Diffusion Models:
Managerial Applications and Software

Gary Lilien,
Arvind Rangaswamy
The Pennsylvania State University
and
Christophe Van den Bulte
University of Pennsylvania

Diffusion Models: Managerial Applications and Software[1]

Gary L. Lilien, Arvind Rangaswamy and Christophe Van den Bulte

(Version **5-20-99**)

Abstract

In this paper we review managerial applications of diffusion models in marketing. We first develop definitions for the basic Bass model and some of its key extensions. Following this, we briefly review a number of applications from which we draw some lessons on building and using diffusion models for predictive and normative applications. We also provide a Microsoft Excel-based software that implements the Bass and generalized Bass diffusion models along with a tutorial to illustrate how to use the software. Finally, we include a case on High Definition Television (HDTV) that allows the reader to explore how to use diffusion models (with our software) to address a business forecasting problem.

The other chapters in this book have focused on analytic developments and extensions of quantitative diffusion models. In this chapter we will focus on the *application* of *quantitative diffusion models* to aid in *new product planning* and *decision making*.

1. *Application.* By application, we mean the use of a model or model results to generate actionable information *by a decision-maker*. The demonstration that a model fits retrospective data, or provides post-hoc explanation for a phenomenon, or helps in generating theoretical insights, falls outside our definition of application.

2. *Quantitative diffusion models.* We focus primarily on aggregate diffusion models representing the market penetration of a new product, process, or technology. Quantitative diffusion models specify mathematical relationships between quantifiable variables and include parameters that allow the model to be customized for a specific application.

3. *New product planning and decision making.* We focus on applications in the product planning area, where the models may be used

   - to *describe* the rate of diffusion, to provide a better understanding of the drivers of adoption;
   - to *predict* the future penetration trajectory, so that growth may be planned for;
   - to *control* the future penetration trajectory to provide inputs for investment, pricing, advertising and product development decisions (normative use).

Our view is that there is no best, single, applied diffusion model. The problem, the available data, and the specific characteristics of the situation combine to suggest the most cost-effective approach.

As with several other papers in this book, we will use the basic Bass model as our foundation and describe some applications. In the sections that follow, we first develop the formal definitions for the basic Bass model and some of its key extensions. We then discuss how to obtain parameter values for these models early in the product life cycle, when limited penetration (or sales) data are available. Following this, we briefly review a number of applications and comment on the uses and limitations of aggregate diffusion models for predictive and normative use.

Afterwards we provide a tutorial that illustrates how to use the spreadsheet software included with this book. The software allows the user to estimate the parameters of the basic and generalized Bass model (except market size) to plot the resulting penetration trajectory and to run sensitivity analyses. Finally, we include a case on High Definition Television that allows the reader to explore the use of the Bass model and our software to address a critical business problem. Interested readers can also use a two-part case study on forecasting compact disc sales at RCA. That case, developed by Professor Barry Bayus, is included in Rao and Steckel (1998).

The Bass Model and Some Extensions

The Bass model is a good starting point for forecasting the long-term penetration pattern of new technologies and products under two types of conditions: (1) the firm has recently introduced the product or technology and has observed its penetration for a few time periods, or (2) the firm has not yet introduced the product or technology, but it is similar in some way to existing products or technologies whose diffusion history is known. The model attempts to predict how many customers will eventually adopt the new product, and when they will adopt. The question of *when* is important, because answers to this question guide the firm in its deployment of resources in marketing the innovation.

The software included with this book uses the basic Bass model as well as some of its extensions. The basic model (Bass 1969), using discrete time notation, can be written as:

$$x(t) = [p + q (X(t-1)/m)] [m - X(t-1)], \quad \text{where} \tag{1}$$

$x(t)$ =   the number of adoptions occurring in period t,

$X(t-1)$ =   the cumulative number of adoptions having occurred before period t,

p =   coefficient of innovation, capturing the intrinsic tendency to adopt as well as the effect of time invariant external influences,

q =   coefficient of imitation or social contagion, capturing the extent to which the probability that one adopts (given that one has not done so yet) increases with the proportion of eventual adopters that has already adopted,

m   =   the number of eventual adopters.

Note that the model is framed in terms of number of adopters (penetration) and not number of units sold (sales). When adopters acquire multiple units (automobile, televisions, etc.) or when adopters replace existing units, penetration and sales series may differ significantly.

Bass, Krishnan, and Jain (1994) have proposed a generalized form that incorporates the effects of marketing mix variables on the likelihood of adoption. They do so by adding a multiplicative factor *Z(t)* to the original model (1):

$$x(t) = [p + q \ (X(t\text{-}1)/m)] \ [m - X(t\text{-}1)] \ Z(t). \tag{2}$$

They operationalize *Z(t)* as:

$$Z(t) = \ 1 + \ \alpha \ [P(t) - P(t\text{-}1)]/P(t\text{-}1) \ + \ \beta \ \max \{0, \ [A(t) - A(t\text{-}1)]/A(t\text{-}1)\} \ , \text{ where} \tag{3}$$

α =   coefficient capturing the percentage increase in diffusion speed resulting from a 1% decrease in price,

P(t)   =   price in period *t*,

β =   coefficient capturing the percentage increase in diffusion speed resulting from a 1% increase in advertising,

A(t)   =   advertising in period *t*.

Note that Bass, Krishnan and Jain (1994) link diffusion speed only to increases in advertising levels but have diffusion speed insensitive to *cuts* in advertising, a formulation that might be questioned. Still, the software included with this book closely follows their work and uses the same operationalization.

Another extension, which can be made to both the original and the generalized Bass model, is to allow the number of eventual adopters to change over time. The number of adopters may change because the overall population increases or decreases. Another driver for change could be an improvement in the product's distribution infrastructure, making the product physically available to an increasing number of consumers or companies. The number of eventual adopters can also change as a function of complementarities with different products. The diffusion of online grocery shopping, for instance, is limited by the penetration of Internet access at home, which in turn is limited by the penetration of home computing. If the potential population is related to the price of the product, one could express the number of potential adopters as:

$$m(t) = \ m(1) \ (1 + r)^{t\text{-}1} \ [P(t)/P(1)]^{-\eta}, \quad \text{where} \tag{4}$$

m(t)   =   the number of eventual adopters in period *t*,

r =   market growth rate (apart from price effects),

$P(t)$ =   price in period t,

η =   elasticity of the number of eventual adopters with respect to the innovation's price.

Obtaining Model Parameter Values

To forecast the adoption path of a new product by means of a diffusion model, we need to have values for the model's parameters. This is a particularly thorny issue since good forecasts for new product adoption are needed early on in the product's life, when very little data is available.

If we have adoption data available, usually for four or more periods, we can obtain parameters for p and q by estimating the basic Bass model (eq. 1) or generalized Bass model (eq. 2) using a variety of estimation techniques discussed elsewhere in this volume. The software included with this book calculates those parameter values for p and q that minimize the squared deviations between the data series and the adoption levels, $x(t)$, predicted by the model. Note that the software does not estimate m, but requires the user to specify a value. We implemented this approach because non-linear least squares estimation of diffusion models often produces estimates of m that are too low, so it is preferable to use exogenous information on m rather than having it be a part of the estimation procedure (e.g., Trajtenberg and Yitzhaki, 1989; Van den Bulte and Lilien, 1997).

If no data are available, one may gain insight from the diffusion history of analogous products. The procedure is rather simple in spirit: identify products whose diffusion path was probably similar to how the new product will diffuse, get parameter values of p and q for these previous innovations, plug them into a diffusion model for the new product (with m estimated separately, perhaps through a survey), and forecast the diffusion path of the new product. The challenge is to find the right analogous products. Some formalized methods have been developed to do so, but one must remain careful and creative in how one chooses the analogs. Analogies based on similarities in expected market behavior work better than analogies based on product similarities. For example, in forecasting the diffusion path of online services, it may be better to use cellular telephone subscriptions as an analog, rather than cable TV subscriptions.

Thomas (1985) recommends that in selecting analogs, we consider similarities along five bases: environmental context (e.g., socioeconomic and regulatory environment), market structure (e.g., barriers to entry, number and type of competitors), buyer behavior (buying situation, choice attributes), marketing mix strategies of the firm, and characteristics of the innovation (e.g., relative advantage over existing products and product complexity). If necessary, we can consider multiple analogs and take the (weighted) average of their p and q values. Choffray and Lilien (1986), Sultan, Farley and Lehmann (1990), and Lenk and Rao (1990) provide more formal procedures to incorporate data from analogous products into the parameter estimation procedure. Exhibits 1a and 1b summarize parameter estimates for various innovations (we have included estimates from Exhibit 1a within the software to help in selecting analogs). Appendix

C expands on the choices we made regarding data and estimation methodology in computing the parameter estimates reported in Exhibits 1a and 1b. Graphs with the original data and the estimated Bass diffusion curve based on Exhibit 1a are also included in the software to help select analogs. We emphasize, though, that analogs should be based on product and market characteristics rather than simply mirroring the path managers hope their new product will follow.

Once we determine the parameter values by estimation or by using analogs, we can put these values into a spreadsheet to develop forecasts (see the software tutorial for an illustration). The software includes options for estimation, for selection of parameters via analogy and for forecasting.

Other papers in this book describe many of the extensions of the basic and generalized Bass model, incorporating issues such as competition, product replacement, multiple purchases, multi-product interactions, multi-market effects, and the like as well as a discussion of associated estimation issues.

Product/Technology	Period of Analysis	p	q	m
Agricultural				
Tractors (thousands of units)	1921-1964	.000	.134	5201.0
Hybrid corn	1927-1941	.000	.797	100.0
Artificial insemination	1943-1959	.028	.307	73.2
Bale hay	1943-1959	.013	.455	92.2
Medical Equipment				
Ultrasound imaging	1965-1978	.000	.534	85.8
Mammography	1965-1978	.000	.729	57.1
CT scanners (50-99 beds)	1980-1993	.044	.350	57.9
CT scanners (>100 beds)	1974-1993	.036	.268	95.0
Production Technology				
Oxygen steel furnace (USA)	1955-1980	.002	.435	60.5
Oxygen steel furnace (France)	1961-1980	.008	.279	88.4
Oxygen steel furnace (Japan)	1959-1975	.049	.333	81.3
Steam (vs. sail) merchant ships (UK)	1815-1965	.006	.259	86.7
Plastic milk containers (1 gallon)	1964-1987	.020	.255	100.0
Plastic milk containers (half gallon)	1964-1987	.000	.234	28.8
Stores with retail scanners (FRG, units)	1980-1993	.001	.605	16702.0
Stores with retail scanners (Denmark, units)	1986-1993	.076	.540	2061.0
Electrical Appliances				
Room air conditioner	1950-1979	.006	.185	60.5
Bed cover	1949-1979	.008	.130	72.2
Blender	1949-1979	.000	.260	54.5
Can opener	1961-1979	.050	.126	68.0
Electric coffee maker	1955-1979	.042	.103	100.0
Clothes dryer	1950-1979	.009	.143	70.1
Clothes washer	1923-1971	.016	.049	100.0
Coffee maker ADC	1974-1979	.077	1.106	32.2
Curling iron	1974-1979	.101	.762	29.9
Dishwasher	1949-1979	.000	.213	47.7
Disposer	1950-1979	.000	.179	50.4
Fondue	1972-1979	.166	.440	4.6
Freezer	1949-1979	.019	.000	94.2
Frypan	1957-1979	.142	.000	65.6
Hair dryer	1972-1979	.055	.399	51.6
Hot plates	1932-1979	.056	.000	26.3
Microwave oven	1972-1990	.002	.357	91.6
Mixer	1949-1979	.000	.134	97.7
Power leaf blower (gas or electric)	1986-1996	.013	.315	26.0
Range	1925-1979	.004	.065	63.6
Range, built-in	1957-1979	.048	.086	21.7
Refrigerator	1926-1979	.025	.126	99.7
Slow cooker	1974-1979	.000	1.152	34.4
Steam iron	1950-1979	.031	.128	100.0
Toaster	1923-1979	.038	.000	100.0
Consumer Electronics				
Cable television	1981-1994	.100	.060	68.0
Calculators	1973-1979	.143	.520	100.0
Camcorder	1986-1996	.044	.304	30.5
CD player	1986-1996	.055	.378	29.6
Cellular telephone	1986-1996	.008	.421	45.1
Cordless telephone	1984-1996	.004	.338	67.6
Electric toothbrush	1991-1996	.110	.548	14.8
Home PC (millions of units)	1982-1988	.121	.281	25.8
Radio	1922-1934	.027	.435	100.0
Telephone answering device	1984-1996	.025	.406	69.6
Television, black and white	1949-1979	.108	.231	96.9
Television, color	1965-1979	.059	.146	100.0
VCR	1981-1994	.025	.603	76.3
Average		.037	.327	
25th percentile, median, 75th percentile		.004, .025, .054	.134, .280, .435	

Unless indicated, the model was estimated on penetration data collected in the USA.

Exhibit 1a Parameters of the Bass model in several product categories based on penetration data and long data series.

Product/Technology	Period of Analysis	p	q	m
Agricultural				
Tractors (thousands of units)	1921-1931	.000	.211	1324.0
Hybrid corn	1927-1939	.000	.798	100.0
Artificial insemination	1943-1953	.000	.567	56.9
Bale hay	1943-1955	.006	.583	80.3
Medical Equipment				
Ultrasound imaging	1965-1977	.001	.510	89.2
Mammography	1965-1976	.000	.738	56.4
CT scanners (50-99 beds)	1980-1990	.036	.572	47.8
CT scanners (>100 beds)	1974-1985	.034	.254	100.0
Production Technology				
Oxygen steel furnace (USA)	1955-1970	.000	.477	56.2
Oxygen steel furnace (France)	1961-1974	.003	.384	58.2
Oxygen steel furnace (Japan)	1959-1968	.048	.324	83.9
Steam (vs. sail) merchant ships (UK)	1815-1900	.000	.311	77.0
Plastic milk containers (1 gallon)	1964-1975	.024	.331	73.6
Plastic milk containers (half gallon)	1964-1973	.040	.630	4.4
Stores with retail scanners (FRG, units)	1980-1993	.001	.605	16702.0
Stores with retail scanners (Denmark, units)	1986-1993	.076	.540	2061.0
Electrical Appliances				
Room air conditioning	1950-1963	.016	.304	24.2
Bed cover	1949-1962	.002	.177	64.2
Blender	1949-1960	.023	.199	10.3
Can opener	1961-1971	.027	.341	51.8
Electric coffee maker	1955-1965	.001	.302	72.8
Clothes dryer	1950-1960	.009	.514	18.2
Clothes washer	1923-1936	.004	.093	100.0
Coffee maker ADC	1974-1979	.077	1.106	32.2
Curling iron	1974-1979	.101	.762	29.9
Dishwasher	1949-1974	.000	.189	57.4
Disposer	1950-1966	.008	.256	15.5
Fondue	1972-1979	.166	.440	4.6
Freezer	1949-1959	.043	.213	25.3
Frypan	1957-1967	.301	.000	51.0
Hair dryer	1972-1979	.055	.399	51.6
Hot plates	1932-1942	.095	.143	18.2
Microwave oven	1972-1983	.012	.383	33.1
Mixer	1949-1959	.000	.145	83.0
Power leaf blower (gas or electric)	1986-1996	.013	.315	26.0
Range	1925-1935	.071	.000	10.2
Range, built-in	1957-1969	.030	.000	41.3
Refrigerator	1926-1940	.015	.290	69.5
Slow cooker	1974-1979	.000	1.152	34.4
Steam iron	1950-1960	.000	.376	63.8
Toaster	1923-1933	.039	.262	46.2
Consumer Electronics				
Cable television	1981-1991	.080	.167	60.8
Calculators	1973-1979	.143	.520	100.0
Camcorder	1986-1996	.044	.304	30.5
CD player	1986-1996	.055	.378	29.6
Cellular telephone	1986-1996	.008	.421	45.1
Cordless telephone	1984-1994	.000	.438	54.0
Electric toothbrush	1991-1996	.110	.548	14.8
Home PC (millions of units)	1982-1988	.121	.281	25.8
Radio	1922-1933	.028	.422	100.0
Telephone answering device	1984-1994	.019	.481	63.4
Television, black and white	1949-1959	.100	.353	90.1
Television, color	1965-1975	.058	.168	97.1
VCR	1981-1991	.011	.832	67.5
Average		.040	.398	
25th percentile, median, 75th percentile		.001, .021, .055	.255, .365, .519	

Unless indicated, the model was estimated on penetration data collected in the USA.

Exhibit 1b Short data series parameters of the Bass model. These parameters are based on penetration data with shorter data series than in Exhibit 1a. Wherever feasible, data length is at least 10 observations and right-censored at 2 observations after the first peak.

Diffusion Model Applications

In line with our goal here, we describe briefly some diffusion model applications and attempt to draw some lessons from each.

Consumer Durable Model with Price, Advertising and Replacement (Kalish, 1985): The product class dealt with here is a consumer durable, with an average lifetime of seven to ten years. The product was relatively new, having been introduced seven years previously and had experienced consistent growth during that period. The diffusion curve was leveling off, and the question facing management was what future sales were going to look like.

Historical growth had been accompanied by a price decline and a competitive entry into the market, factors that needed to be addressed. In addition, the product was in a stage where replacement sales were important. Advertising was considered a major sales influence by management. Consequently, competition, price, and advertising had to be included in the model.

The problem was decomposed into an aggregate product class sales model and a market share model for the different competitors and brands. Market share was structured as a nested logit model. The product class sales model was formulated as a diffusion equation, incorporating replacement, price, and advertising.

The product class sales model was estimated using quarterly data, from introduction until the most recent time period using non-linear least squares. The market share model was estimated using the actual market shares, prices, distribution penetration, advertising expenditures, and features of the existing and past brands and the fit was quite good. The signs of the coefficients turned out as expected, with price, advertising, and distribution statistically significant.

The model was implemented as an interactive, on line decision support system. Several forecasts were produced that turned out to be more accurate than what could be expected from the standard errors of the estimates. Similarly, the market share model's predictions were very good. It forecast sales prior to launch for two new products. The model predicted that the higher priced product would have a larger market share than the second new product, because an added characteristic was valued more than the price difference. This was in contrast to a sales forecast produced elsewhere in the company that predicted equal shares. The model's prediction turned out to be accurate, evidenced by stockouts of the first product, and large inventories of the second. The company subsequently increased the price of the first product.

The procedure was not widely adopted subsequently within the company, perhaps because the model and the entire DSS were not readily understood by management.

Lesson: With a sophisticated diffusion analysis system like this, it may be better to provide the forecasts directly to managers rather than to provide them in the form of a (potentially intimidating) decision support system. Models are often best left in the hands of analysts and modelers. Managers are served best if they help define the issues, are engaged in the model development process, and get quick access to model results.

A Market Entry Timing Model (Kalish and Lilien, 1986a)**:** In 1980 the U.S. Department of Energy's photovoltaic (PV—solar batteries) program was reviewing a proposal from a developer

to build, with government support, 100 photovoltaic homes in a cluster in the southwest. Although this proposal was a welcome initiative from a private sector developer, DOE expressed concern that the technology might not be far enough along to demonstrate it in this way. The question was: when *should* such a demonstration program begin?

The authors developed a diffusion model in which the word of mouth term (the *q* parameter) is a function of the quality of the technology, which is assumed to be changing over time. They expanded that term to carry over (and discount) the effect of good (and bad) word of mouth from previous time periods. They did so because if the technology is improving over time, too early an introduction can provide negative feedback that can delay market acceptance. The authors took as their goal to maximize market penetration by a target year (1990), given a demonstration budget constraint.

In order to assess "quality" levels, the authors worked with staff at Massachusetts Institute of Technology's Lincoln Laboratories, the organization under contract with DOE for technical evaluation, installation, and monitoring of photovoltaic experimental and demonstration programs. Lincoln Laboratories' staff identified three categories of PV system failure: (1) lowered array output due to panel degradation, (2) system breakdown, primarily due to power conditioning unit malfunction, and (3) damage to residence through fire resulting from short circuit. Lincoln Laboratories also provided estimates for the 1980 rate of incidence of these failure categories along with technological forecasts of how those likelihoods would change by 1986.

The authors performed a telephone survey of builders to assess their perceptions of the impact on the marketplace of each of these types of futures relative to the positive effect of a perfectly running demonstration. The rest of the model was calibrated using outputs from other planning models in current use at the Department of Energy. They used the calibrated model, together with assumptions about possible negative effects of degrees of PV system failures to simulate the effect of a demonstration program at any time from 1980 on. The model results pointed to a large improvement in market penetration by the year 1990 if introduction waited until 1986 (versus the planned introduction in 1980). Sensitivity analysis also showed that the model implication—wait until 1986 to introduce—was relatively robust with respect to the intensity of the demonstration and market entry program, the carryover factor and the target year.

Partly on the basis of this analysis, funding for the demonstration program was *not* granted; the developer was asked to wait several years until the technology was more technically proven.

Lesson: *Diffusion models can provide a disciplined approach to support product launch timing decision (a situation where there is clearly no available sales data) where product quality is assumed to vary over time.*

A Prelaunch Sales Forecasting System (Choffray and Lilien, 1986): In 1982, a leading European paper producer, was dissatisfied with the judgment-based sales forecasts for its new products. Its management hoped to be able to reduce forecasting error and lower the rate of new product failure due to market misassessment. They searched for "lookalike" product market situations, hoping that the way a new product is accepted in the marketplace would be close to the way similar products were accepted.

They found that the Center for Research in Management Science at Ecole Supérieure des Sciences Economiques et Commerciales (ESSEC) in conjunction with the French Ministry of Industry and the Novaction Company, a French consulting firm, had developed a data base of individual new products, including information on the development process, the marketing strategy and the rate of market penetration for a five-year period.

For each of the 112 products included in the database, over 500 pieces of information were collected in three categories:

- *R&D process.* Cost structure, financing, duration, methods of evaluation, types of protection, etc.
- *Market introduction strategy.* Bases for decision, success or failure, evaluation criteria, initial marketing mix, etc.
- *Rate of product penetration.* Sales volume and dollar sales for the new product and its prime competitors, market structure, changes in the marketing mix, etc.

Market penetration information was collected on a quarterly basis over a five-year period after market introduction. Other data included managerial judgments about how the new product performed relative to competition, information on the objectives set for the new product, the way these objectives evolved over time, and how they were achieved. These data were related to the parameters of a diffusion model and were incorporated in a computer based decision support system. One module of the system included fifteen questions about market structure, entry strategy and the development process and develops a four-year sales projection.

After using the system, the company reported that they are pleased with the results, giving them four-year cumulative sales projections with less than a 30 percent discrepancy between actual and forecast sales (Virolleaud, 1983).

This system was also used at Vieille Montagne, a world leader in zinc production and associated technologies, to simulate the time growth of cumulative sales for a new product introduced five years earlier. Discrepancy with the actual sales rate was less than 15 percent over that horizon. Subsequently, the approach was used to help plan the entry strategy for a line of new products.

This case illustrates how the analogue approach can be used to parameterize a diffusion model *prior* to market introduction and the consequent collection of sales data. The results were superior to a previously used judgmental approach. The incorporation of the results into a user-friendly DSS, properly introduced to appropriate employees, may have been a factor in the model's acceptance.

Lesson: A formal approach to selecting analogies can help improve the acceptance of a diffusion model as can an appropriate, user friendly software implementation

Forecasting Sales of New Contingent Products (Bayus, 1987): Bayus reports the results of a study done at RCA/Ariola to forecast the sales of a product when those sales (compact disks) are dependent on the sales of another product (CD players). His model accounted for hardware sales (via a standard diffusion model) but linked software sales to the market potential generated by the diffusion model. The model incorporates the effects of different market segments, pricing variations, awareness and purchase intentions and permits analysis of the impact of a new entry (digital audio tape). The model predicted sales quite well. In addition management reported that

a major benefit of the model was the identification of the key data elements necessary to forecast sales.

Lesson: The diffusion modeling process itself can provide key management benefits, such as identification of necessary data and refinement of managerial judgment that are critical model inputs.

Forecasting the Trajectory of a New Auto Model: John Roberts, Glen Urban and John Hauser report that Buick was launching a new electronic dashboard, down-sized, front wheel drive car. Its challenge was to forecast sales over time, the other General Motors models it would cannibalize, brand price sensitivity, and how the car should be positioned. They used Roberts and Urban's (1988) individual-level multiattribute level diffusion model and calibrated it in a laboratory setting with respondent estimates at different stages of the progressive information collection. Sales of a control car were also used to get a transformation function to calibrate what was happening in the laboratory with what would happen in the marketplace. The model helped Buick position the car. Financial planning was based on the model forecasts, including features planning and pricing. In a later application of the model, the prelaunch projected sales were much lower than expected, which encouraged Buick to keep an old car model in production to maintain sales levels (see also Urban, Hulland and Weinberg, 1993; and Urban, Weinberg and Hauser, 1996).

Lesson: Laboratory-based data can be used as inputs to diffusion models, producing pre-launch positioning and long-term product planning implications.

The cases above appeared in the published literature; in addition, we conducted a brief, informal survey amongst some of the key diffusion researchers, with the following results:

Market Planning Model for Mathworks: James M. Lattin and John H. Roberts report an application for forecasting the diffusion of new software for The MathWorks, Inc., where the problems were targeting, cash flow planning, positioning, and feature planning. They used the Lattin and Roberts (1998) individual level diffusion model, calibrated with market research data (intentions to purchase at different time horizons, perceptions, preference, uncertainty), data from analogous software products, management judgment, as well as measures of perceived benefits of new products/features to customers. The model was used to develop pre-launch forecasts that were updated over time. The modeling process itself was a useful exercise in training managers to think about their target market, relevant attributes, segmentation, and market dynamics, etc. An important insight from this study is that non-response bias in purchase intentions surveys is likely, which could lead to misleading results—this bias occurs because prime prospects for the product are the ones that are more likely to respond to the survey.

Defensive Strategies Against Innovation Diffusion: John Roberts reports that Telstra, the Australian telephone company, was facing a new entrant offering a differentiated product which would reduce Telstra's brand share. The challenge to Telstra was how to minimize the equilibrium share gain by the new entrant (m) as well as to slow the trajectory to that share (p and q). They developed a Bass model as well as a four-state diffusion model to model customer flows. They derived a logit model, which not only gave equilibrium share, but also its underlying drivers. The data used were self stated adoption intentions from respondents after

progressive information exposure as well as perceptions, preferences and perceived uncertainty. An embedded conjoint analysis enabled Telstra to test its own innovations, including price changes as a way to protect its share. They estimated the other coefficients using two methods: firstly from the diffusion of new entrants' share in the U.S. and secondly from respondents' self-stated time of adoption (for adopters). Because these self-stated adoption times were available on an individual basis they were able to model the rate parameters as a function of respondent characteristics and perceptions before product launch. The Marketing Director of Telstra described the research as the single most important piece of market research that Telstra had undertaken. The model was used for resource allocation and financial planning, pricing option evaluation, and strategic service initiatives. It was validated post launch and was accurate to within .1 share point after six months. The model was recalibrated after six months by combining early sales data with the other data. Given that the model was tracking well, this did not lead to any substantial new insights (except that defense at the repeat stage was easier than defense at the trial stage).

Movie Performance Forecasting: Josh Eliashberg, Mohan Sawhney, Berend Wierenga and Jedid-Jah Jonker report an application dealing with forecasting and diagnosis of the commercial box-office performance of an American movie prior to its release in Europe for RCV (movie distributor) and Pathé (movie exhibitor) in Holland. They developed a probabilistic flow model in which word of mouth is explicitly incorporated, and calibrated the model based on consumers inputs, similar in spirit to a simulated test market. Their model is an aggregate model based on interactive Markov systems. Calibration data came from consumers who were exposed to the type of information that normally occurs in the market for movies when the movie is released (i.e., advertising, testimonials, trailers, and reviews). The client used the approach to help set its movie introduction budget for TV spots, print ads, and trailers. A lesson learned in this application is that forecasts were very sensitive to estimates of word-of-mouth effects. Thus, careful measurement of word of mouth is an important requirement for applying the Bass model, especially when very little sales data are available (see also Eliashberg and Sawhney, 1994; and Sawhney and Eliashberg, 1996).

Market Planning for Investment Analysis: Frank Bass reports several applications: One was for satellite TV, where an electronics company required a subscriber growth forecast for investment purposes. He applied the Bass model, using survey (intentions) data plus analogies to develop a forecast that showed management that even the lower bound forecast exceeded the necessary conditions for profitability. This helped build management confidence in the product. Another application, for a satellite phone, was for a consortium of companies also concerned with investment decisions based on a demand forecast. He applied the Bass model, the multiple generation Bass model, and the generalized Bass model using survey data, information from earlier generation technologies, and pricing data supplied by the client to produce the forecasts. A main benefit of the model was that it helped the companies feel more confident of the forecasts when making investment decisions. In applying this model, the managers learned that the model does not forecast changing technologies and competitive conditions, but only forecasts the product sales curve under different assumptions about the technology and competitive context. Thus, the diffusion model must be supplemented with other methods to forecast changing market and technological conditions.

Lessons: These cases show that, with sufficient care and blending of survey data, judgments and some sales data, diffusion models can provide a systematic approach to address a wide range of

new product planning challenges. But sample selection and survey-question design is critical in developing valid, supportable results.

Other Applications. There are other applications of diffusion models in one form or another in both the public and the private sector. Some examples include forecasts of the diffusion of new financial instruments (Molyneux and Shamroukh, 1996), linking movie advertising to box office performance (Zufryden, 1996), estimates of the extent of software piracy (Givon, Mahajan and Muller, 1995), forecasting the adoption of new medical technology (Sillup, 1992), an analysis of the causes and prospects for various population segments concerning quitting smoking (Redmond, 1996), forecasting regional penetration of new farming practices (Steffens, 1998), forecasting the sales of fax machines, integrating choice models with diffusion models (Weerahandi and Dalal, 1992), and an application to assess the potential lost sales to Polaroid due to competitive entry by Kodak (Mahajan, Sharma, and Buzzell, 1993). These applications suggest that both the concepts behind diffusion models and the associated technology to estimate calibrate and apply them can be extremely valuable when properly applied.

Assessment

A diffusion model's applied value is determined by its ability to help managers analyze, plan, and control the process of introducing and managing a new product. It is difficult to project and control sales growth when little or no sales data are available and the competitive structure of the market is uncertain. Diffusion models can help in these tasks if they are used appropriately. They are not meant to replace management judgment; rather they should be used to aid that judgment, to help run sensitivity analyses and to compare the attractiveness of alternative market scenarios.

In what follows we group together a few suggestions for applying diffusion models in two categories: those dealing with *model structure* and those dealing with *model use.*

Model Structure

Market Potential. For goods with a long lifetime, market potential is a key factor affecting the model's predictive value. Market potential for the product class should be modeled as a function of population changes over time. There is more ambiguity in whether price, economic conditions, distribution penetration, and marketing communications effort should also be modeled as affecting market potential. One can equally well argue that these factors affect the probability of adoption at a particular time (the hazard rate) rather than the probability of eventual adoption.

Control Variables. The model should incorporate relevant control variables: price, advertising and quality for example. Without those variables, the model may not be useful for policy making or for conducting sensitivity analyses. A model with control variables can be used to evaluate different marketing programs, and study the evolution of the market under different price scenarios.

Competition. Many past models ignored competitive effects, which are of vital managerial concern. A way to incorporate competition is to model product class adoptions using the diffusion model, and brand share using a market share model.

Quality and Risk. Individuals perceive new products as riskier than established ones due to uncertainty about product performance. Product quality and reliability may indeed change over time. The issues of uncertainty, information diffusion, and perceived and actual product quality should be considered in the model.

Replacement Sales. Replacement sales become a major issue later in the life cycle. It is important to model replacement explicitly, particularly for products that are extensions of a product line and are not completely new.

Model Use

Early Calibration. The model is likely to be most needed before the product is introduced. At that time no adoption data are available. Thus the model should be parameterized in such a way that the parameters have an intuitive interpretation, and the model can be calibrated using a combination of survey data and managerial judgment. Methods for updating the parameters once sales data are available should be provided. Alternatively, suitable analogues should be sought and a procedure for calibrating models using those analogues developed.

Simplicity and Validity. The model structure should be made simple and intuitive so that the manager can understand it and be comfortable with it. The results of the modeling effort should be presented visually to highlight key insights from the model. The model should also be externally valid. (One can demonstrate validity by showing how the model would have performed in predicting sales of a previous innovation in the company, for which current sales data are available.)

User Involvement and DSS. As with any model, when the users and stakeholders are involved in the model-building process and the user can assess and analyze the results through an understandable decision support system, the results are more likely to be used. The software described in the next section illustrates some aspects of such a system
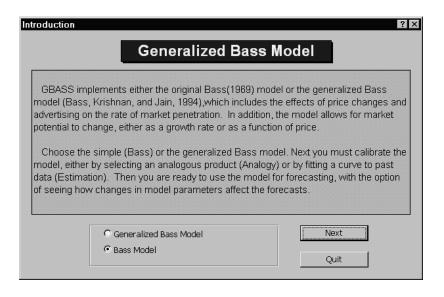
On net, diffusion models *can* be valuable and can be used in managerial settings. They are flexible and widely applicable. The application challenge is not so much to build better diffusion models, as it is to customize models using existing concepts and to make them work in managerial contexts. To meet this challenge we must provide the means to estimate model parameters credibly, to explain the concepts to management, and make the model-manager interface efficient, so these models can provide a net positive value to the user. We also need more published reports of diffusion model application so that we can integrate and build upon the many lessons that have already been learned—what decision contexts are most appropriate for applying diffusion models and what are really the best ways to both calibrate and use them.

Appendix A

Software Tutorial for the (Generalized) Bass Model Tutorial[2]

The software included here runs under Windows 95, Windows 98 or Windows NT. To run this software you will need Microsoft Excel 7 or higher. You will also need access to Excel's Solver tool, not available as part of the default Excel installation. Under the **Tools** menu on your version of Excel, check the list of **Add-Ins** to see whether Solver is included. If not, run the Excel (or MS Office) setup procedure with the original CD and select the Solver Add-in. We also maintain a FAQ page on this software at our web site: www.smeal.psu.edu/courses/mktgeng.

The diskette includes five files: gbass.xls, modgen99.ind, gbass.cnt, gbass.hlp and analogy_data.xls (which contains the raw data for our 54 analogy products). **Copy** all files to any directory on your hard drive. Open Excel and select gbass.xls. to see the **Introduction** screen.

First select either the **Generalized Bass Model** or the **Bass Model** and click **Next**. The generalized Bass model includes two decision variables, pricing and advertising, which are assumed to determine the speed of diffusion. The Bass model sets up the original model without decision variables.
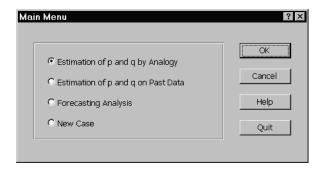
Because the generalized Bass model includes the Bass model, we will describe its use. Both versions have the same setup.

---

[2] Drawn from Lilien and Rangaswamy (1998)
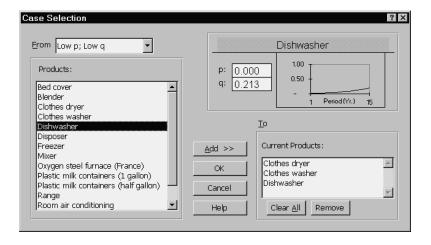
Model calibration by analogy

When there are no past data for the product of interest, calibrating the model by analogy can be useful. This can also be useful when you don't have enough data to feel confident about estimating numerical parameters for the model.

You need to identify an analogous product or technology that has market characteristics similar to those of the product you want to analyze.

The software includes a database that contains actual data points, estimated p and q coefficients, and estimates of market potential several product categories. For further information about the database, see Exhibit 1, which lists the p and q coefficients, the market potential, and the time period of analysis for each case.

To explore the sales patterns of analogous products, select **Estimation of p and q by Analogy** and click **OK**.
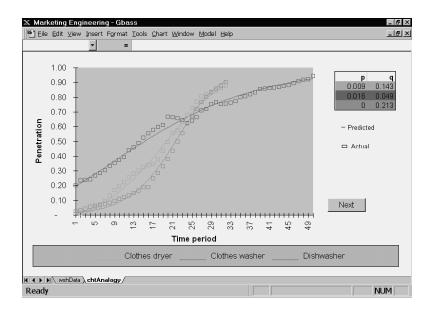
First select a category. All cases have been divided into four categories: (low p, low q), (high p, low q), etc., using the median values of

p and *q* in Exhibit 1 as cut offs. When you click a product in the **Products** list, you will see a preview of its curve and its coefficients.
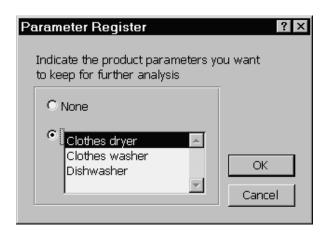
Now choose a product and click **Add>>** to add a product to the group of potential reference cases in the **Current Products** area. (Add no more than three products, since at most three curves can be graphed at a time.)

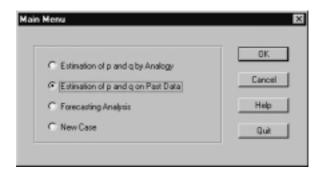After you choose your cases and click **OK**, the program will chart the actual data points and the estimated diffusion curves for them.

To simplify comparisons we have normalized the available cases to a maximum market penetration equal to 1.

Click **Next** to get to the next box. Indicate the product that you think offers the best analogy and that you want to keep for further reference.
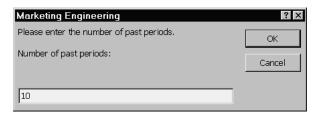

Click **OK** to get to the **Main Menu**.
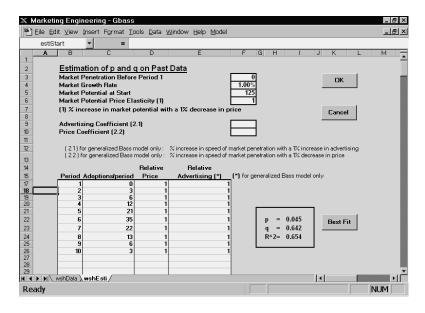
Model calibration by estimation

To estimate the model parameters numerically (using Solver), choose **Estimation of q and p on Past Data** and click **OK**. Enter the number of past periods for which you have data. Click **OK**.

*NOTE: Once you have specified this number you will not be able to change it for subsequent estimations for this product. To make estimates for a different number of periods you must go to the **Main Menu** and select **New Case**.*

Next, enter data for **Market Penetration Before Period 1, Market Growth Rate**, **Market Potential at Start** (your estimate of the total market size at the starting period), and **Market Potential Price Elasticity**.

If you chose the **Bass Model,** you don't need to provide values for the advertising or price coefficients. If you chose the **Generalized Bass Model**, you must provide estimates for these coefficients, because historical data rarely have enough variability to permit estimation of these parameters. The advertising and price coefficients can be roughly thought of as "market acceptance speed elasticities," indicating the speed with which the market adopts the new product:

- The advertising coefficient reflects the percent increase in speed of market acceptance with a one percent increase in advertising. (Documented values for the advertising coefficient typically range between 0.3 and 1.)
- The price coefficient reflects the percent increase in speed of market acceptance with a one percent decrease in price. (Documented values for the price coefficient typically range between 1 and 2.)
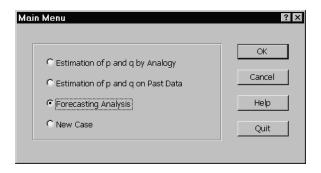
Now enter the data on adoption in each period and, optionally, an index for price and advertising in each period. Note that marketing effort and price should be measured relative to a base level in period 1 indexed to 1.0. Thus, if advertising was \$4 million in period 1 and \$6 million in period 2, the index in period 2 is 1.5.

Click **Best Fit** to start the calibration of the model. The program estimates only the coefficients for p and q. The market potential estimate is fixed at your best guess input.
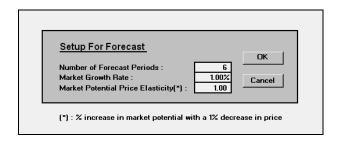
Click **OK** to go back to the **Main Menu**.

NOTE: It is sometimes possible to obtain negative R-Square with nonlinear least squares estimation for some ill-structured data. However, more likely reasons for this situation are (1) you made an error in your input data, or (2) the nonlinear estimation procedure (in Excel's Solver) started the estimation in an infeasible region. For example, you can get a negative R-Square if you enter cumulative sales data instead of sales per period in the data sheet, or if you specify a market potential that is smaller than first period sales.
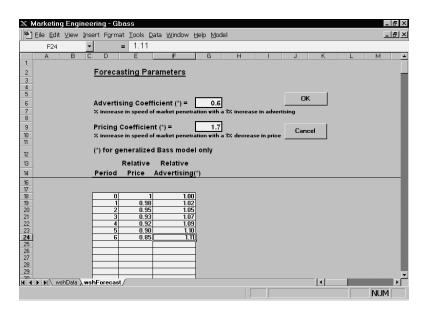
Forecasting analysis

Select **Forecasting analysis** and click **OK**.
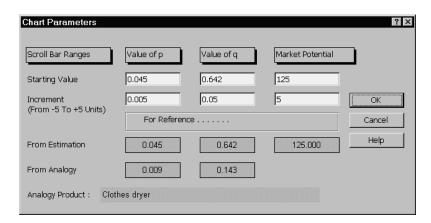You will see the following box:

Enter values for the **Number of Forecast Periods**, the **Market Growth Rate**, and the **Market Potential Price Elasticity** (that is, the percent increase in market potential with a one percent decrease in price). Click **OK**.

Next specify the expected course of price evolution. For the generalized Bass model you must also specify the evolution of advertising effort. The relative price and relative advertising values are indices with

respect to their values in forecasting period 0, the last period for which actual data are available.
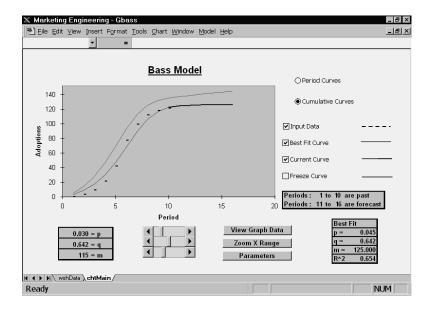
Click **OK** to get to the next screen.

For charting and visual tracking, specify the starting values and increments for the coefficients p and q, and for the market potential. If available, the program displays values from both the best-fit estimates (estimation of p and q on past data) and the reference case (selection of p and q by analogy).

NOTE: The parameters can be varied by up to five increments on either side of the **Starting Value**.

Click **OK** to see the estimated adoptions curves.

You can view the adoption curves either period by period (**Period Curves**) or cumulatively (**Cumulative Curves**). Check the box to the left of **Current Curve** to see the forecasted adoption pattern for the values of p, q, and m shown in the left bottom part of the screen. By using the scroll bars under **Period** you can adjust the parameters for the **Current Curve** and observe how changes in the parameter values affect the shape of the graph.
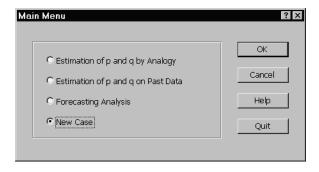
Once you think you have a reasonably good match between your **Input Data** (if available) and the **Current Curve**, you can freeze this curve as a benchmark. All the parameter values for **Freeze Curve** are displayed in the area in the lower right corner. Compare them to the parameter values for the **Best Fit Curve** if you checked that option. You can continue to change the shape of the **Current Curve**, and the **Freeze Curve** will remain fixed.

Clicking the **View Graph Data** button brings up a worksheet listing the raw data. You can only view and not change the data in this data sheet.
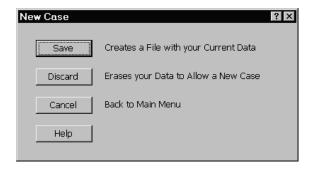
Clicking **Zoom X Range** allows you to limit the number of periods for which the data are plotted.

Clicking the **Parameters** button brings back the display **Chart Parameters**. You can enter new values for the starting points of the coefficients and increments.

If you want to try another case or another analogy, modify your data points, or save the current case, you need to bring up the **Main Menu**. To do so, go to the **Model** menu and choose **Main Menu**.

If you want to analyze another product, select **New Case** from the **Model Menu**. Decide whether you want to **Save** or **Discard** the current scenario.

New Case

Save	Creates a File with your Current Data
Discard	Erases your Data to Allow a New Case
Cancel	Back to Main Menu
Help	

Appendix B

Zenith High Definition Television (HDTV) Case[3]

On August 1, 1990, Jerry Pearlman, CEO of Zenith Electronics Corporation, met with Bruce Huber, VP of marketing, to discuss the market potential for a new technology called high definition TV (HDTV). At the end of the meeting, Mr. Pearlman asked Mr. Huber to develop, within a month, a preliminary forecast of demand for HDTV sets for a 15-year period starting in 1992. Although they both realized that any forecasts they came up with would just be best guesses they still felt that forecasts would be useful in deciding whether and how the company should respond to this emerging technology. Many strategic decisions would depend on these forecasts, including the level and nature of the R&D and marketing research activities the company would undertake, the strategic alliances it would pursue to get a running start in the marketplace, and the extent of its participation in industrywide lobbying efforts with the FCC (Federal Communications Commission) and the US Congress.

HDTV background

As compared to conventional TV sets, HDTV sets produce better quality pictures with higher resolution and superior sound (CD-like). They also have wider screens. According to the Electronic Industries Association, high definition in TV can be measured by the resolution of the picture, that is, the number of horizontal and vertical lines scanned on the TV screen.

To promote the growth of HDTV several stakeholders would have to adopt a common set of standards:

- Technical specifications for the core functions and manufacture of HDTV sets
- Production standards to enable TV and movie studios to develop content to take advantage of the superior display features of HDTV
- Broadcast and transmission standards regulated by the FCC to ensure high quality transmission within the available frequency spectrum

The Japanese government and industry adopted an HDTV standard in 1984 that had 1125 lines per frame, while the US NTSC (National Television Standards Committee) standard is 525 lines per frame. In addition the US NTSC standard has a 4:3 (or 16:12) aspect ratio (ratio of frame width to height) but the committee is considering a wide-screen aspect ratio of 16:9 for HDTV. Movies made after 1950 typically used

---

[3] This is based on Harvard Business School case 5-591-025 and is used here with the permission of HBS Publishing Division.

wide-screen formats although not always with a 16:9 aspect ratio, while TV programs and most movies made before 1950 typically used a 16:12 aspect ratio.
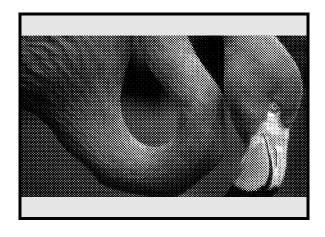
The Japanese standard relied on traditional analog signals for broadcasts, but the transmission was only over satellite channels. Unless consumers had both an HDTV and a way to receive satellite signals, they would not be able to receive these programs.

In 1990, US industry and government were still working together on setting standards. They had to resolve several thorny issues:
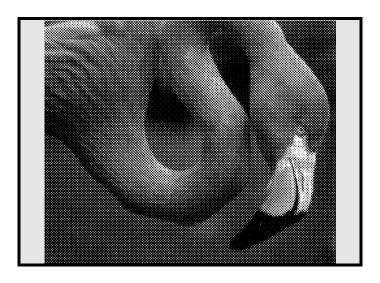
Compatibility with existing TVs: The FCC wanted to ensure that whatever transmission standard the industry adopted for HDTV it would not make existing TV sets obsolete. Even with compatibility ensured, an HDTV program would leave the top and bottom of the screen empty when displayed on a standard TV set (Exhibit B1a). On the other hand, when receiving a standard-broadcast TV program, an HDTV would display a squarish picture in the middle of a wide rectangle (Exhibit B1b).

Digital versus analog standard: Several US firms including Zenith were pushing for adoption of digital standards instead of the analog standard the Japanese had adopted. Under a digital standard all images would be converted to the 0/1 language of computers and compressed before being transmitted by cable, satellite, or over the air. The TV receiver would convert the digital streams back into images.

Although a digital standard seemed to be better aligned with the expected convergence of computer and telecommunication technologies, industry members had several concerns. Analog signals typically degenerate gracefully under interference, i.e., a small loss of signal quality results in only a small loss of picture quality. Digital signals however tend to degrade substantially with a small impairment to the signal quality. This may not be a major problem for cable-based transmission. Also people have had a lot of experience with analog transmission. A digital transmission standard could require experimentation and testing over several years before adoption.

(a)

(b)

Exhibit B1
(a) HDTV broadcast as it appears on standard TV, and
(b) standard NTSC broadcast as it will appear on HDTV.

Regardless of whether the industry adopts a digital or analog transmission standard, content providers, such as TV and movie studios would have to invest in costly equipment to produce images with higher resolution. For example, studios would either need high-definition digital cameras for shooting or equipment to convert images from a high-resolution format, such as 35mm film. A studio-quality camera would cost around $300,000 to $400,000. Production staff at TV studios would also have to adapt to the new wide-screen-aspect ratio. They would have to learn new techniques for composing scenes, editing frames, and so forth. At the same time, broadcasters (e.g., TV stations and cable TV companies)

would have to invest heavily in such equipment as transmitters and towers to broadcast HDTV signals.

Zenith HDTV efforts to date

In 1990, Zenith was working to develop advanced flat-screen picture tubes that could display images in the HDTV format. The development efforts looked promising, so Zenith anticipated marketing 20" and wider screens by 1992. In addition Zenith and its partner, AT&T, had made significant advances in developing a "spectrum compatible" HDTV transmission system that would offer HDTV pictures in the same channel space as existing NTSC standards. (Because of the scarcity of channel bandwidth such a system was considered to be a necessary element in the introduction of HDTV.)

The TV market

Zenith had conducted a number of studies of consumer behavior, which led to the following general conclusions:

- Consumers looked for value for their money and stayed within their budgets. Most consumers were satisfied with their existing TVs.
- Product quality was the most important criterion for evaluating brands. Consumers generally preferred large screens to small screens and considered such product features as stereo, remote control, and style to be important as well.
- Consumers tended to shy away from the lowest priced brands because they were suspicious of poor quality.

Bruce Huber had access to several additional sources of data acquired by Zenith's marketing research department. In particular he thought the data shown in Exhibits B2 to B7 might be useful in forecasting the sales of HDTV sets.

	Size	% units	Average retail price
Small	<19"	42%	$290
Medium	20-25"	40%	$610
Large	27+"	15%	$1,050

Exhibit B2
Breakdown of the TV set size distribution in 1989 and the corresponding average prices.

	TV households	Multi set	Color TV	Cable	VCR	Remote control
1950	10%	–	–	–	–	–
1955	67	4%	–	–	–	–
1960	87	12	–	–	–	–
1965	94	22	7%	–	–	–
1970	96	35	41	7%	–	–
1975	97	43	74	12	–	–
1980	98	50	83	20	–	–
1985	98	57	91	43	14%	29%
1989	98	63	97	53	60	72
1990	98	65	98	56	66	77

Note: Nielsen estimated U.S. TV households = 92.1 million on Jan. 1, 1990.

Exhibit B3
Data on the market's time pattern for adoption of past TV-related technologies. *Source*: *The American Enterprise*, 1990, p. 97.

Year	Total units	Total $	Average $/unit	Total $ in 1989 $[*]	Avg. $/unit in 1989 $[*]
1971	11,197	$2,551,997	$228	$7,831,740	$698
1975	11,606	2,684,121	231	6,184,102	533
1980	18,143	4,798,239	264	7,220,650	398
1985	20,829	5,871,854	282	6,766,820	325
1989	24,669	6,899,762	280	6,899,761	280

[*] Adjusted for the Consumer Price Index

Exhibit B4
Summary of factory shipments of TVs in the U.S. since 1971. *Source*: EIA Electronic Fact Books 1981-1989.

Buyer type

Performance or feature	36%
Experience	34%
Price	30%

Note: Performance or feature-oriented buyers consider primarily the performance and the features of the set when making a TV purchase;

Experience-oriented buyers want technology they can trust, i.e., technology that is stable and has been widely used, before they adopt;

Price-oriented buyers base their purchases primarily on the price of the product.

Exhibit B5
Summary of the results of a market segmentation study of TV buyers conducted by Zenith.

	1989	1990	1991	1992	1993	1994
Color TV forecast (Econometric model)	22.0	22.2	23.4	24.9	25.7	25.9
—Units—						
First purchase	2.1	1.8	1.6	1.6	1.5	1.5
Replacement	7.7	8.3	8.9	9.6	10.3	11.0
Additional	11.6	11.5	12.3	13.0	13.2	12.7
Institutional	0.6	0.6	0.6	0.7	0.7	0.7

Exhibit B6
Zenith's forecast sales of color TVs by purchase occasion (millions of units).

	1992	1993	1994	1995	1996	1997	1998	1999	2000
Industry total (millions of units)	21.4	21.9	22.4	22.9	23.5	24.1	24.7	25.2	25.9
25" and larger (millions of units)	6.0	6.1	6.2	6.4	6.8	7.2	7.5	8.0	8.5
Zenith retail price for HDTV									
26"/31"	$2500	$2000	$1700	$1500	$1400	$1350	$1300	$1300	$1300
22"/27"				1100	1000	900	900	900	
Zenith retail price with conventional tube									
26"/31"	$3000	$2500	$2100	$1900	$1700	$1550	$1550	$1550	$1500
22"/27"				1200	1100	1000	1000	1000	

Exhibit B7
Zenith's forecasts of U.S. sales of large screen TVs, which have price points that are likely to be similar to those of the HDTV.

Forecasts of HDTV sales

A few months earlier, the Electronic Industries Association (EIA) had forecast that HDTV would penetrate 25 percent of US households by the year 2000. Jerry Pearlman was not that optimistic but still predicted that HDTV would garner about 10 percent of the TV industry sales by 1999.

Some industry observers believed that both of these forecasts were optimistic because picture quality alone won't sell HDTV sets without significant levels of HDTV programming and broadcasting. They believed that the projected levels of penetration would occur only if (1) the FCC

settled on a transmission standard immediately, a highly unlikely prospect, and if (2) broadcasters invested substantial amounts of money in new equipment, which is unlikely before studios produce the content for HDTV broadcasting. There are about 1500 TV stations in the country, each of which would have to incur equipment costs of between $2 and 3 million to upgrade to digital transmission. These observers thought that neither of these scenarios was likely to occur for several years and that by the year 2000, sales would perhaps reach "a few hundred thousand units." Until then, HDTV would be used mostly for viewing closed-circuit TV programs, such as training films (e.g., surgery demonstrations), or for home-viewing of rented or owned movies on high-end entertainment systems.

With this preliminary research behind him, Bruce Huber was ready to tackle "the HDTV forecasting problem." He had recently acquired software called GBass for forecasting new-product sales. He wondered whether this software would be of any help in this forecasting task.

EXERCISES

1. Summarize and justify alternative scenarios (i.e., consistent sets of assumptions) ranging from pessimistic to optimistic with regard to market performance of HDTV.

2. Develop forecasts of HDTV penetration in the US market from 1992 through 2006 for each scenario you develop. Justify and explain your forecasts.

Appendix C

How we Obtained the Estimates in Exhibits 1a and 1b[4]

Exhibits 1a and 1b summarize the *p*, *q* and *m* estimates of the basic Bass model for 54 innovations. We had to make several choices about the data to use and the estimation method in obtaining these estimates. In this Appendix, we explain our choices and also describe other alternatives available for estimating model parameters.

Penetration versus unit sales data:

Penetration is the percentage of a target population (e.g., all U.S. households) that has adopted a product by a particular time. Penetration data closely captures the spirit of the diffusion phenomenon being modeled, because they measure only adoptions of a product or technology, and ignore multiple purchases and replacement sales. Unit sales data, on the other hand, incorporate ***all*** purchases made by individuals or households. For durables with long inter-purchase cycles, penetration data and sales data are equivalent in the early phases of a product's life cycle, and the parameter estimates using either type of data should produce similar estimates.

Penetration data are obtained through periodic "ownership" surveys of samples of households, whereas sales data are obtained from a near census of factory shipments or retail sales. Penetration data are likely to have more measurement error (on a percentage basis) than sales data, and as a result, also exhibit greater variation from one period to the next. Thus, the standard errors of the parameter estimates are likely to be larger for penetration data, even though the estimates themselves are consistent. On the other hand, parameter estimates obtained from sales data may have higher reliability, but only if they do not include significant levels of replacement sales or multiple purchases. If they do, then the estimates will vary over (Putsis 1998) and possibly be inconsistent as well.

Managerially, sales data are more meaningful than penetration data. Managers evaluate performance in terms of sales, and are often rewarded for meeting certain sales goals. This suggests a need for more sophisticated diffusion models based on sales data that incorporate first purchase, replacement sales, as well as multiple purchases (see for example, Hahn, Park, Krishnamurthi, and Zoltners, 1994; Kamakura and Balasubramanian, 1987;Ratchford, Balasubramanian, and Kamakura 1999).

On balance, we decided to use penetration data, and the parameter estimates given in Exhibits 1a and 1b are based on such data. An important reason for our choice is that for products introduced in the recent past (e.g., answering machines, PDA), multiple purchases and replacement purchases have tended to occur sooner, which makes penetration data more appropriate for estimating model parameters.

Short versus long data series

A long data series provides a more complete picture of the evolution of an innovation, and

---

[4] We are indebted to Professor Frank Bass for helping us clarify the various issues described in this Appendix.

minimizes the bias in the estimates. Specifically, a longer series helps avoid underestimating *m* and overestimating *q* (cf. Debecker and Modis 1994; Gillis and Ratkowsky 1978; Hardie, Fader and Wisniewski 1998; Van den Bulte and Lilien 1997).

At the same time, one can make good arguments for using shorter data series. Of particular concern is the possibility that the process generating the underlying data may change over time. Later adoptions may occur because of enhancements to a product or technology. An example in the home PC market is the introduction of the Windows operating system in the 1980s, which encouraged both replacement sales (from DOS) and adoptions by newer segments. If we lack the data to appropriately control for such dramatic product enhancements (e.g., in price or performance), we may want to limit our parameter estimates to the very first product generation. In fact, some managers may be interested in predicting just the initial peak in demand, rather than the global peak that can occur several years down the road. Peak sales prediction is useful for capacity planning.

We opted to use the long data series to minimize bias in our estimates included as default value in our software, assuming that the Bass model captures the diffusion phenomenon throughout the entire life cycle of a product. Users who believe that shorter data series may be more appropriate for a particular application, can select and input estimates from Exhibit 1b. Also, in our software, we provide access to all our data, available in the Excel file analogy_data.xls. Thus, advanced users can select any appropriate subset of the data, estimate model parameters using the selected data, and then incorporate the estimated parameters to predict the diffusion of analogous products.

Parameter Estimation Method

We obtained parameter estimates from equation 1 using nonlinear least squares, imposing the constraints that all parameters are non-negative, and that $m \leq 100\%$, when the model was estimated using penetration data. We chose to estimate equation 1 rather than other operationalizations of the Bass model to avoid having to use a specific date for the launch of the product, an issue about which scholars disagree for many products listed in Exhibits 1a and 1b.

The average values reported for *p* and *q* in Exhibit 1a (0.037 and 0.327) and in Exhibit 1b (0.040 and 0.398) are close to the (0.03 and 0.38) average values reported by Sultan et al. (1990). Still, some of the entries in Exhibit 1a may look odd to readers familiar with other published estimates. Specifically, as Professor Bass has brought to our attention, some of our estimates for *p* and *q* seem low compared to estimates obtained in previous studies analyzing the same products. For room air conditioners for instance, Exhibit 1a indicates $p = 0.006$ and $q = 0.185$, while Bass, Krishnan and Jain (1994) report $p = 0.009$ and $q = 0.380$. A similar pattern exists for clothes dryers: Exhibit 1a indicates $p = 0.009$ and $q = 0.143$, whereas Bass, Krishnan and Jain (1994) report $p = 0.013$ and $q = 0.332$.

The differences between our estimates and those reported by Bass, Krishnan, and Jain (1994) could occur due to a number of reasons:

1. We use penetration data whereas they use sales data.
2. We use longer time series than their data
3. We estimate equation (1), i.e. the discrete-time version of the Bass model specified in

the adoption domain, whereas they estimate the continuous-time version of the Bass model specified in the time domain (cf. Srinivasan and Mason 1986).

To get some insights about the factors that may be most responsible for differences in parameter estimates, we summarize below our exploratory assessment of parameter estimates for room air conditioners and clothes dryers. We estimated four different models for these two products (all estimation was performed using nonlinear least squares implemented in the MODEL procedure in SAS):

1. Sales data, short time series, time domain. This choice of data type, time series length and model variant is used by Bass, Krishnan and Jain (1994).
2. Penetration data, short time series, time domain.
3. Penetration data, long time series, time domain.
4. Penetration data, long time series, adoption domain. This choice of data type, time series length and model variant is identical to the ones used to construct Exhibit 1a.

Comparing the estimates from analyses 1 and 2 provides some information on the impact of using sales rather than penetration data for short time series. Comparing 2 with 3 provides information on the impact of using short rather than long data series. Comparing 3 with 4, finally, provides information on the impact of formulating the Bass model in the time or adoption domain.

Exhibit C1 summarizes the parameter estimates for these four models. While we performed the exercise for only two products and the results only have heuristic value, the patterns are remarkably similar for the two products and in agreement with prior research (Schmittlein and Mahajan 1982; Van den Bulte and Lilien 1997). The divergence between the results in Exhibit 1 and those reported by Bass et al. (1994) stems to a very large degree from the difference in time series length. Using sales versus penetration data (provided the series are short) hardly affects the estimates for p and q, nor does formulating the model in the time versus adoption domain. The pattern that q can decrease rather dramatically as one increases the length of the data series has been documented earlier by Van den Bulte and Lilien (1997), while the (quite minor) increase in p observed when one switches from the time to the adoption domain is in agreement with both the time aggregation bias argument and the empirical findings presented by Schmittlein and Mahajan (1982).

Product	Data type	Time series	Model	p	q	m
Air conditioners	Sales	1949-61	Time domain	.009	.375	18711
Air conditioners	Penetration	1950-61	Time domain	.011	.322	23.46%
Air conditioners	Penetration	1950-79	Time domain	.005	.195	59.65%
Air conditioners	Penetration	1950-79	Adoption domain	.006	.185	60.60%
Clothes dryers	Sales	1949-61	Time domain	.014	.327	16497
Clothes dryers	Penetration	1949-61	Time domain	.017	.345	24.37%
Clothes dryers	Penetration	1949-79	Time domain	.008	.145	70.40%
Clothes dryers	Penetration	1949-79	Adoption domain	.009	.143	70.06%

Exhibit C1
The impact of data type, time series length, and model specification on the value of parameter estimates.

Another way to compare the p and q estimates, is to see how well the predicted time to peak adoptions implied by their value, which can be computed as $\ln(q/p) / (p+q)$, corresponds to the actual time of peak adoptions. For room air conditioners, the estimates by Bass et al. suggest the peak occurred 9.5 years after launch. Assuming that air conditioners were launched in 1949 (when their data start), this suggests that adoptions peaked in 1959. Our estimates, in contrast, suggest a peak time of 18 years, i.e. adoptions peaking around 1967. So which peak time is in best agreement with the data? The answer depends on which data and time window one uses. If we limit ourselves to the 1949-1961 period, then 1959 indeed appears to have been the year of peak sales (Exhibit C2). However, when we consider the product's entire diffusion history, then we must conclude that 1959 was only a minor local peak; the true peak time in the latter half of the 1960s, as suggested by our estimates. A similar pattern exists for clothes dryers (Exhibit C3): the true peak occurred not in the mid 1950s, as predicted by the results published by Bass et al. but in the late 1960s, as suggested by the parameter values presented in Exhibit 1. Using long data series helps separate these local and global peaks.
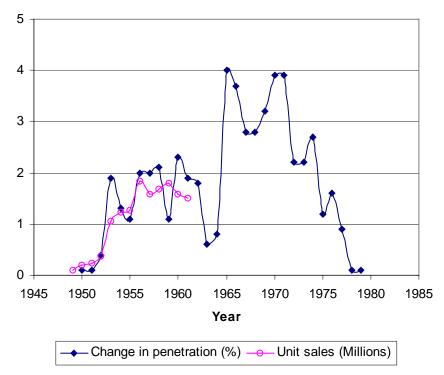
Exhibit C2
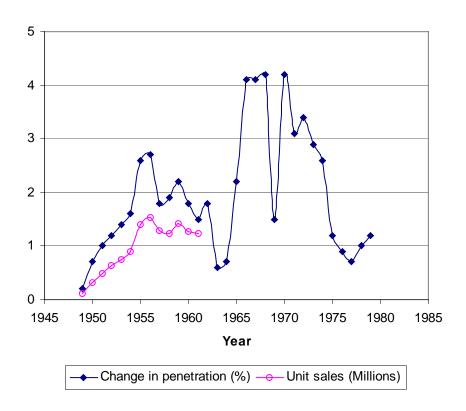The adoption path of room air conditioners: Changes in household penetration versus unit sales.

Exhibit C3
The adoption path of clothes dryers: Changes in household penetration versus unit sales

In addition to the factors described above, other factors can also impact the validity and reliability of parameter estimates. For example, sometimes new products are available for several years before penetration data are collected for those products -- although Cable TV was introduced in the 1970s, we were able to get data only from 1981 onwards. Such left-censored data may inflate the estimates of the p parameter and deflate the estimates of q.

Overall, while for some products our estimates reported in Exhibit 1a and implemented in our software are at variance with some previous analyses, the differences stem primarily from our use of more data than were available at the time some of those previous studies were published (e.g., Bass 1969).

References

Bass, Frank M. (1969) "A New Product Growth Model for Consumer Durables," *Management Science*, 16 (January), 215-227.

Bass, Frank M., Trichy V. Krishnan and Dipak Jain (1994) "Why the Bass Model Fits Without Decision Variables," *Marketing Science*, 13 (Summer), 204-223.

Bayus, Barry L. (1987) "Forecasting Sales of New Contingent Products: An Application to the Compact Disc Market," *Journal of Product Innovation Management*, 4 (December), 243-255

Choffray, Jean-Marie and Gary L. Lilien (1986) "A Decision Support System for Evaluating Sales Prospects and Launch Strategies for New Products," *Industrial Marketing Management*, 15 (February), 75-85.

Debecker, A. and T. Modis (1994) "Determination of the Uncertainties in S-Curve Logistic Fits," *Technological Forecasting and Social Change*, 46, 153-173.

Eliashberg, Jehoshua and Mohan Sawhney (1994) "Modeling Goes to Hollywood: Predicting Individual Differences in Movie Engagement," *Management Science*, 40 (September), 1151-1173.

Gillis, P. R. and D. A. Ratkowsky (1978) "The Behaviour of Estimators of the Parameters of Various Yield-Density Relationships," *Biometrics*, 34 (June), 191-198.

Givon, Moshe, Vijay Mahajan and Eitan Muller (1995) "Software Piracy: Estimation of Lost Sales and the Impact on Software Diffusion," *Journal of Marketing*, 59 (Winter), 29-37.

Hahn, Minhi, Sehoon Park, Lakshman Krishnamurthi, and Andris A. Zoltners (1994), "Analysis of New Product Diffusion Using a Four-Segment Trial-Repeat Model," *Marketing Science*, 13 (Summer), 224-247.

Hardie, Bruce G. S., Peter S. Fader and Michael Wisniewski (1998) "An Empirical Comparison of New Product Trial Forecasting Models," *Journal of Forecasting*, 17 (June), 109-229

Kalish, Shlomo (1985)   "A New Product Diffusion Model with Price, Advertising and Uncertainty," *Management Science,* 31 (December), 1569-1585.

Kalish, Shlomo and Gary L. Lilien (1986a)   "A Market Entry Timing Model for New Technologies," *Management Science*, 32 (February), 194-205.

Kalish, Shlomo and Gary L. Lilien (1986b) "Applications of Innovation Diffusion Models in Marketing" in Vijay Mahajan and Yoram Wind, eds, *Innovation Diffusion Models of New Product Acceptance*, Ballinger: Cambridge, MA, 235-279.

Kamakura, Wagner A. and Siva K. Balasubramanian (1987), "Long-term Forecasting with Innovation Diffusion Models: The Impact of Replacement Purchases," *Journal of Forecasting*, 6 (January-March), 1-19.

Lattin, James M. and John H. Roberts (1998) "Calibrating an Individual Level Diffusion Model Prior to Launch," working paper, Graduate School of Business, Stanford University, October.

Lenk, Peter J. and Ambar Rao (1990) "New Products from Old: Forecasting Product Adoption by Hierarchical Bayes Procedures," *Marketing Science,* 9 (Winter), 42-57.

Lilien, Gary L. and Arvind Rangaswamy (1998) *Marketing Engineering*, Addison-Wesley, Reading, MA.

Mahajan, Vijay, S. Sharma and R. D. Buzzell (1993) "Assessing the Impact of Competitive Entry on Market Expansion and Incumbent Sales," *Journal of Marketing*, 57 (July), 39-52.

Molyneux, Phil and Nidal Shamroukh (1996) "Diffusion of Financial Innovations: The Case of Junk Bonds and Note Issuance Facilities," *Journal of Money, Credit, and Banking*, 28 (August, part 2), 502-522.

Rao, Vithala and Joel Steckel (1998) *Analysis for Strategic Marketing*, Addison-Wesley, Reading, MA.

Putsis, William P., Jr. (1998) "Parameter Variation and New Product Diffusion," *Journal of forecasting*, 17, 231-257.

Ratchford, Brian T., Siva K. Balasubramanian, and Wagner A. Kamakura (1999), "Diffusion Models with Replacement and Multiple Purchases," in *New Product Diffusion Models*, Vijay Mahajan, Eitan Muller and Jerry Wind (Editors), Sage Press 2000.

Redmond, William H. (1996) "Product Disadoption: Quitting Smoking as a Diffusion Process," *Journal of Public Policy and Marketin*g, 15 (Spring), 87-97.

Roberts, John H. and Glen L. Urban (1988) "Modeling Multi-Attribute Utility, Risk and Belief Dynamics for New Consumer Durable Brand Choice," *Management Science*, 34 (February), 167-185.

Sawhney, Mohan and Jehoshua Eliashberg (1996) "A Parsimonious Model for Forecasting Gross Box Office Revenues of Motion Pictures," *Marketing Science*, 15 (2), 113-131.

Schmittlein, David and Vijay Mahajan (1982) "Maximum Likelihood Estimation for an Innovation Diffusion Model of New Product acceptance," *Marketing Science*, 1 (Winter), 57-78.

Sillup, George P. (1992) "Forecasting the Adoption of New Medical Technology Using the Bass Model," *Journal of Health Care Marketing*; 12 (December), 42-51.

Srinivasan, V. and Charlotte H. Mason (1986) "Nonlinear Least Squares Estimation of New Product Diffusion Models," *Marketing Science*, 15 (4), 169-178.

Steffens, Paul (1998) "Applying Diffusion Models with Regional Heterogenity*," Marketing Letters*, 9 (4), 361-370.

Sultan, Fareena, John U. Farley and Donald R. Lehmann (1990) "A Meta Analysis of Applications of Diffusion Models," *Journal of Marketing Research*, 27 (February), 70-77.

Thomas, R. J. (1985) "Estimating Market Growth for New Products: An Analogical Diffusion Models Approach," *Journal of Product Innovation Management*, 2 (March), 45-55.

Trajtenberg, Manuel and Shlomo Yitzhaki (1989) "The Diffusion of Innovation: A Methodological Reappraisal," *Journal of Business and Economic Statistics*, 7 (January), 35-47

Urban, Glen L., Bruce D. Weinberg, and John R. Hauser (1996) "Premarket Forecasting of really-New Products," *Journal of Marketing*, 60 (January), 47-60.

Urban, Glen L., John S. Hulland, and Bruce D. Weinberg (1993) "Premarket Forecasting for New Consumer Durable Goods: Modeling Categorization, Elimination, and Consideration Phenomena," *Journal of Marketing*, 57 (April), 47-63.

Van den Bulte, Christophe and Gary L. Lilien (1997) "Bias and Systematic Change in the Parameter Estimates of Macro Level Diffusion Models," *Marketing Science,* 16 (4), 338-354.

Virolleaud, P. (1983) "Les ventes des trois premières années prévisibles à 30%," *Usine Nouvelle*, June.

Weerahandi, S. and S. R. Dalal (1992) "A Choice-Based Approach to the Diffusion of a Service: Forecasting Fax Penetration by Market Segments," *Marketing Science,* 11 (1), 39-53.

Zufryden, Fred S. (1996) "Linking Advertising to Box Office Performance of New Film Releases—A Marketing Planning Model," *Journal of Advertising Research,* 36 (July/August), 29-41.