABLES*, 1

FRANK M. BASS

Purdue University

A growth model for the timing of initial purchase of new products is developed and tested empirically against data for eleven consumer durables. The basic assumption of the model is that the timing of a consumer's initial purchase is related to the number of previous buyers. A behavioral rationale for the model is offered in terms of innovative and imitative behavior. The model yields good predictions of the sales peak and the timing of the peak when applied to historical data. A long-range forecast is developed for the sales of color television sets.

The concern of this paper is the development of a theory of timing of initial purchase of new consumer products. The empirical aspects of the work presented here deal exclusively with consumer durables. The theory, however, is intended to apply to the growth of initial purchases of a broad range of distinctive "new" generic classes of products. Thus, we draw a distinction between new classes of products as opposed to new brands or new models of older products. While further research concerning growth rate behavior is currently in process for a wider group of products, attention focuses here exclusively upon infrequently purchased products.

Haines [6], Fourt and Woodlock [5], and others have suggested growth models for new brands or new products which suggests exponential growth to some asymptote. The growth model postulated here, however, is best reflected by growth patterns similar to that shown in Figure 1. Sales grow to a peak and then level off at some magnitude lower than the peak. The stabilizing effect is accounted for by the relative growth of the replacement purchasing component of sales and the decline of the initial purchase component. We shall be concerned here only with the timing of initial purchase.

Long-range forecasting of new product sales is a guessing game, at best. Some things, however, may be easier to guess than others. The theoretical framework presented here provides a rationale for long-range forecasting. The theory stems mathematically from the contagion models which have found such widespread application in epidemology [2]. Behaviorally, the assumptions are similar in certain respects to the theoretical concepts emerging in the literature on new product adoption and diffusion, [7, 8, 9, 13] as well as to some learning models [3, 12]. The model differs from models based on the log-normal distribution [1, 4, 10] and other growth models [11] in that the behavioral assumptions are explicit.

The Theory of Adoption and Diffusion

The theory of the adoption and diffusion of new ideas or new products by a social system has been discussed at length by Rogers [13]. This discussion is largely literary. It is, therefore, not always easy to separate the premises of the theory from the conclusions. In the discussion which follows an attempt will be made to outline the major ideas of the theory as they apply to the *timing* of adoption.

^{*} Received March 1967 and revised August 1967.

¹ Some of the basic ideas in this paper were originally suggested to the author by Peter Frevert, now of the University of Kansas. Thomas H. Bruhn, Gordon Constable, and Murray Silverman provided programming and computational assistance.

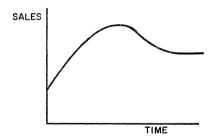


Fig. 1. Growth of a new product

Some individuals decide to adopt an innovation independently of the decisions of other individuals in a social system. We shall refer to these individuals as *innovators*. We might ordinarily expect the first adopters to be innovators. In the literature, the following classes of adopters are specified: (1) Innovators; (2) Early Adopters; (3) Early Majority; (4) Late Majority; and (5) Laggards. This classification is based upon the timing of adoption by the various groups.

Apart from innovators, adopters are influenced in the timing of adoption by the pressures of the social system, the pressure increasing for later adopters with the number of previous adopters. In the mathematical formulation of the theory presented here we shall aggregate groups (2) through (5) above and define them as *imitators*. Imitators, unlike innovators, are influenced in the timing of adoption by the decisions of other members of the social system. Rogers defines innovators, rather arbitrarily, as the first two and one-half percent of the adopters. Innovators are described as being venturesome and daring. They also interact with other innovators. When we say that they are not influenced in the timing of purchase by other members of the social system, we mean that the pressure for adoption, for this group, does not increase with the growth of the adoption process. In fact, quite the opposite may be true.

In applying the theory to the timing of initial purchase of a new consumer product, we formulate the following precise and basic assumption which, hopefully, characterizes the literary theory: The probability that an initial purchase will be made at T given that no purchase has yet been made is a linear function of the number of previous buyers. Thus, P(T) = p + (q/m)Y(T), where p and q/m are constants and Y(T) is the number of previous buyers. Since Y(0) = 0, the constant p is the probability of an initial purchase at T = 0 and its magnitude reflects the importance of innovators in the social system. Since the parameters of the model depend upon the scale used to measure time, it is possible to select a unit of measure for time such that p reflects the fraction of all adopters who are innovators in the sense in which Rogers defines them. The product q/m times Y(T) reflects the pressures operating on imitators as the number of previous buyers increases.

In the section which follows, the basic assumption of the theory will be formulated in terms of a continuous model and a density function of time to initial purchase. We shall therefore refer to the linear probability element as a likelihood.

Assumptions and the Model

The following assumptions characterize the model:

a) Over the period of interest ("life of the product") there will be m initial purchases of the product. Since we are dealing with infrequently purchased products, the unit sales of the product will coincide with the number of initial purchases during that part

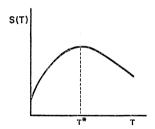


Fig. 2. Growth rate (q > p)

of the time interval for which replacement sales are excluded. After replacement purchasing begins, sales will be composed of both initial purchases and replacement purchases. We shall restrict our interest in sales to that time interval for which replacement sales are excluded, although our interest in initial purchase will extend beyond this interval.

b) The likelihood of purchase at time T given that no purchase has yet been made is

$$[f(T)]/[1-F(T)] = P(T) = p + q/m Y(T) = p + q F(T),$$

where f(T) is the likelihood of purchase at T and

$$F(T) = \int_0^T f(t) dt,$$
 $F(0) = 0.$

Since f(T) is the likelihood of purchase at T and m is the total number purchasing during the period for which the density function was constructed,

$$Y(T) = \int_0^T S(t) dt = m \int_0^T f(t) dt = mF(T)$$

is the total number purchasing in the (0, T) interval. Therefore, sales at T =

$$S(T) = mf(T) = P(T)[m - Y(T)] = \left[p + q \int_0^T S(t) dt/m\right] \left[m - \int_0^T S(t) dt\right].$$

Expanding this product we have

$$S(T) = p m + (q - p)Y(T) - q/m[Y(T)]^{2}.$$

The behavioral rationale for these assumptions are summarized:

- a) Initial purchases of the product are made by both "innovators" and "imitators," the important distinction between an innovator and an imitator being the buying influence. Innovators are not influenced in the timing of their initial purchase by the number of people who have already bought the product, while imitators are influenced by the number of previous buyers. Imitators "learn," in some sense, from those who have already bought.
- b) The importance of innovators will be greater at first but will diminish monotonically with time.
- c) We shall refer to p as the coefficient of innovation and q as the coefficient of imitation.

Since f(T) =

$$[p + q F(T)][1 - F(T)] = p + (q - p)F(T) - q[F(T)]^{2},$$

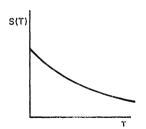


Fig. 3. Growth rate $(q \le p)$

TABLE I
Growth Model Regression Results For Eleven Consumer Durable Products

Product	Period covered	a (10 ²)	b	c (10 ⁻⁷)	R2	a/σ _a	b/σ _b	c/σ _c	m (10³)	Þ	q
Electric refrig- erators	1920–1940	104.67	.21305	053913	.903	1.164	6.142	-2.548	40,001	.0026167	.21566
Home freezers	1946-1961	308.12	.15298	077868	.742	4.195	4.769	-3.619	21,973	.018119	.17110
Black and white television	1946-1961	2,696.2	. 22317	025957	.576	3.312	3.724	-3.167	96,717	.027877	.25105
Water softeners	1949-1961	.10256	.27925	-512.59	.919	3.593	8.089	-6.451	5,793	.017703	.29695
Room air con- ditioners	1946-1961	175.69	.40820	24777	.911	1.915	8.317	-6.034	16,895	.010399	.41861
Clothes dryers	1948-1961	259.67	.33968	23647	.896	2.941	7.427	-5.701	15,092	.017206	.35688
Power lawn- mowers	1948-1961	410.98	.32871	075506	.932	1.935	7.408	-4.740	44,751	.0091837	.33 7 9 0
Electric bed coverings	1949-1961	450.04	.23800	031842	.976	3.522	6.820	-1.826	76,589	.005876	.24387
Automatic cof- fee makers	1948-1961	1,008.2	.28435	051242	.883	3.109	6.186	-4.353	58,838	.017135	.30145
Steam irons	1949-1960	1,594.7	.29928	058875	.828	3.649	5.288	-4.318	55,696	. 028632	.32791
Recover players	1952-1961	543.94	.62931	29817	.899	1.911	5.194	-3.718	21,937	.024796	.65410

Data Sources: Economic Almanac, Statistical Abstracts of the U.S., Electrical Merchandising, and Electrical Merchandising Week.

in order to find F(T) we must solve this non-linear differential equation:

$$dT = dF/(p + (q - p)F - qF^2).$$

The solution is:

$$F = (q - pe^{-(T+G)(p+q)})/q(1 + e^{-(T+G)(p+q)}).$$

Since F(0) = 0, the integration constant may be evaluated:

$$-C = (1/(p+q)) \operatorname{Ln}(q/p)$$
 and $F(T) = (1 - e^{-(p+q)T})/(q/pe^{-(p+q)T} + 1)$

Then,

$$f(T) = ((p+q)^2/p)[e^{-(p+q)T}/(q/pe^{-(p+q)T}+1)^2],$$

and

$$S(T) = (m(p+q)^2/p[e^{-(p+q)T}/(q/pe^{-(p+q)T}+1)^2].$$

To find the time at which the sales rate reaches its peak, we differentiate S,

$$S' = (m/p(p+q)^3 e^{-(p+q)T} (q/p e^{-(p+q)T} - 1))/(q/p e^{-(p+q)T} + 1)^3$$

Thus, $T^* = -1/(p+q) \operatorname{Ln}(p/q) = 1/(p+q) \operatorname{Ln}(q/p)$ and if an interior maximum exists, q > p. The solution is depicted graphically in Figures 2 and 3. We note that $S(T^*) = (m(p+q)^2/4q \text{ and } Y(T^*) = \int_0^{T^*} S(t) dt = m(q-p)/2q$. Since for

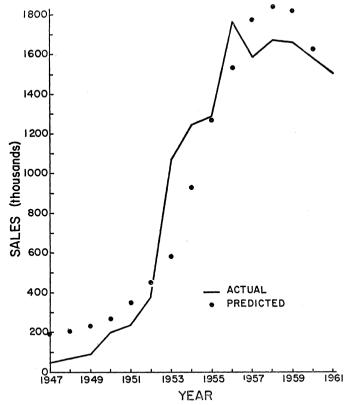


Fig. 4. Actual sales and sales predicted by regression equation

successful new products the coefficient of imitation will ordinarily be much larger than the coefficient of innovation, sales will attain its maximum value at about the time that cumulative sales is approximately one-half m. We note also that the expected time to purchase, E(T), is $1/q \operatorname{Ln}((p+q)/p)$.

The Discrete Analogue

The basic model is: $S(T) = pm + (q - p)Y(T) - q/mY^2(T)$. In estimating the parameters, p, q, and m from discrete time series data we use the following analogue: $S_T = a + bY_{T-1} + cY_{T-1}^2$. T = 2, $3 \cdots$ where: $S_T =$ sales at T, and $Y_{T-1} = \sum_{t=1}^{T-1} S_t =$ cumulative sales through period T - 1. Since a estimates pm, b estimates q - p, and c estimates -q/m: -mc = q, a/m = p.

Then q - p = -mc - a/m = b, and $cm^2 + bm + a = 0$, or $m = (-b \pm \sqrt{b^2 - 4ca})/2c$, and the parameters p, q, and m are identified. If we write $S(Y_{T-1})$ and differentiate with respect to Y_{T-1} , $dS_T/d_{Y_{T-1}} = b + 2cY_{T-1}$. Setting this equal to $S_{T-1} = -b/2c = m(q - p)/2q = S(T^*)$, and $S_T(Y_{T-1}^*) = a - b^2/2c + b^2/4c = m(p + q)^2/4q = S(T^*)$. Therefore, the maximum value of S_T as a function of time coincides with the maximum value of S_T as a function of cumulative sales.

Regression Analysis

In order to test the model, regression estimates of the parameters were developed using annual time series data for eleven different consumer durables. The period of analysis was restricted in every, case to include only those intervals in which repeat

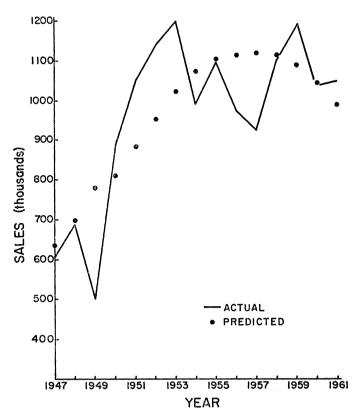


Fig. 5. Actual sales and sales predicted by regression equation (home freezers)

purchasing was not a factor of importance. These intervals were determined on the basis of a subjective appraisal of the durability of the product as well as from limited published data concerning "scrappage rates" and repurchase cycles.

Table 1 displays the regression results. The data appear to be in good agreement with the model. The \mathbb{R}^2 values indicate that the model describes the growth rate behavior rather well. Furthermore, the parameter estimates seem reasonable for the model. The regression estimates for the parameter c are negative in every case, as required in order for the model to make sense, and the estimates of m are quite plausible. One of the more important contributions derived from the regression analysis is the implied estimate of the total number of initial purchases to be made over the life of the product. Figures 4, 5, and 6 show the actual values of sales and the values predicted by the regression equation for three of the products analyzed. For every product studied the regression equation describes the general trend of the time path of growth very well. In addition, the regression equation provides a very good fit with respect to both the magnitude and the timing of the peaks for all of the products. Deviations from trend are largely explainable in terms of short-term income variations. This is especially apparent in Figure 5, where it is easy to identify recessions and booms in the years of sharp deviations from trend.

Model Performance

The performance of the regression equation relative to actual sales is a relatively weak test of the model's performance since it amounts to an ex post comparison of the

regression equation estimates with the data. A much stronger test is the performance of the basic model with time as the variable and controlling parameter values as determined from the regression estimates. Table 2 provides a comparison of the model's prediction of time of peak and magnitude of peak for the eleven products studied.

Since, according to the model S(0) = pm, we identify time period 1 as that period

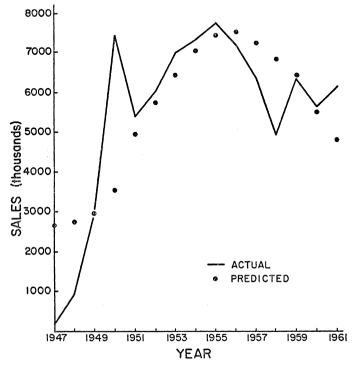


Fig. 6. Actual sales and sales predicted by regression equation (black & white television)

TABLE 2

Comparison of Predicted Time and Magnitude of Peak with Actual Values
for Eleven Consumer Durable Products

Product	q/Þ	Predicted time of peak $T^* = (1/(p + q))$ Ln (q/p)	Actual time of peak*	Predicted magnitude of peak $S(T^*) = m(p + q)^2/4q$ (10°)	Actual magni- tude of peak (10°)
Electric refrigerators	82.4	20.1	†	2.20	†
Home freezers	9.4	11.6	13	1.2	1.2
Black & white television	9.0	7.8	7	7.5	7.8
Water softeners	16.7	8.9	9	.5	.5
Room air conditioners	40.2	8.6	7	1.8	1.7
Clothes dryers	20.7	8.1	7	1.5	1.5
Power lawnmowers	36.7	10.3	11	4.0	4.2
Electric bed coverings	41.6	14.9	14	4.8	4.5
Automatic coffee makers	18.1	9.0	10	4.8	4.9
Steam irons	11.4	6.8	7	5.5	5.9
Recover players	26.3	4.8	5	3.8	3.7

^{*} Time period one is defined as that period for which sales equal or exceed p m for the first time.

[†] Interrupted by war. Prewar peak in year 16 (1940) at 2.6×10^6 units.

Product	Period of forecast	72	
Electric refrigerators	1926–1940	.762	
Home freezers	1947–1961	.473	
Black & white television	1949–1961	.077*	
Water softeners	1950–1961	.920	
Room air conditioners	1950–1961	.900	
Clothes dryers	1950–1961	.858	
Power lawnmowers	1949–1961	.898	
Electric bed coverings	1950–1961	.934	
Automatic coffee makers	1951–1961	.690	
Steam irons	1950–1961	.730	
Recover players	1953–1958	.953	
	1		

TABLE 3
Forecasting Accuracy of the Model for Eleven Consumer Durable Products

in which sales equal or exceed pm for the first time. It is clear from the comparison shown in Table 2 that the model provides good predictions of the timing and magnitude of the peaks for all eleven products studied.

In order to determine the accuracy with which it would have been possible to "forecast" period sales over a long-range interval with prior knowledge of the parameter values, the regression estimates of the parameters were substituted in the basic model, $S(T) = (m(p+q)^2/p)[e^{-(p+q)T}/(q/pe^{-(p+q)T}+1)^2]$, and sales estimates generated for each of the products for each year indicated in the intervals shown in Table 3. In most cases the model provides a good fit to the data. Even in the few instances of low r^2 values, the model provides a good description of the general trend of the sales curve, the deviations from trend being sharp, but ephemeral. Figures 7, 8, and 9 illustrate the predicted and actual sales curves for three of the products.

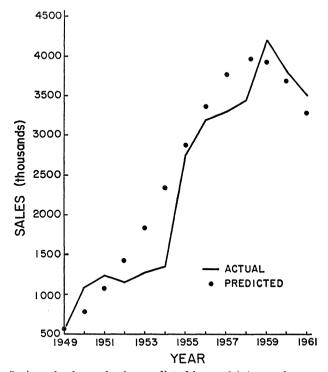
It would appear fair to conclude that the data are in generally good agreement with the model. The model has, then, in some sense, been "tested" and verified. We may now claim to know something about the phenomenon we set out to explore. The question is, however, will this knowledge be useful for purposes of long-range forecasting?

Long-Range Forecasting

There are two cases worth considering in long-range forecasting: the no-data case and the limited-data case. For either of these possibilities one may well ask: is it easier to guess the sales curve for the new product or easier to guess the parameters of the model? No attempt will be made here to answer this question, in general, but it does seem likely that for some products it would be possible to make plausible guesses of the parameters. Analysis of the potential market and the buying motives should make it possible to guess at m, the size of the market, and of the relative values of p and q, the latter guess being determined by a consideration of buying motives. If the sales curve is to be determined by means other than the model suggested in this paper, the implications of this forecast in terms of the parameters of the model might be useful as a test of the credibility of the forecast.

In order to illustrate the forecasting possibilities in the limited data case, we shall develop a forecast for color television set sales. In principle, since there are three

^{*} The low "explained" variance for this product is accounted for by extreme deviation from trend in two periods. Actually, the model provides a fairly good description of the growth rate, as indicated in Figure 9.



Frg. 7. Actual sales and sales predicted by model (power lawnmowers)

parameters to be estimated, some kind of estimate is possible with only three observations of the first of these observations occurs at T=0. Any such estimate should be viewed with some skepticism, however, since the parameter estimates are very sensitive to small variations in the three observations. Before applying estimates obtained from a limited number of observations, the plausibility of these estimates should be closely scrutinized.

In substituting $\sum_{t=0}^{T-1} S_t$ in the discrete analogue for $\int_0^T S(t) dt$ in the continuous model, a certain bias was introduced. This bias is mitigated when there are several observations, but can be crucial when there are only a few. Thus, the proper formulation of the discrete model, if $S_T = S(T)$ is: $S_T = a + bk(T)Y_{T-1} + ck^2(T)Y_{T-1}^2$, where $k(T) = Y(T)/Y_{T-1}$. We note that for any probability distribution for which: a) f(x) = 1/k[F(x+1) - F(x)], and b) F(0) = 0, $\sum_{t=0}^{x-1} f(t) = 1/k F(x)$. In particular, these two properties hold for the exponential distribution. Therefore, for this distribution $\sum_{t=0}^{x-1} F(x)/f(t) = k$. The density function f(T) in the growth model developed in this paper is approximately exponential in character when p and T are small. Thus, $f_{apx}(T) = 1/k[F_{apx}(T+1) - F_{apx}(T)]$ and $1/k = (p+q)/[e^{(p+q)} - 1]$. For small values of T we therefore write: $S_T = a + b'Y_{T-1} + c'Y_{T-1}^2$, where b' = kb, and $c' = k^2c$. Then m = k m', q = 1/kq', and p = 1/kp'. The value of 1/k for each of several different values of p + q has been calculated and appears in Table 4.

On the basis of the relationship between k and (p+q) indicated in Table 4: 1/k = .97 - .4(p+q), $(p'/q')q = q = 1/kq' = .97q'/(1 + .4(1+4\theta)q')$, where $\theta = q'/p' = q/p$, and $p = (p'/q')q' = .97p'/(1 + .4(1+\theta)p')$.

We turn now to the forecast of color television set sales. The following data are

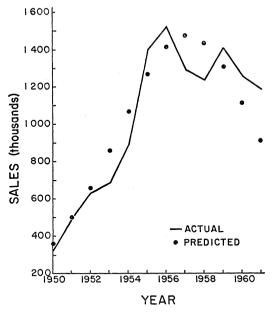


Fig. 8. Actual sales and sales predicted by model (clothes dryers)

available:

Sales (Millions of Units)	Year
.7	1963
1.35	1964
2.50	1965

Solving the following system of equations:

$$S_0 = .7 = a$$

 $S_1 = 1.35 = a + .7b' + .49c'$
 $S_2 = 2.50 = a + 2.05b' + 4.20c'$, we find:
 $a' = .7, b' = .954, c' = -.0374,$
 $m' = 26.2, q' = .96, p' = .0267,$
 $q = .67, p = .018, m = 37.4.$

Since these parameter values appear plausible, they have been used in the basic model to generate the series of estimates of sales shown in Table 5 and Figure 10. The projected peak occurs in 1968 at around 7 million units. This forecast differs somewhat from some industry forecasts. At this writing, one company's research department has estimated that sales will "top out" in 1967 at between 7 and 8 million units. The forecast speaks for itself and the ultimate reality of actual sales and one's personal criterion of "goodness" will determine whether or not the forecast was a good one.

The preceding forecast was made in late 1966. The following report was published in the May 19, 1967 issue of the Wall Street Journal:

^{...} Industry sources say most color set producers are continuing to trim production. "Only R.C.A., Zenith Radio Corp. and Philco-Ford Corp. aren't cutting back now, and they're taking a big gamble by banking on a big sales pickup this fall," asserts one industry analyst.

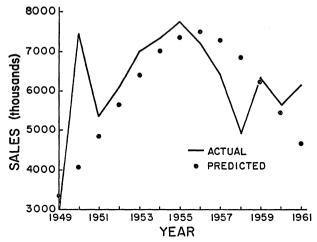


Fig. 9. Actual sales and sales predicted by model (black & white television)

TABLE 4
Calculated Values of 1/k and (p+q)

(p+q)	$1/k = (p+q)/[e^{(p+q)}-1]$		
.3	.85		
.4	.81		
.5	.77		
.6	.73		
.7	.69		
.8	.65		
.9	.61		

TABLE 5
Forecast of Color Television Sales 1966-1970

Year	Sales (millions)
1964	1.35
1965	2.5
1966	4.1
1967	5.8
1968	6.7
1969	6.3
1970	4.7

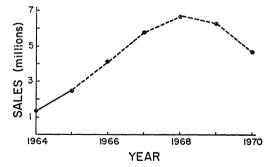


Fig. 10. Projected sales—color television

Executives at R.C.A., the biggest producer of color sets, have turned reticent about publicly forecasting industry sales for 1967. Their silence contrasts markedly with last fall, when R.C.A. and other companies were exuberantly predicting industry sales of 7 million color sets in 1967.

Admiral recently pared it's 1967 forecast to 6.1 million color-set sales for the industry. . . .

If the model developed in this paper does nothing else, it does demonstrate vividly the slowing down of growth rates as sales near the peak. In focusing upon the vital theoretical issues the model may serve to aid management in avoiding some of the more obviously absurd forecasts as have been made in the past.

While our forecast of color television sales was objectively determined in the sense that it was derived from data, it is also based upon a subjective judgment of the plausibility of the parameters. Since the parameter estimates are very sensitive to small variations in the observations when there are only a few observations, the importance of the plausibility test cannot be overemphasized. The parameter for which one has the strongest intuitive feeling is m. The plausibility test for this parameter is therefore perhaps the most easily derived. An examination of the parameter values implied by early sales of the other products analyzed in this paper indicates that in some instances these are remarkably close to the regression estimates, while in others they differ markedly. In every instance in which the early year estimates differ substantially from the regression estimates, these estimates are easily rejected on the basis of the implausibility of the implied value of m.

Conclusion

The growth model developed in this paper for the timing of initial purchase of new products is based upon an assumption that the probability of purchase at any time is related linearly to the number of previous buyers. There is a behavioral rationale for this assumption. The model implies exponential growth of initial purchases to a peak and then exponential decay. In this respect it differs from other new product growth models.

Data for consumer durables are in good agreement with the model. Parameter estimates derived from regression analysis when used in conjunction with the model provide good descriptions of the growth of sales. From a planning viewpoint, probably the central interest in long-range forecasting lies in predictions of the timing and magnitude of the sales peak. The model provides good predictions of both of these variables for the products to which it has been applied. Insofar as the model contributes to an understanding of the process of new product adoption, the model may be useful in providing a rationale for long-range forecasting.

References

- Bain, A. D., The Growth of Television Ownership in the United Kingdom, A Lognormal Model, Cambridge, The University Press, 1964.
- 2. Bartlett, M. S., Stochastic Population Models in Ecology and Epidemiology, London, Methuen, 1960.
- 3. Bush, R. R. and Mosteller, F., Stochastic Models for Learning, New York, Wiley, 1955.
- 4. Dernburg, T. F., "Consumer Response to Innovation: Television," in Studies in Household Economic Behavior, Yale Studies in Economics, Vol. 9, New Haven, Connecticut: Yale, 1958.
- FOURT, L. A. AND WOODLOCK, J. W., "Early Prediction of Market Success for New Grocery Products," Journal of Marketing, No. 2, Vol. 26, October, 1960.
- HAINES, G. H., JR., "A Theory of Market Behavior After Innovation," Management Science, No. 4, Vol. 10, July, 1964.
- 7. KATZ, E. AND LAZARSFELD, F., Personal Influence, New York: The Free Press, 1955.

- King, C. W., "Adoption and Diffusion Research in Marketing: An Overview," in Science, Technology and Marketing, 1966 Fall Conference of the American Marketing Association, Chicago, 1966.
- 9. Mansfield, E., "Technological Change and the Rate of Imitation," *Econometrica*, No. 4, Vol. 29. October, 1961.
- 10. Massy, W. F., "Innovation and Market Penetration," Ph.D. Thesis, Massachusetts Institute of Technology, 1960, Cambridge, Massachusetts.
- 11. Pessemier, E. A., New Product Decisions, An Analytic Approach, New York: McGraw-Hill, 1966.
- 12. RASHEVSKY, N., Mathematical Biology of Social Behavior, Chicago: The University Press, 1959.
- 13. Rogers, E. M., Diffusion of Innovations, New York: The Free Press, 1962.