A Diffusion Theory Model of Adoption and Substitution for Successive Generations of High-Technology Products

Author(s): John A. Norton and Frank M. Bass

Source: *Management Science*, Sep., 1987, Vol. 33, No. 9 (Sep., 1987), pp. 1069-1086

Published by: INFORMS

Stable URL: https://www.jstor.org/stable/2631875

MANAGEMENT SCIENCE
Vol. 33, No. 9, September 1987
*Printed in U.S.A.*

# A DIFFUSION THEORY MODEL OF ADOPTION AND SUBSTITUTION FOR SUCCESSIVE GENERATIONS OF HIGH-TECHNOLOGY PRODUCTS*

JOHN A. NORTON AND FRANK M. BASS
*Colgate Darden Graduate School of Business Administration, The University of Virginia, Charlottesville, Virginia* 22906-6550
*School of Management, The University of Texas at Dallas, Richardson, Texas*

This study deals with the dynamic sales behavior of successive generations of high-technology products. New technologies diffuse through a population of potential buyers over time. Therefore, diffusion theory models are related to this demand growth. Furthermore, successive generations of a technology compete with earlier ones, and that behavior is the subject of models of technological substitution. Building upon the Bass (1969) diffusion model, we develop a model which encompasses both diffusion and substitution. We demonstrate the forecasting properties of the model by estimating parameters over part of the data and projecting shipments for later periods.
(MARKETING; NEW PRODUCTS)

## Introduction

Newer technologies are continually replacing older ones. In marine power, steam replaced sail and, in turn, was replaced by the internal combustion engine. As a source of industrial fuel, oil substituted for coal, and natural gas and nuclear energy have substituted for oil. In recent years, the time interval between successive generations of high-technology electronic products has been demonstrated to be relatively brief in comparison with the time interval between replacing technologies using historical norms. As the time interval between technologies decreases the importance of understanding the impact of recent technologies on earlier ones increases. No matter what their advantages, newer technologies are not adopted by all potential buyers immediately. Rather, a diffusion process is set into motion. The newer technology may widen the market by allowing applications which were not feasible before. It will also provide an opportunity for buyers of earlier technologies to substitute the more recent technology for earlier ones. These substitution effects will ultimately diminish the potential, if not the actual sales, of earlier technologies, in the following ways. First, customers who would otherwise have adopted the earlier device will, instead, adopt the later one. Second, customers who have already adopted the first device may switch from (disadopt) the earlier device in favor of the later one. When the time interval between technologies is short the earlier technology may continue to diffuse through a population of potential buyers even as the substitution process is under way. Therefore, the demand for an earlier technology may continue to grow even as the substitution process occurs, as illustrated in Figure 1. We develop a model which explains the conjunction of diffusion and substitution. In this way we provide a basis for assessing and forecasting the influence of recent technologies on earlier ones. We study here certain electronic devices, but other work underway indicates that the model and concepts apply more widely.

There is a substantial literature on the diffusion of innovations (that is, literature dealing with the growth in level of demand, acceptance, or use of something as a

* Accepted by John R. Hauser; received February 25, 1986. This paper has been with the authors 3 months for 2 revisions.



0025-1909/87/3309/1069$01.25

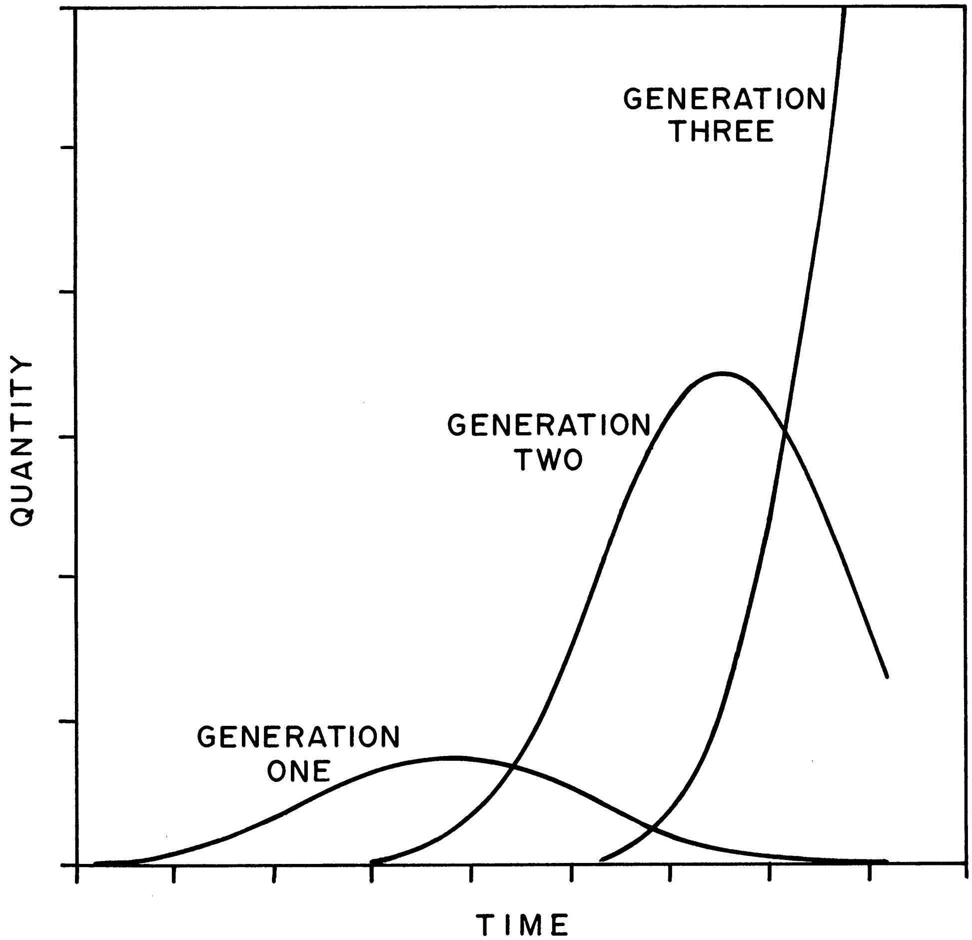


FIGURE 1. A Series of Technological Generations.

function of time and other variables). There is also a considerable literature dealing with technological substitution. Most substitution models are market share models; indeed, a principal distinction between substitution and diffusion models is that models of technological substitution assume there is a market there to be substituted, and, many times, the size of that market is known, whereas diffusion models make no such assumptions, in general. In fact, it is often the case that one of the principal reasons a diffusion model might be of interest in the first place is its ability to estimate or forecast a market potential.

## The Literature

### *Diffusion Models*

In the model to be developed for demand growth and decline of successive generations of technological innovations we shall make use of diffusion effects as well as substitution effects. In this section we review briefly some of the more popular diffusion models.

The Bass (1969) model is one of the more well-known and widely used models of first-purchase demand. It is a model of the timing of adoption of an innovation and will be central to subsequent developments.

The Bass model has a behavioral rationale that is consistent with studies in the social science literature on the adoption and diffusion of innovations (e.g., Rogers 1983) and is based on a simple premise about the hazard function (the conditional probability that an adoption will occur at time $t$ given that an adoption has not yet occurred). Thus if $f(t)$ is defined as the probability of adoption at time $t$, or the fraction of the ultimate potential which adopts the innovation at time $t$ and $F(t)$ is the fraction of the ultimate potential which has adopted by time $t$, the fundamental premise is that the likelihood of adoption at time $t$ given that one has not yet occurred is:

$$f(t)/[1 - F(t)] = p + qF(t). \tag{1}$$

The parameter $p$ is called the coefficient of innovation and $q$ the coefficient of imitation. Adoptions of the product are made by "innovators" and "imitators." The importance of innovators will be greater at first but will diminish monotonically with time, while the imitation effect will increase with time. Equation (1) leads to the differential equation:

$$f(t) = p + (q - p)F(t) - q[F(t)]^2. \tag{2}$$

If $F(0) = 0$ the solution to equation (2) is:

$$F(t) = [1 - \exp(-bt)]/[1 + a\exp(-bt)], \tag{3}$$

and the density function of time to adoption will be:

$$f(t) = (b^2/p)\exp(-bt)/[1 + a\exp(-bt)]^2, \tag{4}$$

where $a = q/p$ and $b = p + q$. It is easy to differentiate $f$ to find the time of the peak in $f$ or the inflection point of $F$ to be $t^* = (1/b)\ln(a)$.

If the innovation is a consumer durable such as are television sets, power lawn mowers, room air conditioning units, or cellular mobile telephones, then sales during the period in which demand consists of initial purchases of the product will be proportional to $f$. On the other hand, in other circumstances when sales consist of repeat purchases or when adopters make multiple purchases of the innovation it will be necessary to measure the adoption rate separately from the sales measure.

Other distributions used in connection with time to first purchase include the negative exponential, the logistic, and the Weibull. (For applications of these, see Fourt and Woodlock 1960, Mansfield 1961, and deKluyver 1982.)[1] Bass' original extended logistic formulation initiated a number of modifications and extensions. (For examples, see Bass 1980; Dodson and Muller 1978; Horsky and Simon 1983; Kalish 1985; Lekvall and Wahlbin 1973; Mahajan and Peterson 1978 and 1979; Mahajan, Peterson, Jain and Malhotra 1979. For reviews of a number of diffusion models, see Kotler 1971; Lilien and Kotler 1983; and Mahajan and Muller 1979.) While most of these refinements add something to the paradigm, none overrides the essential character of the model as originally formulated. Further, as Jeuland (1981) points out, given the nature of the data available for most new product forecasting tasks (approximate) and the quantity of information available (scant), there is considerable difficulty in reliably

[1] Because we anticipated the question of whether one density function is just as good as another in the model formulation, we did investigate the model formulated with these alternative distributions. We found that the negative exponential distribution was clearly inappropriate, which makes sense when one considers that it implies a memoryless process. We further found that the logistic and Weibull formulations did almost as well as that of Bass in terms of fit to the data over many observations. There were, however, considerable differences in the forecasts resulting from these distributions when few data points were used to parameterize the model. The resulting conclusion is that the function characterized as equation (3) above was uniformly superior for forecasting, especially when few observations are available. The results are not published here in the interest of brevity, but are available from the first author upon request.

extrapolating the early sales path of a new product. He argues that "if improved forecasting models are to be built, it must be within the constraint of highly parsimonious models" (Jeuland 1981, p. 3). From the point of view of developing a first model incorporating the effects of diffusion and substitution, the original formulation is conceptually sound and captures the essence of the process.

*Substitution Models*

An important and often-used model of technological substitution was proposed by Fisher and Pry (1971). Their model is based upon three assumptions. The first of these is that "many technological advances can be considered as competitive substitutions of one method of satisfying a need for another." Second, they observed that new technologies often completely supplant older ones. Third, they expressed their market share model in terms of Pearl's Law: "The fractional rate of fractional substitution of new for old is proportional to the remaining amount of the old left to be substituted." They assert that the rate constant of a substitution, once begun, does not change. Rather, it is the number and magnitude of such substitutions that measure more properly the pace of technological change in society. Their model is:

$$d/dt[s(t)] = ks(t)[1 - s(t)], \tag{5}$$

where $s(t)$ is the fractional market share of the innovation at time $t$, and $k$ is a constant of proportionality. The time scale $t$ may be chosen such that $s(0) = \frac{1}{2}$. When this is done it is possible to solve equation (5) to yield the following expression for the fractional market share of the newest innovation:

$$s(t) = 1/[1 + \exp(-kt)]. \tag{6}$$

Equation (6) is a form of logistic function. Using the assumption that there are only two competing technologies, Fisher and Pry derive a more convenient form for purposes of estimation. The result is that the log of the ratio of the market share of the succeeding technology to that of the first is a linear function of time:

$$\ln [s/(1 - s)] = kt. \tag{7}$$

(In their notation $k = 2a$ and $t = t - t_0$.)

The Fisher-Pry model has been shown to fit the data well for a number of successive innovations. Peterka (1977) extended their model to apply to several generations. His formulation was:

$$\ln (s_j/s_i) = kt, \tag{8}$$

where $i$ and $j$ are technologies, $j$ newer than $i$.

There have been very few studies in the technological substitution literature of a series of technological substitutions. The most intense and interesting competition may indeed be taking place between the two newest technologies, but there are a number of examples of simultaneous competition of more than two generations. (But see, for example, Peterka 1977 and Sharif and Kabir 1976.) As each innovation is studied in finer detail, it is often clear that the process is evolutionary, not revolutionary. The process of multilevel substitution is central to the development of the model we propose.

Blackman (1971) modified the Mansfield (1961) model and suggested the model:

$$\ln [s/(S - s)] = a + bt \tag{9}$$

where $S$ is the upper limit on market share obtainable by the new technology, $s$ = the market share of the new technology at time $t$, $a$ is a constant, and $b$ is a linear function of investment and profitability. A number of other variations of the Fisher-Pry share substitution models exist, such as those by Floyd (1968) and Sharif and Kabir (1976). The existing substitution models either deal in share or in number of adopting firms as opposed to sales. Each of them has a dynamic time component thus implying delayed learning about the advantage of the innovations.

The Fisher-Pry model is particularly nice in that it describes in very simple terms a very powerful generalization. However, it does not address the level of sales for each generation. And, as we will observe later, market shares exhibit a much nicer regularity than do absolute sales levels. In certain environments, managers require insight into the development of a market in absolute as well as relative terms. Clearly, if we can successfully forecast in absolute terms, we can use that knowledge to derive the relative or market share information.

The model we develop deals with sales and it explicitly incorporates diffusion effects and substitution effects. Time-varying factors are subsumed in the model and not explicitly taken into account.

## The Model

The model development will be facilitated by introducing some simple definitions. We will use the word "device" to refer to a component part of a larger product. We will use the word "application" to refer to a use of the device in a particular product, and "adoption" to refer to the incorporation of the device into the design of the product. Applications are things, not people, but it may be useful to think of applications as an adopting population. More properly, design engineers decide to incorporate a certain level of integrated circuit (IC) content (say) into a particular product. That IC content presumably makes the product better in some way, and becomes a standard part of the product's design. The product has "adopted" the use of ICs.

Let us make a few assumptions: First, we assume that once an application incorporates some level of the new technology, it does not revert to earlier technologies during the relevant period of time.

In the case where adopters buy multiple units of a new innovation, sales will consist of the number of units purchased per user times the number of users. Let us assume that the average rate of consumption per time period approaches a constant. More specifically, let the usage rates associated with applications be independently and identically distributed with a constant overall mean rate. (The argument for the constancy of this rate is based on the fact that, as the number of consuming units gets large, the change in the average amount consumed will be little affected by the inclusion of an incremental user.)

Based on several phenomena, we can assume that the number of uses of a given device changes over time. The device may have been invented to go into a particular product or solve a particular problem. It may also have uses which are not imagined until it exists. Finally, a device may perform better than something currently available. It is readily observable that the uses of some generations of innovations are growing in number, some explosively. Nonetheless, however rapid its growth appears, each device must have some upper limit on its applicability, and we assume that limit to be constant. We assume also that the rate of growth in applications incorporating a given device is related to the number of applications already existing and the number of potential applications yet to be made.

We recognize the simplification introduced by the second of the above assumptions. The assumption of the constancy of average per-period consumption implies that the growth of applications constitutes all of the variability in the process over time. Since in many situations we can observe neither the average rate of consumption of devices nor the number of applications but only their product, this may be a reasonable and workable simplification.

*Model Development*

Denote by $M$ the upper limit of the number of applications for which the innovation is appropriate. Denote by $r$ the rate at which an average application consumes the output of interest. These two variables, assumed constant, multiply together to yield the upper limit on sales per time period, notationally designated by $m$. The only thing changing, as discussed earlier, is the number of applications which adopt.

The process of incorporating a new technology into a given product is a process which involves the diffusion of knowledge about the characteristics of the technology. The specific functional form chosen to represent that process is that proposed by Bass and written above as equation (3). Thus in the case of a single generation with no successor, sales would be written:

$$S(t) = mF(t). \tag{10}$$

Sales would be proportional to the ($S$-shaped) cumulative distribution function of the adoption rate. First differences in sales would be proportional to the adoption rate [approximately the (unimodal) density function of time to adoption].

We deal in this model with a series of generations of innovations. Each generation is introduced to the market before its predecessor has been fully diffused to its potential population. Each successive generation will obtain sales by: (a) expanding applications, thus obtaining sales that would not have otherwise gone to earlier generations and (b) capturing sales that would otherwise have gone to earlier generations. Some part of the sales that would have otherwise gone to earlier generations will consist of customers switching from the earlier to the later generation. Another part will consist of customers who would have adopted the earlier device but instead adopt the later one. The two may or may not be distinguishable in practice; in the empirical section which follows, they are not.

Let us introduce some additional notation: Let $i$ index generations of a particular device type. Then we can denote by $S_i$ the shipments of the $i$th generation, and, in the case of two generations, we may write, for the first generation's sales:

$$S_1(t) = F_1(t)m_1 - F_2(t-\tau_2)F_1(t)m_1 = F_1(t)m_1[1 - F_2(t-\tau_2)] \qquad \text{for } t > 0 \tag{11}$$

and for the second,

$$S_2(t) = F_2(t-\tau_2)[m_2 + F_1(t)m_1] \qquad \text{for } t > \tau_2 \tag{12}$$

where $S_i(t)$ refers to sales of the $i$th generation in time period $t$, $m_1$ refers to the potential for the first generation, $m_2$ refers to the potential uniquely served by the second generation, and where $F_i = [1 - \exp-(b_it)]/[1 + a_i \exp-(b_it)]$. Indexing the generations with $i$, of course, means that $a_i = q_i/p_i$ and $b_i = p_i + q_i$. Note that $\tau_2$ is the time at which the second generation is introduced, and $F_2(t-\tau_2) = 0$ for $t < \tau_2$.

The simultaneous model indicated in equations (11) and (12) captures both adoption and substitution. $S_2(t-\tau_2)$ will increase monotonically after $\tau_2$. The peak in $S_1(t)$ will occur at or after $\tau_2$ depending on the relationship of $F_2$ to $F_1$. The ultimate level of sales of the second generation will be the sum of the potentials for both generations.[2]

The model may be applied to a number of generations simultaneously even though as the number of generations expands the sales of earlier generations will approach zero because of substitution. For three generations, generation one loses to two. Generation

[2] In theory, $m_i$ may be any nonnegative number; it is an incremental market for all $i > 1$. For example, in the case of near perfect substitutes, $m_2$ might be very nearly zero. For purposes of estimation, there is no constraint on any of the $m_i$. The use of an appropriate functional measure (e.g., tons of freight moved instead of number of steam or diesel locomotives) should avoid the consequence of a negative number for any of the $m_i$.

two gains from one, but loses to generation three, including the loss of actual and potential gained from one.

The model assumes (a) the existence of a series of advancing generations, each of which can do everything the previous generation could do, and possibly more, (b) a density function of time to adoption for each generation applying against a time-varying potential, and (c) the substitution of actual and potential sales from earlier to later generations.

If we relax the assumption that the adoption process differs by generations, we can write $p_i = p$ and $q_i = q$ for all $i$. This assumption has a certain plausibility in that for a given technology class the behavioral processes for the adoption of advancing generations could be expected to be similar. If the assumption is approximately correct, the number of parameters to be estimated will be drastically reduced. For example, if $F$ has $k$ parameters and there are $n$ generations, there will be $kn + n$ parameters to estimate, but if the $k$ parameters of $F$ are the same for each generation the number of parameters to estimate will be $k + n$. The very strong assumption that $p$ and $q$ are constant across equations should be easily falsifiable on the basis of model fit: If the $p_i$ and $q_i$ are really different for different $i$, it should be difficult to obtain reasonable fits to the empirical data and reasonable forecasts using the simpler model.

## The Data

The data employed in this study are from the semiconductor industry and were collected and supplied to us by Dataquest. We shall model the demand growth and decline for successive generations of two basic types of integrated circuits (ICs): memory and logic circuits. These devices are similar in that they are products of the highest technology, they are sold to industrial as opposed to consumer markets, and their markets are dynamic in the sense that technological advances influence their environment. They are also similar in that they are very important to the nations that manufacture and use them for geopolitical, economic, and military reasons. They are different in their applications, their served markets, and in that the technology which leads to improvement in one may not be applicable to the other. Descriptions of the devices are included in an appendix available from the TIMS office, in case the terms dynamic random access memory (DRAM), static random access memory (SRAM), microprocessor and microcontroller (MPU and MCU) are unfamiliar. The point is that they are sequential innovations. The successive generations studied are the 4k, 16k, 64k, and 256k DRAM, the 4k, 16k, and 64k SRAM, and eight-bit logic devices (MPU and MCU). The data comprise 44 quarterly observations of DRAM shipments from 1974 through 1984, 36 quarterly observations of SRAM shipments ending in mid-1984, and 32 quarterly observations of eight-bit logic devices, ending with the fourth quarter of 1983.

## Estimation and Fitting

### *DRAM Shipments*

Operating under the assumption of constant diffusion parameters $p$ and $q$ for all generations of DRAMs, we can write the following system of joint nonlinear equations to describe their growth and interplay:[3]

[3] Let $G_k = m_k \prod_{i=k}^{M} F_i$, $i, k = 1, 2, \ldots, n$, and $F_i = F(t - \tau_i) \equiv 0$ for $t < \tau_i$. Then in general,

$$S_n(t) = \sum_{k=1}^{M} G_k(1 - F_{n+1}). \tag{17}$$

$$S_1(t) = F(t)m_1[1 - F(t - \tau_2)], \tag{13}$$

$$S_2(t) = F(t - \tau_2)[m_2 + F(t)m_1][1 - F(t - \tau_3)], \tag{14}$$

$$S_3(t) = F(t - \tau_3)[m_3 + F(t - \tau_2)[m_2 + F(t)m_1]][1 - F(t - \tau_4)], \tag{15}$$

$$S_4(t) = F(t - \tau_4)[m_4 + F(t - \tau_3)[m_3 + F(t - \tau_2)[m_2 + F(t)m_1]]], \quad \text{where} \tag{16}$$

$S_i(t)$ = shipments of generation $i$,
$F(\cdot) = [1 - \exp(-b\cdot)]/[1 + a\exp(-b\cdot)]$ (where $a$ is $q/p$ and $b$ is $p + q$, as before), and
$m_i$ = the incremental potential served by the $i$th generation, that is, that not capable of being served by any generation $j < i$.

This set of equations requires the estimation of six parameters: the coefficient of innovation $p$, the coefficient of imitation $q$, and the incremental potentials $m_1$ through $m_4$. These equations have been estimated jointly using the nonlinear three-stage least squares[4] procedure SYSNLIN, available as part of the Statistical Analysis System (SAS). For details of the program, please refer to the documentation in SAS 1984, pp. 505–550. Table 1 shows the parameter estimates using DRAM data to estimate equations (13)–(16). All of the parameter estimates are positive as expected and highly statistically significant with small standard errors relative to the estimates. Moreover, the parameter estimates are plausible. The $m_i$ increase monotonically with $i$ as might be expected for DRAMs and the values of $p$ and $q$ are within the range of experience of parameter estimates for other innovations.

Table 2 shows the fit of the model to the data for DRAMs. The $R^2$ values are remarkably high, the lowest value exceeding 0.96. Any test of the hypothesis that the parameters $p$ and $q$ are equal across generations would reject the hypothesis only if the fit of the model to the data without the constraints that $p_i = p$ and $q_i = q$ were markedly superior to the fit with the constraint. Clearly, the data and the model are in such a high degree of correspondence that we may assert that the data do not reject the notion that the assumption of constant $p$ and $q$ for all generations is reasonable.[5]

Figure 2 shows the fit of the model to the data. Clearly, there is close correspondence between the model and data. A plot of fitted versus actual aggregate demand derived from these data (not reproduced here) similarly shows very close correspondence.

As previously indicated, when there is only one device and no substitution, the first differences in sales will be proportional to the adoption rate. Figure 3 shows the theoretical graph of first differences when substitution comes into play. At some point, first differences will turn negative and will remain so, but at a later point the rate of substitution will diminish as the higher technology begins to saturate its potential from substitution. Eventually the sales of the first generation will approach zero and there will be little potential left for the second to capture. Figure 4 shows, for the 4k DRAM, the fit of the estimated derivative of shipments for the devices to the actual first differences. Although the data are noisy, the trend supports the model predictions. The prediction of the upper peak of differences corresponds chronologically to the time of the inflection point of sales.[6] We note also that as soon as the next generation is

[4] Jain and Rao (1985) and Srinivasan and Mason (1985) demonstrate NLS estimation of the Bass diffusion model in its intrinsically nonlinear form. Our thanks to R. Rao for his helpful suggestions.

[5] In this case, systems of equations are estimated with intercepts constrained to be zero. Therefore, $R^2$ cannot be unambiguously interpreted as a percentage of explained variance. Moreover, because of the large number of zero values which are perfectly "predicted" by the estimation, the significance estimates are very likely to be overstated. Nevertheless, visual inspection of the fits indicates that they are very good.

[6] Given two competing devices, one the successor to the other, when will sales peak, and how high will that peak be? Rather than graph the equations, we can get an analytical answer, by taking the time derivative of equation (14) to derive an expression for $t^*$, the peak time, and $S_1(t^*)$, the peak magnitude, in terms of the estimates of $p$, $q$, and $m_1$. See the appendix (available from the TIMS office) for details.

TABLE 1

*Parameter Estimates for Equations* (13) *Through* (16)

| Parameter | Estimate | Approx. Std. Error | $t$ Ratio | Approx. Prob. $>\|t\|$ |
|---|---|---|---|---|
| $m_1$ | 22523.24 | 654.24 | 34.43 | 0.0001 |
| $m_2$ | 59789.50 | 1778.39 | 33.62 | 0.0001 |
| $m_3$ | 338834 | 9796.27 | 34.59 | 0.0001 |
| $m_4$ | 762917 | 33995.63 | 22.44 | 0.0001 |
| $p$ | 0.00370603 | 0.00013041 | 28.42 | 0.0001 |
| $q$ | 0.33692 | 0.00710933 | 47.39 | 0.0001 |

introduced, the derivative of sales of the older generation starts to slow, clearly indicating the link between the two generations. Graphic representations of the first differences of later devices (again, not reproduced in the interest of brevity) showed similar correspondence.

*SRAM Shipments*

Applying the model to the market for three generations of SRAMs, we write:

$$S_1(t) = F(t)m_1[1 - F(t - \tau_2)], \tag{18}$$

$$S_2(t) = F(t - \tau_2)[m_2 + F(t)m_1][1 - F(t - \tau_3)], \tag{19}$$

$$S_3(t) = F(t - \tau_3)[m_3 + F(t - \tau_2)[m_2 + F(t)m_1]], \qquad \text{where} \tag{20}$$

$S_i(t)$ = shipments of generation $i$,

$F(\cdot) = [1 - \exp(-b\cdot)]/[1 + a\exp(-b\cdot)]$ (where $a$ is $q/p$ and $b$ is $p + q$, as before), and

$m_i$ = the incremental potential served by the $i$th generation, that is, that not capable of being served by any generation $j < i$.

The estimates of equations (18)–(20) are shown in Table 3. As is the case for DRAMs, the estimates have the expected sign and reasonable values, and are all statistically significant.

Table 4 shows the fit of the model to the data for SRAMs. Again, we find a failure of the data to reject the model, and in particular the constrained model.

*Logic Device Shipments*

Logic and memory differ in that an advanced generation logic device will not necessarily easily substitute for an earlier generation device. Our data indicated that the eight-bit microprocessor (MPU) did not take over from the four-bit device, nor did the 16-bit from the eight-bit device. There are applications for which accessing a larger word or operating at very high speed is irrelevant. Furthermore, the peripherals used to control an MPU cost much more, in general, than the MPU itself. As a generation of

TABLE 2

*Model Fit to Data*

| Endogenous Variable | $R$-square |
|---|---|
| $S_1(t)$ | 0.9672 |
| $S_2(t)$ | 0.9646 |
| $S_3(t)$ | 0.9993 |
| $S_4(t)$ | 0.9676 |

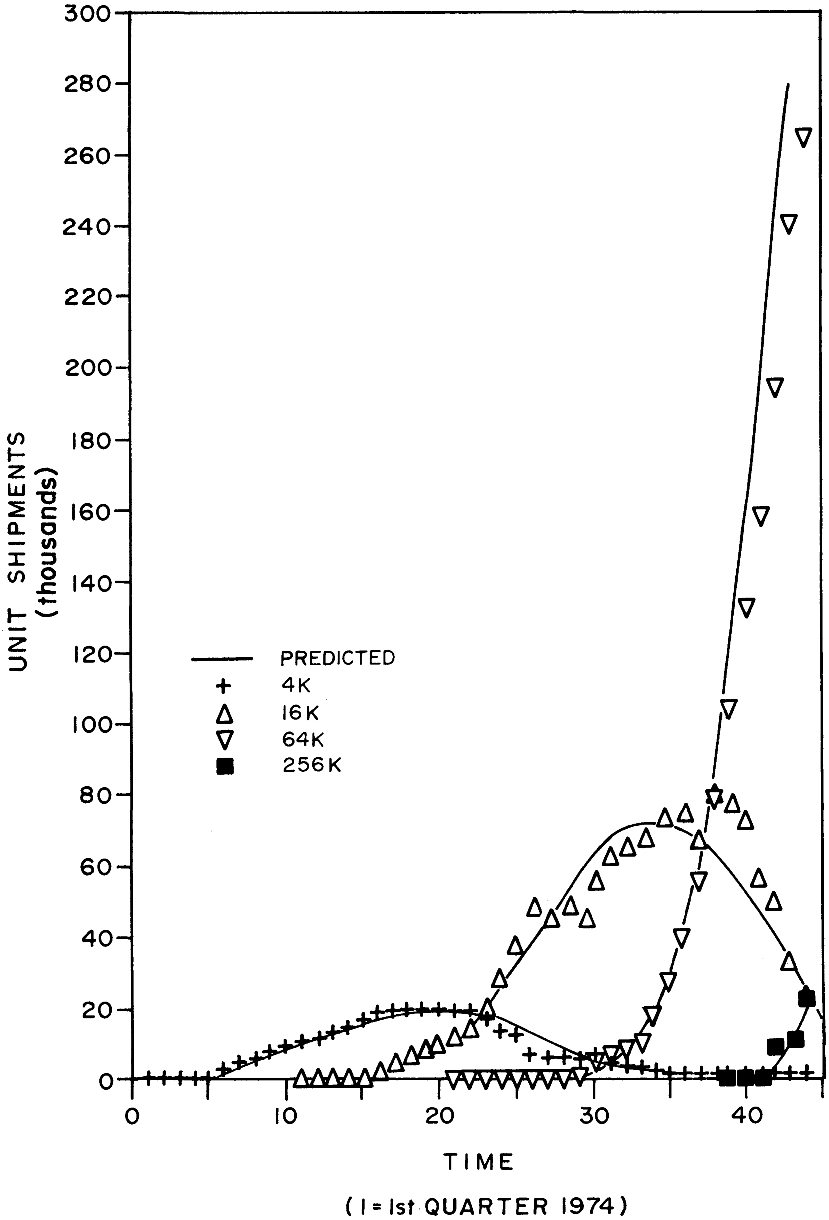


FIGURE 2. Model Fit to DRAM Data.

MPUs matures it will be integrated into more and more designs. With maturity it becomes expedient to put all the controlling devices, together with some instruction and a little memory, on a single chip. That chip is called a microcontroller (MCU). Therefore, one should expect that an MCU of a given word length would take over from an MPU of the same word length. We estimate the model of the relationship between both types of eight-bit logic, a set of equations similar to those presented earlier, with quarterly data during the eight-year period ending in 1983. Tables 5 and 6 show the parameter estimates and fit statistics. As was the case for both types of memory, the model fits the data extremely well. Eight-bit MPU sales are beginning to peak as MCUs take over.

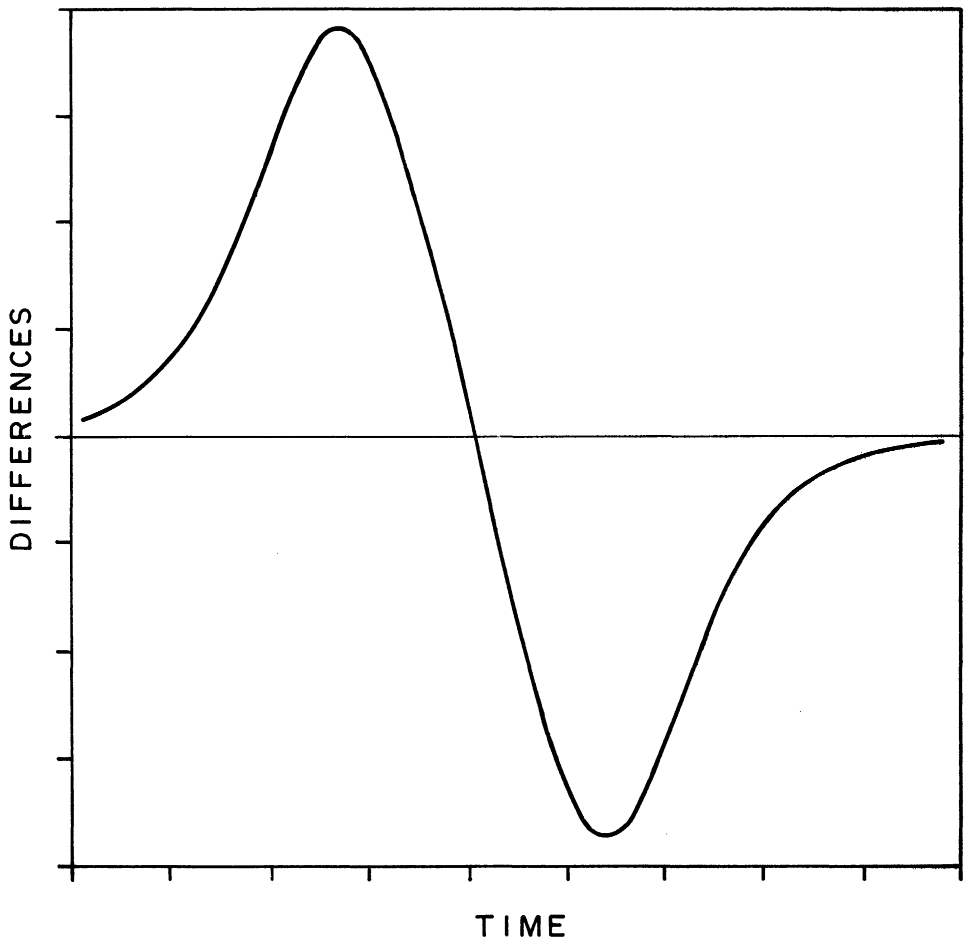


FIGURE 3. 1st Differences, with Substitution.

*Market Shares*

Because of many successful applications the Fisher-Pry model qualifies as a powerful generalization. Its limitations are that it is a share model only and does not account for sales levels of successive generations and that it considers only a single technological substitute at a time. The generalization of the model by Peterka (1977) permits the estimation of share relationships for a series of generations of a technology. A generalized form of the Fisher-Pry model is:

$$\ln [s_j(t)/s_{j-1}(t)] = a + bt. \tag{21}$$

We note that while sales or shipments may exhibit highly irregular patterns, the process of substitution in relative or market share terms shows remarkably regular patterns. With that in mind, we examined the performance of the Fisher-Pry model, comparing the results with market share projections based on our model, to see whether the regularity captured by the Fisher-Pry formulation resulted in better market share forecasts. Given the fits of our model to the shipment data, it may not be surprising that ours produced somewhat better results, measured in terms of mean absolute percentage deviation. Again, we omit the results because of constraints on length, but the results are available from the first author upon request.

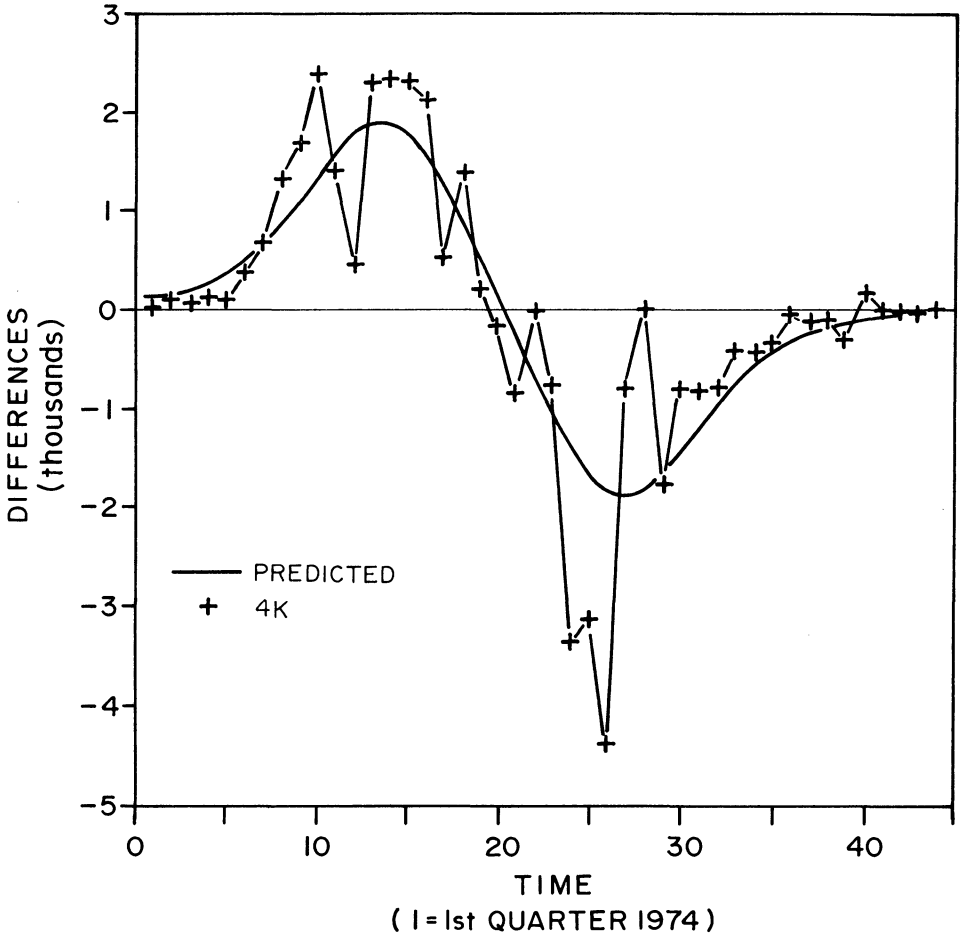


FIGURE 4. 1st Differences, 4k DRAM.

## Forecasting

In order to examine the forecasting properties of the model we have estimated the parameters using different data intervals and then projected shipments of the devices using these parameter estimates. In general, we estimate using data for $i$ devices ($i$ = 1, 2, 3, 4) up until the time when the $i$th-plus-one is introduced and then project shipments for all $i + 1$ devices.

Several studies of the sensitivity of the parameters of the Bass model indicate that estimates using data for first purchase adoption rates of single products are not very

TABLE 3

*Parameter Estimates for Equations* (18) *Through* (20)

| Parameter | Estimate | Approx. Std. Error | $t$ Ratio | Approx. Prob. $>\|t\|$ |
|---|---|---|---|---|
| $m_1$ | 32537.61 | 424.86 | 76.58 | 0.0001 |
| $m_2$ | 180055 | 6135.39 | 29.35 | 0.0001 |
| $m_3$ | 280094 | 19820.41 | 14.13 | 0.0001 |
| $p$ | 0.002459 | 0.00015964 | 15.40 | 0.0001 |
| $q$ | 0.23165 | 0.00548958 | 42.20 | 0.0001 |

TABLE 4
*Model Fit to SRAM Data*

| Endogenous Variable | $R$-square |
|---|---|
| $S_1(t)$ | 0.9911 |
| $S_2(t)$ | 0.9962 |
| $S_3(t)$ | 0.9811 |

TABLE 5
*Parameter Estimates: Logic*

| Parameter | Estimate | Approx. Std. Error | $t$ Ratio | Approx. Prob. $>\|t\|$ |
|---|---|---|---|---|
| $m_1$ | 79287.59 | 12687.67 | 6.25 | 0.0001 |
| $m_2$ | 54756.73 | 17962.66 | 3.05 | 0.0049 |
| $p$ | 0.00095211 | 0.00015277 | 6.23 | 0.0001 |
| $q$ | 0.13370 | 0.01094 | 12.22 | 0.0001 |

TABLE 6
*Model Fit to Logic Data*

| Endogenous Variable | $R$-square |
|---|---|
| $S_1(t)$ | 0.9916 |
| $S_2(t)$ | 0.9858 |

stable unless the period over which the estimates are formed extends past the peak of the curve. We expect something similar here. Fortunately, because there are a series of innovations, there will be a series of peaks and thus we expect the estimates to improve with the data interval, making it possible to predict with some accuracy more recent and future innovation sales and to assess the impact of these upon the sales of earlier generation innovations. In estimating in the way that we have we develop estimates of $p$, $q$, and $m_1$, $m_2$, . . . , $m_i$. Thus in order to project sales of the $i$th-plus-one it will be necessary to guess $m_{i+1}$.

*DRAMs*

The 16k DRAM first captured 2% of the market in period 14. We have estimated $p$, $q$, and $m_1$ with these 14 observations. In order to forecast sales we must guess $m_2$. We have taken $m_2$ to be equal to 4 times $m_1$, an arbitrary guess, but taken for the reason that the 16k device is four times as powerful (as dense) as the 4k. The resulting forecast values of shipments for 4k and 16k DRAMs captured the general shape of the shipment curves for the two devices, but the actual sales of the 4k device exceeded the forecast values, while the actual sales of the 16k device were below the forecast values. It should be emphasized that these forecasts were (a) developed from estimates made from data which do not extend to the peak in 4k shipments and (b) projected without the benefit of updating for 8 periods (two years). We expect the relaxation of either (a) or (b) above would improve the quality of the forecasts.

The 64k DRAM first captured 2% of the market in period 29. Estimates of $p$, $q$, $m_1$, and $m_2$, made on the basis of 29 observations, and $m_3$ (again taking as an estimated value for the new device's potential four times the previous potential), yield the results shown in Figure 5. As expected, the quality of the fit using data which include a turning point is superior to that without one. From a strategic viewpoint the projections in Figure 5 might be considered to be adequate inasmuch as they capture the rapid growth of the 64k device and the decline in the 16k device.

The 256k DRAM first captured two percent of the market in period 38. The fitted values and projections through period 43 for the four devices are only slightly different from those shown using Figure 5. Table 7 shows the parameter estimates for DRAMs for different data intervals. The estimates of $p$ diminish and estimates of $q$ increase with the number of observations used to obtain estimates. Systematic and monotonic changes in the $m_i$ are also observed. Insofar as changes of this type occur over different innovations, there is the suggestion that one could estimate parameters based on observations from the first one or two generations and modify these to accord with the expectation for estimates based on a larger number of observations. Table 8 shows that the changes in the estimates for SRAMs for different data intervals are of the same character as those for DRAMs. Although we do not include the graphs, the forecast quality for SRAMs was similar to that of DRAMs for data over similar intervals.

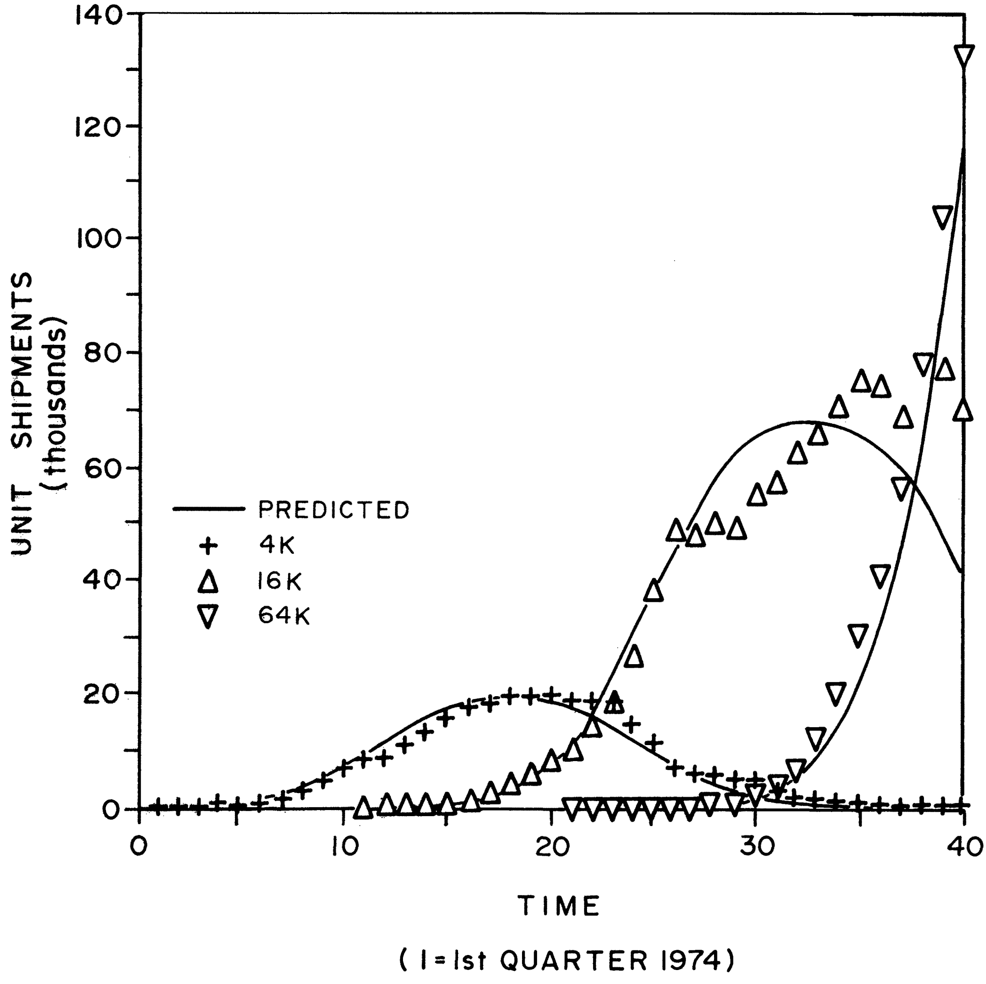


FIGURE 5. Forecast 4k–64k, Using 4k and 16k.

TABLE 7

*Parameter Estimates for DRAM Subsets*

| Observations | $m_1$ | $m_2$ | $m_3$ | $m_4$ | $p$ | $q$ |
|---|---|---|---|---|---|---|
| 14 | 15171 | | | | 0.0020 | 0.5065 |
| 29 | 22037 | 50499 | | | 0.0028 | 0.3938 |
| 40 | 22398 | 61230 | 390948 | | 0.0030 | 0.2521 |
| 44 | 22523 | 59790 | 338834 | 762917 | 0.0037 | 0.3369 |

TABLE 8

*Parameter Estimates for SRAM Subsets*

| Observations | $m_1$ | $m_2$ | $m_3$ | $p$ | $q$ |
|---|---|---|---|---|---|
| 21 | 21875 | | | 0.0013 | 0.3474 |
| 32 | 29659 | 206159 | | 0.0016 | 0.2764 |
| 36 | 32537 | 180055 | 280094 | 0.0025 | 0.2317 |

*Forecasts*

Since the principal use of our model is in forecasting and planning, it seems reasonable to supply some forecasts based on the data on hand thus far. The following forecasts were made in early 1985, using data available at the time. The data series for memory devices ended at the fourth quarter of 1984. Therefore, we forecast quarterly shipments of 16k, 64k, and 256k DRAMs for three years starting with the first quarter of 1985. The data series for logic devices ended with the fourth quarter of 1983, and therefore the logic forecasts are for three years starting with the first quarter of 1984.

Please note that all the following forecasts were developed in April of 1985, and are three year (12-step-ahead) forecasts, not one-step-ahead forecasts. None of the models is updated with information obtained since that time, with one exception: in June of 1985, we obtained from a planner in one of the semiconductor firms an estimate of the time when the 1M (one megabit) DRAM would capture two percent of the memory market. We used that guess to start the process of substitution for the earlier devices. (Only that introduction time matters, since $m_5$ would not enter the equation for the fourth generation's sales.)

Table 9 indicates the model forecasts of 16k, 64k, and 256k DRAMs beginning in the first quarter of 1985 and ending the fourth quarter of 1987. The model forecasts that,

TABLE 9

*Three-Year DRAM Forecast*

| Period | 16k | 64k | 256k |
|---|---|---|---|
| 45 | 19823 | 309257 | 34515 |
| 46 | 14946 | 327619 | 54283 |
| 47 | 11056 | 338055 | 81710 |
| 48 | 8058 | 340318 | 118931 |
| 49 | 5806 | 334375 | 168157 |
| 50 | 4149 | 320364 | 231237 |
| 51 | 2947 | 298716 | 308997 |
| 52 | 2084 | 270381 | 400467 |
| 53 | 1469 | 237012 | 502364 |
| 54 | 1033 | 200924 | 609238 |
| 55 | 725 | 164752 | 714489 |
| 56 | 509 | 130917 | 811907 |

TABLE 10

*Three-Year SRAM Forecast*

| Period | 4k | 16k | 64k |
|---|---|---|---|
| 37 | 22319 | 62739 | 7858 |
| 38 | 20701 | 73064 | 11120 |
| 39 | 18947 | 83737 | 15345 |
| 40 | 17102 | 94359 | 20774 |
| 41 | 15221 | 104463 | 27678 |
| 42 | 13359 | 113560 | 36355 |
| 43 | 11568 | 121176 | 47107 |
| 44 | 9890 | 126901 | 60222 |
| 45 | 8358 | 130421 | 75945 |
| 46 | 6988 | 131543 | 94438 |
| 47 | 5789 | 130210 | 115741 |
| 48 | 4758 | 126503 | 139732 |

regardless of the price of the devices, the 64k DRAM will decline in shipment volume permanently following the fourth quarter of 1985. Table 10 summarizes three-year forecasts for 4k, 16k, and 64k SRAMs.

Table 11 lists the estimates of shipments of eight-bit logic devices for the period from the first quarter of 1984 through the fourth quarter of 1986. We forecast a peak in eight-bit MPU shipments in the second quarter of 1985. Comparing the prediction with 1984 data which have since become available from Dataquest, sales appear to be flattening in accordance with the model predictions. This finding is interesting for a number of reasons, not the least of which is that a flattening of eight-bit MPU shipments was completely unexpected by a number of industry planners.

## Conclusions

The model put forth in this paper derives from prior work in the areas of diffusion and substitution, building upon knowledge gained in those studies. The unique contribution of the model is the casting of the process in such a way as to marry the two concepts to explain the process in a relatively accurate and yet parsimonious form. The model also has value in its ability to forecast overall industry demand as well as market shares of the various devices. Clearly, there are omitted variables, but from a practical forecasting viewpoint, the model serves in describing a generalization. To illustrate,

TABLE 11

*Three-Year Logic Forecast*

| Period | MPU | MCU |
|---|---|---|
| 33 | 18540 | 31437 |
| 34 | 19109 | 35203 |
| 35 | 19520 | 39235 |
| 36 | 19760 | 43507 |
| 37 | 19819 | 47990 |
| 38 | 19697 | 52642 |
| 39 | 19399 | 57419 |
| 40 | 18934 | 62271 |
| 41 | 18321 | 67146 |
| 42 | 17578 | 71992 |
| 43 | 16731 | 76759 |
| 44 | 15802 | 81403 |

consider the results of forecasting the 64k device using data from the 4k and 16k. A strategic planner using the model parameterized with data for 1974 through 1980 would not have been seriously wrong in forecasting global demand for each of the several generations for at least three years, even without updating.

*Validation*

By building upon the contributions of Mansfield, Fisher and Pry, and Bass, we can claim convergent validity. From the evidence supplied by the fit of the model to various data sets, we can assert event validity. We demonstrated predictive validity by using subsets of the data for parameterizing the model, then forecasting over remaining data points. We can assert hypothesis validity only to the extent that (a) the model does not reject the assumption of a constant substitution rate for a product class and (b) these data, at least, do not reject the assumption of the decline of sales to virtually zero for the earlier generations.

*Future Research*

The model is applied to data from the semiconductor industry. How would the model work in other applications? That is for future research to discover, but we offer some suggestions: First, the model needs to be applied to data wherein the distinction between the two types of succession can be discerned. That is, we need to be able to track separately those applications that choose the second generation (say) instead of the first and those that, having already adopted the first, switch to the second. Further, the model needs to be modified to allow switching by only a fraction of the market. For example, consider the substitution of acetaminophen for aspirin, and ibuprofin for both of those. The patient might need the anti-inflammatory qualities of aspirin and not take acetaminophen at any price.

The incorporation of price seems necessary, in some cases, particularly where the device cost is a large fraction of the total value of the product. This matter has been addressed in the case of diffusion models, but much remains to be done.

*Applications*

The finding that the parameters stabilize considerably once a single generation has reached a turning point is a potentially powerful generalization, one that suggests that the model is particularly well suited to deal with industries undergoing rapid change.

We suggest also that the model could be incorporated as a component of larger models, in the investigation of such research issues as optimal phasing out of "old" technological generations and/or optimal timing of the entry of new ones.

*Conclusion*

We have developed what we believe to be a unique and original model of demand growth and decline for a series of technological innovations. Our model is consistent with and builds upon well-known work on the diffusion of innovations. It incorporates both diffusion and substitution. The model works well for the technologies we have studied: In summary, it is simple, plausible, and historically accurate.

We believe that the model holds great promise for forecasting conditional on the timing of the introduction of successive generations of technology.[7]

[7] We are grateful for a grant from United Technologies Mostek to the Morris Hite Center for Product Development and Marketing Science at the University of Texas at Dallas which supported this research. Our thanks to R. L. Schultz, Director of the Hite Center, for arranging this support; to Chris Hann and Bert Kehren of United Technologies Mostek for information, advice, and counsel; and to Dataquest for supplying the data used in this study.

We gratefully acknowledge the thoughtful critical review by three anonymous referees. Any remaining errors and omissions are, of course, our responsibility.

## References


BASS, F. M., "A New-Product Growth Model for Consumer Durables," *Management Sci.*, 15 (January 1969), 215–227.

———, "The Relationship Between Diffusion Rates, Experience Curves, and Demand Elasticities for Consumer Durable Technological Innovations," *J. Business*, 53 (July 1980), S51–S67.

BLACKMAN, A. W., "The Market Dynamics of Technological Substitutions," *Technological Forecasting and Social Change*, 6 (February 1974), 41–63.

DEKLUYVER, C. A., "A Comparative Analysis of the Bass and Weibull New Product Growth Models for Consumer Durables," *New Zealand Oper. Res.*, 3, 2 (1982), 99–130.

DODSON, J. A. AND E. MULLER, "Models of New-Product Diffusion Through Advertising and Word-of-Mouth," *Management Sci.*, 24 (November 1978), 1568–1578.

FISHER, J. C. AND R. H. PRY, "A Simple Substitution Model of Technological Change," *Technological Forecasting and Social Change*, 3 (March 1971), 75–88.

FLOYD, A., "A Methodology for Trend Forecasting of Figures of Merit," in J. Bright (Ed.), *Technological Forecasting for Industry and Government: Methods and Applications*, Prentice-Hall, Englewood Cliffs, N.J., 1968, 95–109.

FOURT, L. A. AND J. W. WOODLOCK, "Early Prediction of Market Success for New Grocery Products," *J. Marketing*, 25 (October 1960), 31–38.

HORSKY, D. AND L. S. SIMON, "Advertising and the Diffusion of New Products," *Marketing Sci.*, 2 (1983), 1–17.

JAIN, D. AND R. RAO, "The Effect of Price on the Demand for Durables: Modeling, Estimation, and Findings," Paper presented at the Third Marketing Science Meeting, Nashville, TN.

JEULAND, A. P., "Parsimonious Models of Diffusion of Innovation Part A: Derivations and Comparisons," Working paper, Graduate School of Business, University of Chicago, 1981.

KALISH, S., "A New Product Adoption Model with Price, Advertising and Uncertainty," *Management Sci.*, 31 (December 1985), 1569–1585.

KOTLER, P., *Marketing Decision Making: A Model Building Approach*, Holt, Rinehart, and Winston, New York, 1971.

LEKVALL, P. AND C. WAHLBIN, "A Study of Some Assumptions Underlying Innovation Diffusion Functions," *Swedish J. Economics*, 75 (1973), 362–377.

LILIEN, G. AND P. KOTLER, *Marketing Decision Making: A Model-Building Approach*, Harper and Row, New York, 1971.

MAHAJAN, V. AND E. MULLER, "Innovation Diffusion and New Product Growth Models in Marketing," *J. Marketing*, 43 (October 1979), 55–68.

——— AND R. A. PETERSON, "Innovation Diffusion in a Dynamic Potential Adopter Population," *Management Sci.*, 24 (November 1978), 1589–1597.

——— AND ———, "Integrating Time and Space in Technological Substitution Models," *Technological Forecasting and Social Change*, 14 (1979), 231–241.

———, ———, A. K. JAIN AND N. MALHOTRA, "A New Product Growth Model with a Dynamic Market Potential," *Long Range Planning*, 12 (August 1979), 51–58.

MANSFIELD, E., "Technical Change and the Rate of Inflation," *Econometrica*, 29 (October 1961), 741–765.

PETERKA, V., *Macrodynamics of Technological Change: Market Penetration by New Technologies*, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1977.

ROGERS, E. M., *The Diffusion of Innovation*, Third Ed., Free Press, New York, 1983.

SAS Institute, *SAS/ETS User's Guide, Version* 5 *Edition*, SAS Institute, Cary, N.C., 1984.

SHARIF, M. N. AND C. KABIR, "A Generalized Model for Forecasting Technological Substitution," *Technological Forecasting and Social Change*, 8 (1976a), 353–364.

——— AND ———, "System Dynamics Modeling for Forecasting Multilevel Technological Substitution," *Technological Forecasting and Social Change*, 9 (1976b), 89–112.

SRINIVASAN, V. AND C. H. MASON, "Nonlinear Least Squares Estimation of New Product Diffusion Models," *Marketing Sci.*, 5 (1986) 169–178.