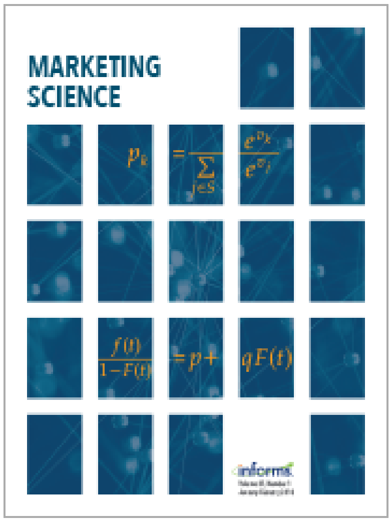


Marketing Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

The Impact of Heterogeneity and Ill-Conditioning on Diffusion Model Parameter Estimates

Albert C. Bemmaor, Janghyuk Lee,

Research Note
The Impact of Heterogeneity and Ill-Conditioning on Diffusion Model Parameter Estimates

Albert C. Bemmaor • Janghyuk Lee
Ecole Supérieure des Sciences Economiques et Commerciales (ESSEC), 95021 Cergy-Pontoise Cedex, France
Groupe HEC, 78351 Jouy-en-Josas Cedex, France
bemmaor@essec.fr • lee@hec.fr

Abstract

Assessment of accurate market size and early adoption patterns is essential to strategic decision making of managers involved in new-product launches. This article proposes methodology that explains changes in parameter estimates of the Bass model, p (coefficient of innovation), q (coefficient of imitation), and c (market penetration rate) by direction of ''extra-Bass'' skew in the data, or equivalently, by underlying heterogeneity of the population. This research shows significantly opposite patterns of these parameter estimates, depending on skew of the diffusion curve detected by a generalized model, i.e., the gamma/shifted Gompertz (G/SG) model, which embeds the Bass model as a special case. The G/SG model originally presented in Bemmaor (1994) is based on two assumptions: (1) Individual-level times to first purchase are distributed shifted Gompertz and (2) individual-level propensity to buy follows a gamma distribution across the population. We assume that the scale parameter of the shifted Gompertz distribution is constant across consumers. The advantage the G/SG model has over alternative diffusion models such as the nonuniform influence model is that its cumulative distribution function takes a closed-form expression.

In line with Van den Bulte and Lilien (1997), we analyze these opposite patterns from simulated data using the G/SG model as the true model and 12 real adoption data sets. The patterns are: (1) as the level of censoring decreases, the estimates of p and c decrease and those of q increase when data exhibit more right skew than the Bass model and (2) the estimates of p and c increase and those of q decrease when data exhibit more left skew than the Bass model. For the simulated data, we manipulated four dimensions: (1) ''extra-Bass'' skew in the data, (2) ratio q/p, (3) speed of diffusion, and (4) error variance. Both results of the simulated data and the real adoption data sets confirm the existence of two opposite patterns of parameter estimates of the Bass model depending on ''extra-Bass'' skew.

When the model is correctly specified with simulated data, estimates of c increase and those of q decrease for both the Bass and the G/SG models. The estimates of p increase as one adds data points only for the G/SG model. No significant tendency in parameter estimates of p was detected for the Bass model. As for ill-conditioning issues, systematic changes in the parameter estimates of the G/SG model can be substantially larger in some cases than those obtained with the Bass model, even though the data were generated by taking the G/SG model as the true one. Therefore, model complexity can aggravate the tendency for parameters to change systematically as one adds data points.

The forecasting results from the simulated data show the supremacy of the G/SG model. It provides more accurate results than the Bass model in the one-step ahead, two-step ahead, and three-step ahead forecasts. With the real data set, the G/SG model provides more accurate one-step ahead forecasts than the Bass model, but the model's forecasting performance deteriorates more rapidly than the Bass model when one shifts to two-step ahead and three-step ahead forecasts. The systematic changes in parameter estimates are larger for the more complex model.

Our research shows that the G/SG model is a flexible model used to analyze the systematic changes in parameter estimates when specification error and ill-conditioning occur. As our findings incorporate two possible types of parameter estimate bias, compared to the previous single-direction view, they can provide essential information to enhance forecasting accuracy of products and services using new technological innovations. Our forecasting results of simulated and real adoption data raise a question about the optimal horizon of forecasting in applying flexible models of diffusion. The G/SG model also provides grounds to investigate jointly ''the speed of takeoff'' and ''the diffusion speed after takeoff''.

(*Diffusion; New-Product Diffusion; Forecasting*)

0732-2399/02/2102/0209/$05.00
1526-548X electronic ISSN

Introduction

Studies of the diffusion of innovations have provided insights to marketing managers involved in commercializing products in their early stage of life cycle. However, recent work has raised questions about the reliability of the parameter estimates obtained with diffusion models. Van den Bulte and Lilien (1997) have shown that estimates of diffusion models often vary systematically as one data point is added. In particular, they showed that for the Bass model (1969) the market size m increases, the coefficient of imitation q decreases, and the coefficient of innovation p increases as the level of censoring decreases or, equivalently, as the number of data points increases. These systematic changes are explained by ill-conditioning, i.e., "lack of richness in the data compared to the complexity of the model" (p. 338). However, these results do not seem to capture the observed changes in the parameter estimates fully. For example, prior research has shown that the parameter estimate $\hat{p}$ may decrease and that the parameter estimate $\hat{q}$ may increase as one data point is added, as shown below with the maximum likelihood estimates for dishwasher (Srinivasan and Mason 1986, Table 5).

No. of Years of Data	$\hat{p}$	$\hat{q}$
4	0.0056	0.0461
8	0.0039	0.0898
12	0.0031	0.1219

In a recent study, Venkatesan et al. (2000) argued that the pattern of instability hinges on the value of the exponent coefficient of the non-uniform influence model (Easingwood et al. 1983). This value is closely related to the amount of skew unaccounted by the Bass model.

The systematic changes may be the consequence of ill-conditioning or of model misspecification. Both causes were mentioned by Van den Bulte and Lilien (1997), but their analysis focused mostly on ill-conditioning. The objective of this paper is to assess in more depth the effect of model misspecification, both separate from and in combination with ill-conditioning. The gamma/shifted Gompertz (G/SG) model presented by Bemmaor (1994) is applied to analyze the consequences of misspecification. The reason for using this model is twofold: (1) The Bass model is nested within the G/SG model and (2) the G/SG model can also generate diffusion curves that show more left or right skew at given levels of p and q, hence allowing one to investigate whether the amount of "extra-Bass" skew[1] is related to the amount and direction of change in the parameter estimates.

Through an extensive simulation where we generate "perturbed" adoption data with the use of the G/SG model, we explain that (1) when the data show "extra-Bass" right skew at given values of p and q, the estimates of p and m decrease and those of q increase,[2] and (2) when the data show "extra-Bass" left skew, the estimates of p and m increase and those of q decrease. The finding of two opposite patterns of parameter estimate bias leads the marketer to better understand the dynamics of diffusion and to enhance forecasting accuracy by taking two patterns into consideration for necessary adaptations. We test these predictions on 12 real data sets and show them to be consistent with the evidence.

The Model: The Bass Model as a Nested Model

The Individual-Level Adoption Model

Following Bemmaor (1994), we assume that the individual-level model of adoption timing is consistent with a two-parameter shifted Gompertz distribution whose cumulative distribution function is given by

$$F(t \mid \eta, b) = (1 - e^{-bt})\exp(-\eta e^{-bt}), \qquad t > 0. \tag{1}$$

The parameter η captures the individual-level propensity to buy. For a fixed value of b, the lower η is, the lower the mean time to adoption (see Appendix 1[3]). Therefore, the lower η is, the stronger the individual-level propensity to buy. The parameter b is a

[1]The skew embedded in more flexible diffusion models, i.e., G/SG, than the Bass model is labeled as "extra-Bass" skew.

[2]The "extra-Bass" skew in the true data generating process is actually manipulated and used in the simulation analysis.

[3]All appendices can be found at http://mktsci.pubs.informs.org

scale parameter that is constant across consumers. The shifted Gompertz distribution is a flexible distribution with a mode at $t^* = (1/b)\ln \eta$ when $\eta > 1$ and $t^* = 0$ when $0 < \eta \leq 1$.

Heterogeneity Assumption and Aggregate-Level Distribution

Let us assume that the individual-level propensity to buy η varies according to a gamma distribution across consumers. The gamma distribution takes a variety of shapes, depending on the value of α. When $\alpha = 1$, this density function becomes an exponential distribution. The coefficient of variation (standard deviation/mean) of a gamma distribution is equal to $\alpha^{-1/2}$. The smaller α is, the larger the heterogeneity. Typically, ''small'' values of α (<1) apply when new products diffuse at differing acceptance rates across strata of the population. Hybrid corn is an example; farmers with higher yields per acre and more acres per farm tended to adopt hybrid corn more quickly than the others (Dixon 1980). ''Large'' values of α apply when new products can be adopted (almost) randomly across the population at any given time t. Examples of this include color television and cellular telephones in some countries. The other parameter, β, of the gamma distribution is a scale parameter (mean $= \alpha\beta$). It follows that the aggregate-level distribution of adoption times is

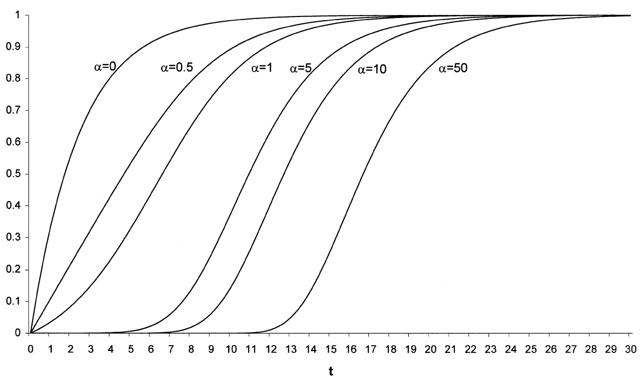
$$F(t) = (1 - e^{-bt})/(1 + \beta e^{-bt})^{\alpha}. \tag{2}$$

This is the G/SG model. We can reparametrize Equation (2) by letting $b = p + q$ and $\beta = q/p$. It follows that we obtain p and q from b and β by using the following equations.

$$p = b/(1 + \beta) \quad \text{and} \quad q = b\beta/(1 + \beta). \tag{3}$$

When $\alpha = 1$, the model is reduced to the Bass model. Although its analytical form is distinct from that proposed by Easingwood et al. (1983), the G/SG model captures an (apparent) nonuniform influence over time. In an empirical study, Parker (1992, Table 2) showed this (so-called) nonuniform influence parameter to be an essential factor that drives diffusion, thereby demonstrating the importance of ''extra-Bass'' skew for descriptive fit. When $\alpha = 0$, Equation

Figure 1 Shapes of the Cumulative Distribution Function of the Gamma/Shifted Gompertz Model



Notes: Fixed values of p and q are $p = 0.03$ and $q = 0.38$.

(2) reduces to an exponential distribution, and when $\alpha = \infty$, it becomes a shifted Gompertz model (see, e.g., Mahajan and Peterson 1985). Therefore, the G/SG model comprehends several models of diffusion. An essential advantage of the G/SG is its closed-form expression, compared with alternative generalizations of the Bass model, e.g., the nonuniform influence model. In Appendix 2, we show some key formulas.

Figure 1 shows the various shapes taken by Equation (2) as the parameter α varies. As α approaches 0, $F(t)$ resembles an exponential diffusion curve and as α becomes large, $F(t)$ resembles a logistic curve. The model explains the clustering of adoption by consumers' homogeneity with respect to their propensity to purchase rather than by contagion (the impact of previous buyers on nonbuyers). Inversely, an exponential type of diffusion can be explained by consumers' heterogeneity rather than by ''independent'' purchasing (lack of contagion). The role of the driver of diffusion does matter for the managerial implications of the model. As shown in Figure 1, the G/SG model captures ''extra-Bass'' skew in the data for given levels of p and q. When α is less than 1, there is more right skew than Bass, and when α is larger than 1, there is more left skew than Bass.

The Pattern of Parameter Estimates by Using Simulated Data

To assess the impact of specification error and ill-conditioning on the parameter estimates, we simulated adoption data in which we assume that (1) the data generation process is given by a G/SG model with known parameters and (2) that the specified model is the Bass model or the G/SG model itself. In addition, we assume that the "observed" pattern of adoption is perturbed by random error ("noise"). Letting $\psi(t)$ be the true number of adopters in period t over the total population and $u(t)$ be a measurement error that is normally distributed with mean 0 and variance σ^2, the "observed" (perturbed) number of adopters $x(t)$ in period t is given by

$$x(t) = \psi(t)\exp(u(t)),$$
$$\text{where } \psi(t) = Mc(F(t) - F(t-1))$$
$$\text{with } t = 0, 1, \ldots, t^{+}. \tag{4}$$

$F(t)$ is the G/SG model (Equation (2)), M is the (known) population size, c is the eventual proportion of adopters ($c = m/M$, where m is unknown), and t^{+} represents the number of data points in the observation period. We fixed the time to peak adoption t^* to be equal to 7 years, the "true" market size c to be equal to 0.4 ($0 < c < 1$), and the population size M to be equal to 10,000. The parameter β is equal to the ratio q/p in the Bass model; this ratio captures the shape of the diffusion model. For a given value of α, the larger β is, the slower the increase of the penetration curve. This parameter can be interpreted as a signal of the diffusion process.

The amount of ill-conditioning is affected by the speed of diffusion: The lower the speed of diffusion is, the larger the level of censoring for a given number of data points. We varied (1) the amount and direction of "extra-Bass" skew, or equivalently, the level of consumer heterogeneity (α); (2) the ratio q/p (signal); (3) the speed of diffusion (low speed corresponds to 80% of normal p and to 80% of normal q); and (4) the error variances (σ^2). Once we fix t^*, q/p and α, we can solve for b using the equations that appear in Appendix 2.[4] We selected the following levels of the variables:

—Six values of α: 0.5, 0.7, 1, 5, 10, and 50;
—Four values of β: 2, 5, 13, and 50 when $\alpha \geq 1$; and 5, 50, 500, and 2500 when $\alpha < 1$[5];
—Two values for speed: normal and low;
—Three values of error variances with $\sigma = 0.06$, 0.24, and 0.42.[6]

For each combination, we generated 50 time series. Therefore, the total number of time series generated is equal to: $6 \times 4 \times 2 \times 3 \times 50 = 7{,}200$. We selected four values of t^{+}: 8, 9, 10, and 11 years. For each time series generated and for each value of t^{+}, we assume that researchers use observed data on cumulative adoptions $N(t)$ to compute the empirical number of adopters in period t:

$$x(t) = N(t) - N(t-1), \tag{5}$$

and then use nonlinear least squares to fit the empirical number of adopters $x(t)$ to the theoretical number of adopters in period t (Srinivasan and Mason 1986):

$$y(t) = Mc(F(t) - F(t-1)), \tag{6}$$

where $F(t)$ is the theoretically derived function (G/SG or Bass, Equation (2)). The actual estimation equation is given by

$$x(t) = Mc(F(t) - F(t-1)) + \epsilon(t). \tag{7}$$

Consistent with Srinivasan and Mason's (1986) estimation procedure and common research practice, the estimation error is operationalized as an additive i.i.d. random term. We analyze the systematic changes in the nonlinear least-squares estimates of the Bass model and the G/SG model as the number of data points increases. From this simulation, we obtain a total of 27,600 estimates.[7] However, we restrict the analysis to the positive parameter estimates and the estimates for which the estimated peak is bounded between 3 to

[4]For example, when $\alpha = 5$ and $\beta = 50$, we find that $b = 0.789$ when $t^* = 7$; hence, $p = 0.0155$ and $q = 0.774$. When $\alpha = 50$ and $\beta = 50$, we find $b = 1.12$ when $t^* = 7$; it follows that $p = 0.0219$ and $q = 1.096$.

[5]The reason for selecting a wider range of values of β when $\alpha < 1$ is to reflect real data more accurately.

[6]The error variances are lower than those used by Van den Bulte and Lilien (1997) because they are based on the estimation of the G/SG model, rather than on the Bass model, to the 12 data sets. They correspond to the three lowest values used by Van den Bulte and Lilien (1997, p. 345).

[7]Two data sets (two speeds) when $\alpha = 0.5$ and $\beta = 5$ were dropped out due to missing parameter values. Therefore, the total number of estimates is 27,600 instead of 28,800.

18. We were left with a total of 23,884 estimates.[8] These estimates form the basis for analyzing the impact of each factor as the level of censoring decreases.

Letting r_{ijklms} denote the nonlinear least-squares estimates $\hat{p}$, $\hat{q}$, and $\hat{c}$ (and $\hat{\alpha}$ for the G/SG model) where $i = 1, 2$, and 3 (and 4 for the G/SG model), respectively, we run two separate regressions for both the Bass model and the G/SG model, one when $\alpha < 1$, and another when $\alpha \geq 1$ to capture the potential interaction effects. The notations are as follows: $\alpha = 0.5$, 0.7, 1, 5, 10, 50 ($j = 1$ to 6), $q/p = 2$, 5, 13, 50, 500, 2500 ($k = 1$ to 6), speed = normal, low ($l = 1$ and 2), error variance, $\sigma = 0.06$, 0.24, 0.42 ($m = 1$ to 3), $\alpha < 1$ ($n = 1$) and $\alpha \geq 1$ ($n = 2$). We run the following regression:

$$\ln r_{ijklms} = a_{inklm} + \delta_{in} \ln Z_{ijklms} + u_{ijklms}, \qquad (8)$$

where a_{inklm} = intercept term, Z_{ijklms} = regressor (t^+), δ_{in} = elasticity with respect to t^+, u_{ijklms} = normally, independently distributed random disturbances with mean zero and constant variance.

When we use the nonlinear least-squares estimates $\hat{p}$, $\hat{q}$, and $\hat{c}$ of the Bass model as the dependent variables in Equation (8), we obtain the results reported in Table 1. The results show the following.

When the data exhibit more right skew than Bass for given values of p and q ($\alpha < 1$),

- The elasticity of $\hat{p}$ with respect to t^+ is negative.
- The elasticity of $\hat{q}$ with respect to t^+ is positive.
- The elasticity of $\hat{c}$ with respect to t^+ is negative.

When the data show more left skew than Bass for given values of p and q ($\alpha > 1$) or when the data are consistent with the Bass model ($\alpha = 1$):

- The elasticity of $\hat{p}$ with respect to t^+ is positive (nonsignificant estimate for $\alpha = 1$).
- The elasticity of $\hat{q}$ with respect to t^+ is negative.
- The elasticity of $\hat{c}$ with respect to t^+ is positive.

The results show consistency, for the most part, with the findings of Van den Bulte and Lilien (1997). They are summarized in Table 2.

[8]Note that 3716 estimates are removed due to this filtering (overall 13% of total estimates). Also note that 25%, 15%, 8%, and 5% of estimates were dropped out, respectively, for the value of $t^+ = 8$, 9, 10, and 11.

For the G/SG model, the nonlinear least-squares estimates $\hat{p}$, $\hat{q}$, $\hat{c}$, and $\hat{\alpha}$ are presented in Table 1. As t^+ increases, the estimates of p and c increase and the estimates of q decrease, regardless of the level of α. When $\alpha \leq 1$, the parameter estimates of α do not show any systematic pattern as one data point is added. When $\alpha > 1$, the estimates of α increase as one data point is added. Note that in certain cases the changes in the parameter estimates can be larger than those obtained with the Bass model even though the G/SG is applied as the true model. For example, this occurs with the parameter estimate $\hat{p}$ when $\alpha \leq 1$. Although large changes do not occur systematically across parameters, this finding enlightens a potential risk of applying complex models that may generate relatively large biases despite the model being true. One question of interest concerns the assessment of the potential effect of the changes in the parameter estimates on the one-step ahead, two-step ahead, and three-step ahead forecasts given by the Bass model and the G/SG model. We computed these forecasts with the simulated data. The Bass model is assessed on data series with model misspecifiation for which α may not equal 1. The mean absolute deviation and the mean squared error corresponding to these forecasts given by each model are shown in Table 3. The G/SG model provides better one-step ahead, two-step ahead, and three-step ahead forecasts than the Bass model. With simulated data where the G/SG is true but also more complex, the relative predictive performance does not seem to be affected by the increased complexity of the model.

Overall, we have shown through simulation that the changes in the parameter estimates of the Bass model hinges on the direction and magnitude of the ''extra-Bass'' skew of the data at given levels of p and q. When the data are more right skew than Bass, the estimates of p and c decrease and those of q increase as one adds data points. When the data are more left skew than Bass, the estimates of p and c increase and those of q decrease as the number of data points increases. When we estimate the G/SG model, we find that the estimates of p and c increase and those of q decrease as the number of observations increases, despite the G/SG being true. The changes in the param-

Table 1 **Relationship Between the Parameters of the Bass Model and the G/SG Model and the Number of Data Points in the Observation Period**

Explanatory Variables	Bass Model						Gamma/Shifted Gompertz Model							
	ln $\hat{p}$		ln $\hat{q}$		ln $\hat{c}$		ln $\hat{p}$		ln $\hat{q}$		ln $\hat{c}$		ln $\hat{\alpha}$	
	$\alpha < 1$	$\alpha \geq 1$	$\alpha < 1$	$\alpha \geq 1$	$\alpha < 1$	$\alpha \geq 1$	$\alpha < 1$	$\alpha \geq 1$	$\alpha < 1$	$\alpha \geq 1$	$\alpha < 1$	$\alpha \geq 1$	$\alpha < 1$	$\alpha \geq 1$
Intercept	−5.899	−28.509	−0.295	3.056	−0.818	−2.707	−14.614	−8.579	0.574	2.525	−2.023	−1.811	−0.0620	1.907
	(0.182)[a]	(0.256)	(0.071)	(0.046)	(0.059)	(0.046)	(0.762)	(0.664)	(0.238)	(0.239)	(0.063)	(0.052)	(0.259)*	(0.361)
Normal speed	0.0663	1.260	0.258	0.0363	−0.0271	0.145	0.575	0.968	0.237	0.00928	−0.0892	0.0916	−0.0510	0.215
	(0.019)	(0.023)	(0.007)	(0.004)	(0.006)	(0.004)	(0.078)	(0.060)	(0.024)	(0.021)*	(0.006)	(0.005)	(0.027)*	(0.032)
Error variance														
$\sigma = 0.06$	0.240	0.462	−0.252	−0.129	0.214	0.106	1.223	2.072	0.184	0.0590	0.116	0.177	−0.497	−0.196
	(0.023)	(0.028)	(0.009)	(0.005)	(0.007)	(0.005)	(0.095)	(0.073)	(0.030)	(0.026)	(0.008)	(0.006)	(0.032)	(0.040)
$\sigma = 0.24$	0.199	0.282	−0.132	−0.0571	0.09305	0.0527	0.564	0.863	0.131	0.0768	0.0509	0.0850	−0.341	−0.136
	(0.023)	(0.029)	(0.009)	(0.005)	(0.007)	(0.005)	(0.097)	(0.074)	(0.030)	(0.027)	(0.008)	(0.006)	(0.033)	(0.040)
Heterogeneity														
$\alpha = 0.5$	1.032		−0.455		0.131		0.281		−0.125		0.00196		−0.352	
	(0.020)		(0.008)		(0.006)		(0.082)		(0.026)		(0.007)*		(0.028)	
$\alpha = 1$		21.975		−2.910		0.530		−1.852		−2.437		−0.630		−1.334
		(0.527)		(0.095)		(0.096)		(1.368)*		(0.491)		(0.107)		(0.744)*
$\alpha = 5$		4.479		−0.560		−0.0812		−0.856		−0.739		−0.0337		−1.890
		(0.032)		(0.006)		(0.006)		(0.083)		(0.030)		(0.006)		(0.045)
$\alpha = 10$		2.965		−0.325		−0.0653		−0.789		−0.504		−0.0383		−1.227
		(0.032)		(0.006)		(0.006)		(0.083)		(0.030)		(0.007)		(0.045)
q/p ratio														
2		3.939		−0.797		−0.160		1.207		−1.545		−0.127		−0.0901
		(0.033)		(0.006)		(0.006)		(0.085)		(0.030)		(0.007)		(0.046)
5	3.377	2.799	−2.196	−0.526	−0.172	−0.0695	5.377	0.854	−2.638	−1.039	−0.318	−0.0385	0.839	0.00506
	(0.034)	(0.032)	(0.013)	(0.006)	(0.011)	(0.006)	(0.144)	(0.084)	(0.045)	(0.030)	(0.012)	(0.007)	(0.049)	(0.045)*
13		1.638		−0.257		−0.0379		0.548		−0.483		−0.0191		−0.0456
		(0.032)		(0.006)		(0.006)		(0.083)		(0.030)		(0.007)		(0.095)*
50	2.565		−1.114		0.136		3.959		−1.101		0.0245		0.233	
	(0.024)		(0.009)		(0.008)		(0.101)		(0.031)		(0.008)		(0.034)	
500	1.253		−0.347		0.0577		1.672		−0.299		0.0138		7.29×10^{-6}	
	(0.023)		(0.009)		(0.008)		(0.098)		(0.031)		(0.008)		(0.033)*	

Table 1 **Continued**

Explanatory Variables	Bass Model						Gamma/Shifted Gompertz Model							
	$\ln \hat{p}$		$\ln \hat{q}$		$\ln \hat{c}$		$\ln \hat{p}$		$\ln \hat{q}$		$\ln \hat{c}$		$\ln \hat{\alpha}$	
	$\alpha < 1$	$\alpha \geq 1$	$\alpha < 1$	$\alpha \geq 1$	$\alpha < 1$	$\alpha \geq 1$	$\alpha < 1$	$\alpha \geq 1$	$\alpha < 1$	$\alpha \geq 1$	$\alpha < 1$	$\alpha \geq 1$	$\alpha < 1$	$\alpha \geq 1$
$\alpha<1$														
$\ln t^{+}$	−0.425		0.0901		−0.0923		2.018		−0.343		0.426		0.0547	
	(0.080)		(0.031)		(0.026)		(0.333)		(0.104)		(0.028)		(0.113)*	
$\alpha=1$														
$\ln t^{+}$		−0.0701		−0.480		0.474		1.692		−0.512		0.564		−0.00416
		(0.205)*		(0.037)		(0.037)		(0.532)		(0.191)		(0.042)		(0.289)*
$\alpha>1$														
$\ln t^{+}$		6.274		−1.004		0.705		1.239		−1.031		0.339		0.980
		(0.112)		(0.020)		(0.020)		(0.290)		(0.104)		(0.023)		(0.158)
R^2	0.719	0.814	0.846	0.868	0.268	0.192	0.278	0.078	0.395	0.211	0.199	0.122	0.120	0.263

[a]The standard errors are shown in parentheses. All the coefficients are significant at the 99% confidence level, except for those indicated by $*$. $N = 6700$ for $\alpha < 1$ and $N = 17184$ for $\alpha \geq 1$. For $\alpha < 1$, the base is given by: low speed, $\sigma = 0.42$, $\alpha = 0.70$, $\beta = 2500$. For $\alpha \geq 1$, the base is given by: low speed, $\sigma = 0.42$, $\alpha = 50$, $\beta = 50$.

Table 2 Summary of the Simulation Results: Impact of Specification Error and Ill-Conditioning on the Parameters Estimates of the Bass Model

Value of the Parameter α	Results of the Simulation
The data exhibit more right skew than Bass ($\alpha < 1$) for given values of p and q	$\hat{p}$ decreases as the number of data points increases $\hat{q}$ increases as the number of data points increases $\hat{c}$ decreases as the number of data points increases A high error variance results in a downward bias for $\hat{p}$ *and* $\hat{c}$ and an upward bias in $\hat{q}$ A low diffusion speed results in a downward bias in $\hat{p}$ and $\hat{q}$ and an upward bias in $\hat{c}$
The data are consistent with the Bass model ($\alpha = 1$)	$\hat{q}$ decreases as the number of data points increases $\hat{c}$ increases as the number of data points increases
The data exhibit more left skew than Bass ($\alpha > 1$) for given values of p and q	$\hat{p}$ increases as the number of data points increases $\hat{q}$ decreases as the number of data points increases $\hat{c}$ increases as the number of data points increases A high error variance results in a downward bias for $\hat{p}$ and $\hat{c}$ and an upward bias in $\hat{q}$ A low diffusion speed results in a downward bias in $\hat{p}$, $\hat{q}$ and $\hat{c}$

eter estimates can be larger than those obtained with the Bass model for the corresponding parameter. However, when we computed one-step ahead, two-step ahead, and three-step ahead forecasts, we found that the G/SG model provided relatively more accurate forecasts than the Bass model. Therefore, model complexity did not seem to penalize forecasting accuracy when the model is true.

The Pattern of Parameter Estimates by Using Real Adoption Data

The Adoption Data

The adoption data of 12 new products and services (Van den Bulte and Lilien 1997, Appendix A) provide a mixture of four-year, annual, semiannual, and monthly data.[9] The reason for using them is to cast further light onto the biases and changes in the parameter estimates.

Estimation

For each product or service, we estimated the Bass model and the G/SG model using the total data set (maximum value of t^+). Letting m be the eventual number of adopters and $x(t)$ be the actual number of adopters in year t, we minimized the following sum with respect to the model's parameters (Srinivasan and Mason 1986):

$$S = \sum_{t=1}^{t^+} \{m[F(t) - F(t-1)] - x(t)\}^2, \qquad (9)$$

where $F(t)$ is given in Equation (2) (Bass or G/SG model) and t^+ represents the total number of observations for the new product or service. For the G/SG model, we obtained the following nonlinear least squares estimates of α:

	Product or Service	Estimate of α
Cases with $\hat{\alpha} < 1$	Ultrasound	0.30
	Corn (1948)	0.30
	Foreign language	0.37
	Mammography	0.40
	Corn (1943)	0.70
Cases with $\hat{\alpha} >\approx 1$	Tetracycline	0.97
	Air conditioner	232.08
	Clothes dryer	261.68
	Color TV	571.50
	CT scanner	2500.62
	Compulsory school	3144.28
	Accelerated program	3756.94

To assess the potential changes in the estimates with the number of data points in the observation period, we estimated the model with varying levels of censoring (t^+), always including peak sales to enhance the stability of the parameter estimates. We ob-

[9]The years of launch of color television, room air conditioner, and clothes dryer are 1954, 1929, and 1936, respectively (Golder and Tellis 1997).

Table 3 One-Step, Two-Step, and Three-Step Ahead Forecasts with Simulated Data

	Mean Absolute Deviation			Mean Squared Error		
Model	MAD1[a]	MAD2	MAD3	MSE1[a] ($\times 10^3$)	MSE2 ($\times 10^3$)	MSE3 ($\times 10^3$)
Bass model	241.1	255.6	243.5	208.0	253.2	226.0
G/SG model	214.3	223.4	216.5	155.6	179.0	169.3

[a]MAD1 and MSE1 denote the mean absolute deviation and the mean squared error of the one-step ahead forecasts, respectively. N = 10644.

tained a total of 60 parameter estimates (Appendix 3). Using the estimates of α as a selection criterion, we grouped the products into two categories: those with $\hat{\alpha} < 1$ (five products) and those with $\hat{\alpha} \gtrsim 1$ (seven products). Using the same notations as in Equation (8), we estimated the following equations:

$$\ln r_{ivs} = a_{iv} + \delta_{in} \ln Z_{iv} + u_{ivs}, \tag{10}$$

where v denotes the product or service ($v = 1, \ldots, 12$), and

$$\ln r_{ivs} = a_{iv} + \delta_{i} \ln Z_{iv} + u_{ivs}. \tag{11}$$

Equation (11) is a constrained version of Equation (10) because it ignores the interaction between the slope and the level of heterogeneity. As shown in Table 4, we find that, consistent with the results obtained with simulated data, the parameter estimates of the Bass model change in the following way as one data point is added: (1) When $\hat{\alpha} < 1$, $\hat{p}$ and $\hat{c}$ decrease and $\hat{q}$ increases, and (2) when $\hat{\alpha} \gtrsim 1$, $\hat{p}$, and $\hat{c}$ increase and $\hat{q}$ decreases as the level of censoring decreases. When all of the data are pooled without interaction and analyzed by using Equation (11), the results obtained by Van den Bulte and Lilien (1997) are reproduced, i.e., $\hat{c}$ increases and $\hat{q}$ decreases as the level of censoring decreases. The elasticity of $\hat{p}$ is not significant but shows the expected positive sign. The inclusion of the interaction had little effect on the R^2 s, but this can be explained by (1) the relatively low number of observations per product or service and (2) the inclusion of varying intercept terms in the equations. When we estimate the gamma/shifted Gompertz model, we find a deterioration of the systematic changes in the parameter estimates. For example, the elasticity of $\hat{p}$ with respect to t^+ shifts from 1.202 for the Bass model to 9.759 for the G/SG model when $\hat{\alpha} \geq 1$. The shift is even more drastic when $\hat{\alpha} < 1$: from -1.700 for the Bass model to 35.960 for the G/SG model. Therefore, a more elaborate diffusion model aggravates the tendency for the parameters to exhibit systematic changes. As we saw before, this pattern does not occur with

Table 4 Relationship Between the Parameters of the Models and the Number of Observations

	Bass Model						Gamma/Shifted Gompertz Model			
	ln $\hat{p}$		ln $\hat{q}$		ln $\hat{c}$					
Explanatory Variables	With Interaction	Without Interaction	With Interaction	Without Interaction	With Interaction	Without Interaction	ln $\hat{p}$	ln $\hat{q}$	ln $\hat{c}$	ln α
$\alpha < 1$										
ln $t+$	-1.700		0.445		-0.350		35.960	-1.834	0.401	2.015
	(0.836)[a]		(0.175)		(0.138)		(9.759)	(2.324)*	(0.215)*	(3.122)*
$\alpha \gtrsim 1$										
ln $t+$	1.202		-0.841		0.397		9.759	-7.392	0.559	8.534
	(0.423)		(0.089)		(0.070)		(4.938)*	(1.176)	(0.109)	(1.580)
ln $t+$		0.610		-0.579		0.244				
		(0.411)*		(0.109)		(0.076)				
R^2	0.986	0.983	0.981	0.963	0.990	0.989	0.742	0.902	0.974	0.912

[a]Standard errors of the estimates are shown in parentheses. All of the coefficients are significant at the 95% confidence level, except for those indicated with *. The base is ultrasound. For simplicity of presentation, we omit the intercepts and all product dummy estimates. $N = 60$.

the simulated data, with the G/SG being the true model. Consequently, in the empirical data where neither the Bass model nor the G/SG model is likely to be exactly true, the increased complexity of the G/SG penalizes the stability of the parameter estimates severely.

In sum, the data are overwhelmingly consistent with the expectations. The direction of skew affects the direction of change in the parameter estimates. Both simulated data and empirical data support this argument. We have shown that because of ill-conditioning, more complex models can result in larger changes in the parameter estimates as the number of data points increases. This corresponds to Van den Bulte and Lilien's (1997, p. 348) conjecture that ''Making the model more complex can actually aggravate the tendency of parameter estimates to change systematically as one extends the data set.''

Forecasting Performance of the Bass and G/SG models

The biases and systematic changes in the parameters observed with the Bass model and the even larger changes observed with the G/SG model can lead us to wonder whether a more flexible, and potentially more realistic, model is worth the effort. To address this issue, we compared (1) the fit and (2) the one-step ahead, two-step ahead, and three-step ahead forecasts given by the G/SG model to those given by the Bass model.

Comparing the Models' Fit to Adoption Data

Appendix 4 shows the mean squared error and the mean absolute deviation obtained with each model on the 12 products and services. On the average, the mean absolute deviation corresponding to the G/SG model is 27% lower than that corresponding to the Bass model, and the mean squared error is 42% lower. Adding a single parameter to the Bass model improves the descriptive accuracy of the model substantially.

Comparing the One-Step Ahead, Two-Step Ahead, and Three-Step Ahead Forecasts

An essential purpose of diffusion models is to forecast the penetration of new products. We compared the one-step-ahead, two-step ahead, and three-step ahead forecasts given by the Bass model with those given by the G/SG model for each of the 12 products and services. As shown in Appendix 5, on the average, the G/SG model provides better one-step-ahead forecasts than the Bass model; the mean absolute deviation is 7.9% lower than that corresponding to the Bass model and the mean squared error is 1.9% lower. Therefore, the larger parameter instability of the G/SG model does not seem to damage the model's one-step ahead forecasting performance relative to the Bass model. However, the forecasting performance of the model deteriorates more rapidly than that of the Bass model as one shifts from one-step ahead to two-step ahead and three-step ahead forecasts. The mean absolute deviation and the mean squared error of the three-step ahead forecasts are, respectively, 16.8% and 27.2% larger for the G/SG model than for the Bass model.

Overall, the comparison of the Bass model with a more flexible model shows that, despite its substantial changes in the parameter estimates as the level of censoring decreases, the G/SG model provides a better fit to the data and better one-step ahead forecasts than the Bass model. As this forecasting accuracy deteriorates more rapidly than that of the Bass model for relatively long-range forecasting, model parsimony becomes an essential model characteristic when ill-conditioning takes place.

Conclusion

This research focused on the systematic change of parameter estimates in a standard aggregate-level diffusion model. By using a flexible model of diffusion with simulated data, we have shown how the relationship between the level of censoring and the parameter estimates interacts with the amount and direction of ''extra-Bass'' skew in the data. The G/SG model allows more flexibility to represent the heterogeneity of consumers' propensities to buy. Specification error on the variation of consumers' propensities to buy affects the pattern of changes of the parameter estimates as one data point is added. The results on systematic changes in the parameter estimates can be summarized as follows:

(1) When the model is correctly specified, the estimate of c increases and that of q decreases as one adds data points. This holds for both the Bass and G/SG model. The estimate of p increases for the G/SG model.

(2) Similarly, when the data exhibit more right skew than the Bass model, the Bass model estimates of p and c decrease and that of q increases as one adds data points.

(3) In contrast, when the data exhibit more left skew than the Bass model, the Bass model estimates of p and c increase and that of q decreases as one adds data points.

(4) With real data, the systematic changes are larger in absolute size for the more complex G/SG model. With simulated data where the true model is the G/SG model, the systematic changes in the parameter estimates obtained with the G/SG model can be substantial also when they occur.

(5) On simulated data, the predictive performance of the G/SG model is higher than that of the Bass model for one-step, two-step, and three-step ahead forecasts; for 12 real data sets, the G/SG has higher predictive performance only for one-step ahead forecasts.

Points 1 and 4 are fully in line with the ill-conditioning argument advanced by Van den Bulte and Lilien (1997) for q and c, replicate their findings for the Bass model, and extend them to the G/SG model. Findings 2 and 3 document how model misspecification and specifically the extent and direction of skew unaccounted for by the Bass model affect the direction of the bias. These results qualify the earlier results by Van den Bulte and Lilien (1997) about the direction of change. The fifth finding highlights the disadvantage of flexible models in the presence of noisy data, especially when one does not know the true data-generating process. The above results are specific to the chosen functional form of the G/SG model and may not be extended to other forms. Further studies with alternative functional forms need to be done to assess the scope of our results.

The systematic change of major parameters such as the "imitation" factor q and the market potential c, depending on the level of heterogeneity may expose the user to a high risk of misinterpreting the pattern of diffusion, especially in the early stages. This risk of misinterpretation can be reduced to the minimum only by applying a diffusion model such as the G/SG that is able to capture the heterogeneity of the population. This study can be expanded with the use of even more flexible models. For example, work in the mathematical biosciences has shown that not only does heterogeneity matter, but also the manner in which the heterogeneous populations interact has an important impact on the diffusion process (see, e.g., Kaplan and Lee 1990, Putsis et al. 1997). Simulation of these processes and their impact on the parameter estimates of the Bass model may bring further insights into the sources of parameter instability.

In addition, some characteristics of G/SG (four curves, shown in Figure 1, for which $\alpha \geq 1$, having a similar slope with a different takeoff time) may provide grounds to investigate jointly issues of "the speed of takeoff" (Golder and Tellis 1997) and "the diffusion speed after takeoff" (Van den Bulte 2000). Overall, the G/SG model, with its flexibility and capacity to capture a wide variety of shapes of diffusion, provides a useful basis for analyzing the changes in the parameter estimates of the Bass model and its extension.

Acknowledgments
Janghyuk Lee is grateful for the doctoral program at ESSEC because of its financial assistance. The authors are grateful to the anonymous reviewers for their constructive comments on an earlier draft of this study and to the participants of the Marketing Science Conference in Syracuse, NY, and the joint ESSEC/HEC/INSEAD seminar.

References

Bass, Frank M. 1969. A new product growth model for consumer durables. *Management Sci.* **15** 215–227.

Bemmaor, Albert C. 1994. Modeling the diffusion of new durable goods: Word-of-mouth effect versus consumer heterogeneity. Gilles Laurent, Gary L. Lilien, and Bernard Pras, eds. *Research Traditions in Marketing.* Kluwer, Boston, MA. 201–223.

Dixon, R. 1980. Hybrid corn revisited. *Econometrica* **48** 1451–1461.

Easingwood, Christopher J., Vijay Mahajan, Eitan Muller. 1983. A nonuniform influence innovation diffusion model of new product acceptance. *Marketing Sci.* **2**(3) 273–295.

Golder, Peter N., Gerard J. Tellis. 1997. Will it ever fly? Modeling the takeoff of really new consumer durables. *Marketing Sci.* **16** (3) 256–270.

Kaplan, Edward H., Yew Sing Lee. 1990. How bad can it get? Bounding worst case endemic heterogeneous mixing models of HIV/AIDS. *Math. Biosciences* **99** 157–180.

Mahajan, Vijay, Robert A. Peterson. 1985. *Models for Innovation Diffusion.* Sage Publications, Beverly Hills, CA.

Parker, Philip M. 1992. Price elasticity dynamics over the adoption life cycle. *J. Marketing Res.* **29**(3) 358–367.

Putsis, William P., Jr., Sridhar Balasubramanian, Edward H. Kaplan, Subrata K. Sen. 1997. Mixing behavior in cross-country diffusion. *Marketing Sci.* **16** (4) 354–369.

Srinivasan, V., Charlotte H. Mason. 1986. Nonlinear least squares estimation of new product diffusion models. *Marketing Sci.* **5**(2) 169–178.

Van den Bulte, Christophe. 2000. New product diffusion acceleration: Measurement and analysis. *Marketing Sci.* **19** (4) 366–380.

——, Gary L. Lilien. 1997. Bias and systematic change in the parameter estimates of macro-level diffusion models. *Marketing Sci.* **16** (4) 338–353.

Venkatesan, Rajkumar, Trichy Krishnan, V. Kumar. 2000. Systematic changes in parameters of macro-level diffusion models: A further inquiry. *Marketing Science Conference,* University of California at Los Angeles, Los Angeles, CA.

This paper was received November 1, 2000, and was with the authors 8 months for 2 revisions; processed by Gary Lilien.