EMPIRICAL GENERALIZATIONS AND MARKETING SCIENCE: A PERSONAL VIEW

FRANK M. BASS
The University of Texas at Dallas

Marketing has matured to the point where it seems desirable to take stock of where we are, what we have learned, and fruitful directions for extending the knowledge base that has developed. Science is a process involving the interaction between empirical generalizations and theory. An empirical generalization is "a pattern or regularity that repeats over different circumstances and that can be described simply by mathematical, graphic, or symbolic methods." One of the purposes of the Empirical Generalizations Conference held at Wharton on February 16–18, 1994 was to develop a list of examples of such empirical generalizations in marketing. Empirical generalization can precede a theory to explain it or it can be predicted by a theory. Science is the process of interaction between theory and data that leads to higher level theories. Examples are provided here of empirical generalizations in marketing and their theoretical counterparts. One example is provided of a higher level theory.

(Diffusion; Brand Choice; Pricing Research; Empirical Generalizations)

1. Introduction

Over the past 30 years or so the volume of serious scholarly research in marketing has expanded tremendously. Modelling activity, in particular, has grown to the point where there is now a critical mass of scholars who pursue marketing issues through marketing models. One increasingly finds the word "science" used in conjunction with the word "marketing." Several companies have established marketing science groups. There is a Marketing Science Institute and a scholarly journal, Marketing Science. A good portion of the critical mass attend the annual Marketing Science Conference that is sponsored by the College on Marketing of The Institute of Management Sciences (TIMS). (TIMS merged with the Operations Research Society of America (ORSA) in January 1995 to form the Institute for Operations Research and the Management Sciences (INFORMS).)

The field has matured to the point where it seems desirable to take stock of where we are, what we have learned, and fruitful directions for extending the knowledge base that has developed. One such effort along these lines is to be found in the discussion by Bass (1993). We need to understand not only what we know, but also to develop, if possible, some unifying principles that will guide our activities in the future.

Because of the close connection between science and empirical generalization (phenomena), Jerry Wind and I organized a workshop of leading marketing scholars at the Wharton School (February 16–18, 1994) on "Empirical Generalizations in Marketing." The topic and the underlying philosophies of the authors are diverse, but it is our hope
that unifying principles will emerge that will be useful for further development of marketing science.

This introductory piece stems from the keynote address given at the Wharton conference. In it I shall give my own views, along with examples, of the major issues involved in further development of marketing science as related to empirical generalizations. Neither the views nor the examples will find universal acceptance as correct or most appropriate, but I do hope they will be useful in structuring thought and discussion.

Readers will find extensive reference here, and elsewhere in this volume, to the works of Ehrenberg. There is good reason for this. He was the earliest (dating from the 1950s), most consistent, and most prolific advocate of the “empirical generalization approach” to the development of knowledge about relationships involved in the study of marketing. Many will regard his views on the topic as extreme, but when juxtaposed to the “a priori theory” philosophy commonly associated with economics and other social sciences, they are likely to continue to have a profound impact on the thinking of marketing scientists. A discussion of the Ehrenberg philosophy and examples of empirical generalizations in marketing from his viewpoint are to be found in this volume. (See Ehrenberg 1995.)

2. Empirical Generalizations in Marketing: What Are They?

2.1. Definition and Schema

I would propose as a definition of an empirical generalization that it is “a pattern or regularity that repeats over different circumstances and that can be described simply by mathematical, graphic, or symbolic methods.” The definition does not assert causality and it does not require that the values of the parameters governing the regularity be invariant over the different circumstances. It does require that there be a pattern, but it does not require that the pattern be universal over all circumstances. Thus, the negative binomial distribution may usually describe the distribution of the number of purchase occasions of a brand or product category, but the parameter values may vary, and there may be circumstances and conditions under which the generalization will fail to hold.

It was my intention in constructing the definition to exclude verbal statements unless the verbal description can be given some measure of enhanced quantification. For example, I would tend to exclude statements such as “y increases with x,” but would include “y increases exponentially with x.” Others may well disagree with this restriction in the definition, but few would disagree with the proposition that more precise generalizations are superior to less precise ones.

Ehrenberg (1982) has noted that:

... the lawlike relationships of science are descriptive generalizations, often at quite a low level. But the variables which do not appear in the equation greatly aid our understanding (e.g., that the type of gas, the type of apparatus, etc. do not matter). They are also the building-blocks of higher level theory and explanation.

Ehrenberg (1975) also has characterized lawlike relationships as having the following properties:

- They are of limited generality, rather than universal; they are approximate rather than exact; they are not necessarily derived from theory and they are broadly descriptive rather than directly causal.

The fact that empirical generalizations are only approximate suggests that conventional tests (such as tests of statistical significance) cannot be employed in their evaluation. Simon (1968) wrote:

If the generalization is just that—an approximate summary of the data then it is certainly not falsifiable, or testable. It become falsifiable or testable when (a) it is extended beyond the data from which it was generated, or (b) an explanatory theory is constructed, from which the generalization can be derived, and the explanatory theory has testable consequences beyond the original data.
There is an often-quoted paradoxical statement about generalizations that: “All generalizations are false, including this one.” The statement implies that there are conditions under which a generalization will fail to hold. In marketing, as elsewhere, empirical generalizations will recur, but there will exist conditions under which the regularity will vanish. Discovery of those conditions leads to “higher level” understanding. Barwise (1995) provides in this volume a discussion of “Good Empirical Generalizations,” including a discussion of higher level theory as related to necessary conditions.

Table 1 shows a scheme that summarizes the interrelationships between empirical generalizations and explanation leading to higher level theory and understanding.

The scheme leading to the development of general theories shown in Table 1 is consistent with the discussion by Hunt (1983) based on philosophy of science considerations as applied to marketing theory.

2.2. Examples: Diffusion

I conclude this section with a discussion of some examples of empirical generalizations in marketing. Figure 1 shows the number of households adopting VCRs from 1980 to 1989. The curve rises to a peak in 1985 and then declines. Figure 2 shows similar data from the 1960s for the adoption of color television sets. In Figure 2 the fitted Bass Model (Bass 1969) is shown. The adoption patterns for new technologies or services, whether by household consumers or industrial and commercial buyers, usually look like those shown in Figures 1 and 2. These curves are well described by the Bass Model. In other words, the central proposition in this model is that: The conditional probability that an adoption will be made at time $T$ given that an adoption has not yet been made is a linear function of the number of previous adopters. The mathematical statement of these words is that:

$$P(T) = p + (q/m)Y(T),$$

where $P(T)$ is the fraction of those remaining to adopt who adopt at time $T$, $p$ and $q/m$ are constants, and $Y(T)$ is the number of previous adopters. This simple equation leads directly to a nonlinear differential equation that has a closed-form solution that describes the time pattern of adoption of innovations. Thus, the diffusion pattern for technological innovations can be described graphically, as in Figures 1 and 2, and an equivalent mathematical description can be provided. Although the indicated diffusion pattern does not always prevail, it usually does, and it is appropriate to refer to it as an empirical generalization. See Mahajan et al. (1995) in this issue for a discussion of other empirical generalizations that are related to or derive from Equation (1).

**TABLE 1**

<table>
<thead>
<tr>
<th>Empirical Generalization</th>
<th>Specific Theory or Explanation</th>
<th>General Higher Level Theory Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is it?</td>
<td>A pattern of results which recurs. Possibly, but not necessarily, predicted by prior theory</td>
<td>A story consistent with accumulated data</td>
</tr>
<tr>
<td>Method of Deriving Process</td>
<td>Who knows?/&quot;eyeballing&quot;/Meta Analysis/Pattern Recognition</td>
<td>Simple, Parsimonious Model Development</td>
</tr>
<tr>
<td>Purpose</td>
<td>Further Observation or Replication</td>
<td>Match to Past Data</td>
</tr>
<tr>
<td></td>
<td>Descriptive</td>
<td>&quot;Theorizing&quot;/Model Development</td>
</tr>
</tbody>
</table>

This content downloaded from 140.113.81.39 on Wed, 03 Oct 2018 05:54:21 UTC
All use subject to https://about.jstor.org/terms
2.3. Example 2: The Experience Curve

Figure 3 shows the prices of cellular telephones (actual and projected on the basis of continued exponential decline) between 1986 and 1998. The price pattern is one of approximate exponential decline. This is the usual pattern of prices for new technologies. Further examples of the pattern are shown in Figure 4 for color television, and in Figure
5 for 1-million-bit memory chips. Explanation for these price patterns may be derived from the "Experience Curve" theory of marginal costs for new technologies, but the essential point is that these patterns constitute an empirical generalization.

3. Marketing Science: What Is It?

Support for a claim that there is a marketing science must be based on the definition of science. Such definitions are to be found in the philosophy of science. Nagel (1961) has stated that "science seeks to provide generalized explanatory statements about dis-
parate types of phenomena and to provide critical tests for the relevance of the attempted explanations.” One may substitute “empirical generalizations” for “phenomena” in Nagel’s definition. Science, then, consists of (1) empirical generalizations and (2) generalized explanations of the empirical generalizations. But, as indicated by Simon, a generalized explanation may have consequences that extend beyond the data from which it was generated. In this case the generalized explanation becomes testable because the consequences may be examined and compared with data beyond the empirical generalization from which it was generated. A generalized explanation fosters new issues and questions and, possibly, new empirical generalizations. Marketing science, like science generally, is a process. The process involves empirical generalizations, generalized explanations, as well as testing and revision or extension of the generalized explanation.

There are several examples in marketing of a process, or stream of research, in which data and theory interact to produce generalizations concerning phenomena. One of the purposes of the Empirical Generalizations in Marketing Conference was to highlight these examples and to bring into focus the nature of the interaction between data and theory in generating knowledge.

Although marketing affords several examples of empirical generalizations and of theories to explain (or describe) these generalizations, there are few examples of higher level explanations. A higher level explanation will reduce as a special case to a lower level explanation, and it will possibly suggest the existence of additional phenomena. Just as Einstein generalized Newton by producing a more general theory that reduces approximately to Newtonian theory as a special case and by suggesting the existence of numerous additional phenomena, progress in marketing science requires the development of higher level theories.

4. Data First or Theory First?

If, as suggested, science is a process that involves empirical generalizations and generalized explanation, which of these should come first? Does it matter? There are those
with strong philosophical views on both sides of the issue. Although in science there are probably more instances in which a phenomenon was observed first and then a theory constructed to explain it, there are also cases in which the theory predicted the existence of a phenomenon before it was observed. In marketing, as well, there are instances in which observation of the phenomenon preceded the explanation, but there are also instances in which the theory predicted the existence of a phenomenon. Examples of the former include observation of "experience curve" effects in the costs of producing new technologies and the observation of skewed distributions of the number of purchase occasions of a product or brand that led to Ehrenberg's NBD model and Dirichlet extensions. An example of the latter includes the Bass Model that predicted the existence of the diffusion empirical generalization. Although prior to the publication of this model in 1969, there had been suggestions that diffusion followed a normal distribution, there was little or no theory behind the suggestion and no systematized scheme of documenting the pervasive pattern of the diffusion phenomenon that followed publication of the Bass Model (see Mahajan et al. 1990). Although there are philosophical differences over the issue of whether it is better for observation to precede theory or the other way around, there is general agreement that science is a process in which data and theory interact to produce higher level explanations.

5. ETET or TETE?

Ehrenberg (1994) has recognized and discussed the interaction between data and theory (i.e., science is a process). He contrasts the empirical-then-theoretical (EtT) approach with the theoretical-in-isolation (Til) approach. He suggests that, naturally, EtT is good and Til is bad. He does, however, indicate that both Til and EtT occur in the early stages of investigation of a topic, and then interaction begins. "Subsequently, science traditionally moves into its characteristic looped ETET . . . modes. This can again be characterized in two steps and a criticism:

Empirical-Theoretical-Empirical-Theoretical . . . or ETET . . . (1) Establish some empirically well-grounded theory (as in EtT). (2) Test the theory more widely (the second E), deduce new conjectural theory (more T), test that widely, and continue. Criticism: It takes time."

However, as previously indicated, Empirical does not always precede Theory, and thus Ehrenberg's schema excludes TETE. The distinction is important because the theory may have other empirical implications beyond the immediate empirical generalization. In discussing the development of the Dirichlet model extension of the NBD model for stationary zero-order markets (characterized by Ehrenberg as an example of the ETET process), he notes that he and his colleagues Chatfield and Goodhardt developed the Dirichlet extension in 1975 and 1976 (Chatfield and Goodhardt 1975 and Ehrenberg and Goodhardt 1976). He then states that "... also starting fractionally later by Bass and his colleagues (Bass et al. 1976) but not in the context of the known patterns of buyer behavior (following T-i-I route) and apparently still not grounded in a wider range of generalizable empirical findings."

Although it is possible to quarrel with Ehrenberg's characterization of the approach of Bass et al. as following a Til route, the point should be made that the phenomena observed by Ehrenberg and others in the domain of purchase incidence has equivalent implications in the domain of brand switching. Thus, if the work of Bass et al. is an example of Til, it is also an example of TETE. It produces implications for both purchase incidence measures and brand switching measures. The assumptions underlying the two models are fundamentally equivalent, but the point is that it is useful to recognize that there are equivalent phenomena.
6. ETET or TETE and Extensions

As just discussed, it is sometimes possible to extend a theory (or explanation of an empirical generalization) by maintaining essential assumptions and examining further implications of the theory with respect to different phenomena that reflect the same underlying behavior. In this way new empirical generalizations may be discovered (or at least new ways of looking at the original generalization).

Figure 6 shows the sales growth and decline of different generations of mainframe computers. Each generation captures the market of earlier generations by obtaining customers that disadopt the earlier generation in favor of the later one and by obtaining sales from those who would have adopted the earlier generation but instead adopted the later one. In addition, each generation may expand the market by extending possible applications. The relationships shown in Figure 6 are by no means unique. They have been shown to obtain for successive generations of numerous, industrial, consumer, and pharmaceutical products. For details see Norton and Bass (1987, 1992).

An earlier important and often-used model of substitution of successive generations of technology was developed by Fisher and Pry (1971). This model, however, is a market-share-only model for each generation and is not capable of predicting unit sales of the various generations. The Norton and Bass model may be easily converted to a share model and is thus more general than the Fisher-Pry model. Moreover, Norton and Bass (1987) have shown that the more general model when converted to shares produced better forecasts of shares than the Fisher-Pry model.

The relationships exhibited in Figure 6 constitute an empirical generalization. Figure 7 shows the substitution of disposable diapers for cloth diapers (daily usage rates), and Figure 8 shows the substitution of successive generations of recording media. Figure 9 shows the displacement of 5.25-inch disk drives with 3.5-inch drives.

![Figure 6. Actual, Fit, and Forecast of Computer Performance Units.](image-url)
The empirical generalization of sales decline and growth of successive generations of technology was discovered not by ET or by ETET, but by TETE. The relations shown in Figures 6 through 9 were discovered by a theoretical extension of the Bass Model.
This theoretical development predicted the pattern of behavior for successive generations of technology, and the empirical observations came later.

The theoretical equations used to predict sales patterns for three generations of technology are:

\[ S_{1,t} = F(t_1)m_1[1 - F(t_2)], \]  
\[ S_{2,t} = F(t_2)[m_2 + F(t_1)m_1][1 - F(t_3)], \]  
\[ S_{3,t} = F(t_3)[m_3 + F(t_2)[m_2 + F(t_1)m_1]], \]

where \( m_i \) = the incremental market potential for the \( i \)th generation, \( t_i \) = the time since the introduction of the \( i \)th generation, and \( F(t_i) \) is the cumulative distribution function of the Bass Model.

The underlying assumptions governing these equations is that the Bass Model obtains with respect to the adoption rate for each generation of technology and that the parameters of the adoption process are the same for each generation.

### 7. Higher Level Explanations

#### 7.1. Overview, Definition, and Discussion

Ehrenberg (1994) has expressed a preference for the analysis of “stationary markets.” His methods apply to such markets, but tend to disintegrate when applied to nonequilibrium conditions. It is true, as he says, that most markets are approximately stationary. It is also true that there are few higher level explanations (or theories) in marketing. Suppose that there were a generalized model that reduced as a special case to the Dirichlet model and that explained nonstationary markets as well as stationary markets. Would not that be a worthwhile model to have? Just as Einstein’s equations reduce as a special
case to Newton when things are operating at less than the speed of light, it would be nice
to have higher level theories in marketing. Of course, such theories are difficult to construct.
However, inasmuch as the major purpose of empirical generalizations in marketing project
has been to generate ideas that will be useful in helping to chart the future directions for
marketing science, bold ideas ought not to be rejected out of hand.

7.2. An Example of a Generalized Model in Marketing

The Bass Model describes an empirical generalization of the diffusion of innovations.
Curves such as those shown in Figures 1 and 2 do not always describe the adoption
process, but they usually do. However, the Bass Model is incomplete in that it does not
include decision variables such as price and advertising. Recently, Bass et al. (1994) have
developed a generalized version of the Bass Model that includes decision variables. This
model was developed with the following principles in mind:
1. The model should reduce as a special case to the Bass Model under commonly
observed conditions. In other words, the model should explain why the Bass Model fits
without including decision variables.
2. Because the Bass Model curve usually describes the diffusion process, it seems
reasonable to conjecture that a different set of prices and other decision variables other
than the observed ones would have produced a curve with a similar shape, but the curve
would have been shifted.
3. The model should track the irregular deviations of actual data from the smooth
curve of the Bass Model.
4. The model should maintain the essential carrythrough properties of the Bass Model.
That is, an increase in adoption today should increase adoption tomorrow through the
influence of imitation ("internal influence").
5. The model should yield a closed-form solution.
6. The model should be flexible and encompass a great variety of shapes.

The Generalized Bass Model (GBM) is developed by making the parameters $p$ and $q$
in Equation (1) functions of decision variables. Equation (1) may be written as:

$$f(T)/[1 - F(T)] = p + qF(T), \quad (5)$$

where $f$ is the density function of time to adoption and $F$ is the cumulative function,
and where $p$ and $q$ are fixed coefficients of "external influence" and "internal influence."

In the generalized model we have:

$$f(T)/[1 - F(T)] = x(T)[p + qF(T)], \quad (6)$$

where $x(T)$ depends on percentage changes in the decision variables. Specifically, we
have:

$$x(T) = 1 + \beta_1[\Delta Pr(T)/Pr(T - 1)] + \beta_2[\Delta ADV(T)/ADV(T - 1)], \quad (7)$$

where $Pr(T)$ is the price at time $T$ and $ADV(T)$ is advertising at time $T$, and the $\beta$s are
weights for the two variables.

If the percentage changes in the variables are constant, $x$ will be constant and the
resulting model will be observationally equivalent to the Bass Model. That is, when
percentage changes in the decision variables are precisely constant, the generalized model
will reduce to the Bass Model. If the percentage changes in decision variables are not
exactly but are approximately constant, the solution to the differential equation will have
the same general shape as the Bass Model but will not be precisely a smooth curve.

Shown earlier were examples of the empirical generalization of exponential decline in
price of new products or technological innovations. Exponential fall in price is, of course,
consistent with a constant percentage decline in price. Moreover, it can be shown that,
under myopic optimization, when marginal costs follow the experience curve form, a constant percentage reduction in price will be approximately optimal. Similar results can be shown for percentage increases in advertising up to the time of peak demand.

Because percentage changes in decision variables are approximately constant we have an explanation of why the Bass Model fits the data without including the decision variables. But we have much more than this because we can now evaluate the effects of different patterns of the decision variables on the diffusion process.

The model has been fit to empirical data, and it yields significant and plausible parameter estimates. Moreover, the estimates for the parameters common to the Bass Model and the Generalized Bass Model are close to one another. Because the generalized model tracks irregular deviations of observations from the smooth curve of the Bass Model, it fits the data better than the special case.

Although there have been several earlier attempts to modify the Bass Model to include decision variables, none of these models reduce to the Bass Model as a special case unless the decision variables are constant. Because the Bass Model will usually fit the data, it seems desirable for an extension or generalization to show the connection between the two models.

Figure 10 shows how the curve will be shifted to the left if prices are reduced by 10 percent from baseline prices. The shape of the shifted curve is still of the Bass Model variety and is a smooth curve because the original curve is based on the assumption that prices fall by a constant percentage. However, highly irregular prices can produce highly irregular shapes under the generalized model. Figure 11 shows a comparison of the smooth curve produced by constant percentage declines in prices in contrast to a highly irregular pattern that is produced by widely varying prices. The parameters are the same in these two examples. Just as behavior under conditions near the speed of light looks little like behavior when gravity is the dominant condition, the generalized model looks little like the special case when the conditions are very different from conditions that generate the special case.
Higher level models have substantial power, and because science is a process leading from one level of theory to higher levels, it is desirable for marketing scientists to explore and develop generalized models.

8. Conclusion

The interests and underlying philosophies of marketing scientists are diverse, but it is our hope that unifying principles will emerge that will lead to greater understanding of marketing phenomena and relationships. Empirical generalizations are central to science of any type, and this is no less the case for marketing science. We hope here to take stock of empirical generalizations in marketing. I have attempted to discuss and illustrate the nature of the scientific process as it applies to marketing science. Although there will be those who will disagree with one or more of the ideas I have advanced, few will disagree with the proposition that science is a process in which theory and data interact leading to greater understanding as theories are generalized, revised, or rejected.

References


