The Econometrics Of The Bass Diffusion Model

H. Peter Boswijk, Philip Hans Franses

ERIM Report Series *Research in Management*	
ERIM Report Series reference number	ERS-2002-66-MKT
Publication	2002
Number of pages	42
Email address corresponding author	franses@few.eur.nl
Address	Erasmus Research Institute of Management (ERIM) Rotterdam School of Management / Faculteit Bedrijfskunde Erasmus Universiteit Rotterdam P.O. Box 1738 3000 DR Rotterdam, The Netherlands Phone: +31 10 408 1182 Fax: +31 10 408 9640 Email: info@erim.eur.nl Internet: www.erim.eur.nl

Bibliographic data and classifications of all the ERIM reports are also available on the ERIM website:
www.erim.eur.nl

ERASMUS **R**ESEARCH **I**NSTITUTE OF **M**ANAGEMENT

REPORT SERIES
RESEARCH IN MANAGEMENT

BIBLIOGRAPHIC DATA AND CLASSIFICATIONS

Abstract

We propose a new empirical representation of the Bass diffusion model, in order to estimate the three key parameters, concerning innovation, imitation and maturity. The representation is based on the notion that the observed data may temporarily deviate from the mean path determined by the underlying hazard rate. Additionally, it rests on the idea that uncertainty about the cumulative process should be smaller, the closer it is to the start of the process and to the level of maturity. Taking this into account, we arrive at an extension of the basic representation proposed in Bass (1969), with an additional heteroskedastic error term. The type of heteroskedasticity can be set by the modeler, as long as it obeys certain properties. Next, we discuss the asymptotic theory for this new empirical model, that is, we focus on the properties of the estimators of the various parameters. We show that the parameters, upon standardization by their standard errors, do not have the conventional asymptotic behavior. For practical purposes, it means that the *t*-statistics do not have an (approximate) *t*-distribution. Using simulation experiments, we address the issue how these findings carry over to practical situations. In a next set of simulation experiments, we compare the new representation with that of Bass (1969) and Srinivasan and Mason (1986). We document that these last two approaches often seriously overestimate the precision of the parameter estimators. We also shed light on the effects of temporal aggregation and on the effects of a serious and persistent deviation between the actual data and their mean. Finally, we consider the various empirical representations for a monthly series on installed ATMs.

Library of Congress Classification (LCC)	5001-6182	Business
	5410-5417.5	Marketing
	HB 141	Econometric Models
Journal of Economic Literature (JEL)	M	Business Administration and Business Economics
	M 31 C 44	Marketing Statistical Decision Theory
	C 51	Model construction and estimation
European Business Schools Library Group (EBSLG)	85 A	Business General
	280 G 255 A	Managing the marketing function Decision theory (general)
	150 K	Applied Econometrics
Gemeenschappelijke Onderwerpsontsluiting (GOO)		
Classification GOO	85.00	Bedrijfskunde, Organisatiekunde: algemeen
	85.40 85.03	Marketing Methoden en technieken, operations research
	85.03	Methoden en technieken, operations research
Keywords GOO	Bedrijfskunde / Bedrijfseconomie	
	Marketing / Besliskunde	
	Econometrische modellen, Representatie (wiskunde), Onzekerheid	
Free keywords	Bass Diffusion Model, Representation, Estimation	

The Econometrics of the Bass Diffusion Model

H. Peter Boswijk

Department of Quantitative Economics

Universiteit van Amsterdam

Philip Hans Franses*

Econometric Institute

Erasmus University Rotterdam

July 10, 2002

*Address for correspondence: Econometric Institute H11-34, Erasmus University Rotterdam, P.O.Box 1738, NL-3000 DR Rotterdam, The Netherlands; e-mail: franses@few.eur.nl. The computer programs used for all calculations in this paper can be obtained from the first author at peterb@fee.uva.nl.

The econometrics of the Bass diffusion model

Abstract

We propose a new empirical representation of the Bass diffusion model, in order to estimate the three key parameters, concerning innovation, imitation and maturity. The representation is based on the notion that the observed data may temporarily deviate from the mean path determined by the underlying hazard rate. Additionally, it rests on the idea that uncertainty about the cumulative process should be smaller, the closer it is to the start of the process and to the level of maturity. Taking this into account, we arrive at an extension of the basic representation proposed in Bass (1969), with an additional heteroskedastic error term. The type of heteroskedasticity can be set by the modeler, as long as it obeys certain properties.

Next, we discuss the asymptotic theory for this new empirical model, that is, we focus on the properties of the estimators of the various parameters. We show that the parameters, upon standardization by their standard errors, do not have the conventional asymptotic behavior. For practical purposes, it means that the t-statistics do not have an (approximate) t-distribution. Using simulation experiments, we address the issue how these findings carry over to practical situations.

In a next set of simulation experiments, we compare the new representation with that of Bass (1969) and Srinivasan and Mason (1986). We document that these last two approaches often seriously overestimate the precision of the parameter estimators. We also shed light on the effects of temporal aggregation and on the effects of a serious and persisent deviation between the actual data and their mean.

Finally, we consider the various empirical representations for a monthly series on installed ATMs.

Key words: Bass diffusion model, representation, estimation

1 Introduction

The Bass diffusion model, introduced in Bass (1969), is frequently applied in modeling and forecasting diffusion processes in marketing research and other disciplines. The basic model is fascinatingly simple, also as its basic form only contains three parameters. Its popularity is likely to be due to the fact that the model can capture a wide variety of diffusion patterns observed in practice. The model parameters have an interpretation in terms of internal effects (imitation), external effects (innovation) and saturation level (maturity), and the basic theory behind the model is solid. Additionally, the model allows for various modifications without loosing its interpretation.

Interestingly, although the theoretical framework of the Bass model is beyond doubt, its translation into a representation for observed empirical and discrete data appears not straightforward. There are two main reasons for this phenomenon. The first is that the basic theory is formulated in continuous time, while in practice one always has discrete data. Hence, the question is how one represents the theory, while taking in account that one does not have continuous data. The second is that the theoretical model can be written in various ways, and that translations of this into an empirical representation should involve a decision on where to put the stochastic error term. Indeed, once an error term has been added to a model, it is not easy to write one empirical model into another, even though they are all based on the same theory.

It is conceivable that different model representations lead to different parameter estimates and associated standard errors. Hence, the choice for a model representation matters in practice. In this paper we aim to contribute to the literature by proposing a new empirical representation of the Bass diffusion model, which we believe is the most natural one, given the theory. A key feature is that we assume that the discrete data have a tendency to revert to the underlying mean path of the continuous process, and that this tendency may vary across processes and levels of temporal aggregation. As such, our model should be resistant to aggregation of the data, in the sense that, even though the discrete data are observed at too high an aggregation level, the underlying three key parameters can still be estimated. As this is also the property of the empirical representation in Srinivasan and Mason (1986), we shall compare our model with theirs when it comes to comparing the consequences of aggregation.

A second important feature of our model is that we assume that the uncertainty on the diffusion path should not be constant over time. Indeed, in the beginning and towards the end of a diffusion process, there must be much less uncertainty than somewhere in the middle. For example, if the process concerns first-time purchases, in the beginning sales are zero, and in the end they are zero again, while they obtain peak levels somewhere before halfway the sample. Consequently, there must be more uncertainty about cumulative sales around the inflection point of the S-shaped curve than there is close to the saturation level.

It turns out that our model extends the basic Bass (1969) representation in two dimensions. The first is that our model incorporates an additional variable, representing the dynamic adjustment to the underlying mean diffusion path. The second is that we allow the error process to be heteroskedastic, where the type of heteroskedastic variation is up to the user to decide, subject to the property mentioned above.

In our paper we propose the new model formulation, derive the asymptotic properties of the corresponding estimators and compare these with the models in Bass (1969) and Srinivasan and Mason (1986). A salient finding is that the estimators do not have standard asymptotic properties. We use simulation experiments to examine how this carries over to practical cases. When we compare our model with other models, we observe that the confidence intervals around the estimators in the Bass (1969) and Srinivasan and Mason (1986) models are often too narrow, which is in part due to their assumption that the variance of the error process is constant over time.

The outline of our paper is as follows. In Section 2, we put forward the new model representation for the theoretical Bass diffusion model. We start with an analysis in continuous time, proposing a stochastic differential equation for the diffusion process, and then we discuss the discretization of the continuous-time process. In Section 3, we focus on parameter estimation. We show that the parameters in the Bass models are not consistent in the usual sense. We examine the consequences of this result for practically relevant cases through simulations. In Section 4, we compare our model with two other models, thereby focusing on the effects of variable omission and the consequences of having a heteroskedastic error process. In Section 5, we illustrate our model on real-life data, and we find that most of our predictions based on theory and simulations hold true for this example. In Section 6, we discuss various topics for further research.

2 Representation

In this section, we first discuss the basic premises of the Bass diffusion theory. Next, we put forward our new representation, which accounts for temporary deviations of sales from their expected value implied by the Bass model, and for level-dependent heteroskedasticity. In the last subsection, we compare it with the representations in Bass (1969) and Srinivasan and Mason (1986).

2.1 The Bass diffusion model

The Bass model starts with a population of m potential adopters. For each of them, the time to adoption is a random variable with a distribution function $F(t)$ and density $f(t)$, such that the hazard rate satisfies

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

The cumulative number of adopters at time t, $N(t)$, where t is measured in continuous time, is therefore a random variable with mean $\bar{N}(t) = E[N(t)] = mF(t)$. The function $\bar{N}(t)$ then satisfies the following differential equation, that is,

$$\bar{n}(t) = \frac{d\bar{N}(t)}{dt} = p[m - \bar{N}(t)] + \frac{q}{m}\bar{N}(t)[m - \bar{N}(t)]. \tag{2}$$

The solution of this differential equation is given by

$$\bar{N}(t) = mF(t) = m\left[\frac{1-e^{-(p+q)t}}{1+\frac{q}{p}e^{-(p+q)t}}\right], \quad (3)$$

$$\bar{n}(t) = mf(t) = m\left[\frac{p(p+q)^2 e^{-(p+q)t}}{(p+qe^{-(p+q)t})^2}\right]. \quad (4)$$

The parameter p can be interpreted as the innovation parameter (or, external effect), q can be seen as the imitation parameter (or, internal effect), and m is the saturation level (or, maturity). In practice, one is interested in estimating these three key parameters. An additional motivation is that the inflection point (which occurs at the time of peak adoptions) of $F(t)$ is a function of p and q, and a first good guess of the timing of this peak can be important for managers.

It is important to recognize that $N(t)$ is a random variable, for which the law of large numbers implies that as m increases, $N(t) \to \bar{N}(t)$. For the size of m typically encountered in empirical applications, $N(t)$ may in fact be equated to its mean. This is illustrated in Figure 1, which concerns simulated data from the model with $m = 100000$, $p = 0.01$ and $q = 0.15$. Note that $N(t)/m$ (obtained from simulating m waiting times with density $f(t)$) and its mean $F(t)$ (given in (3)) are indistinguishable. Similarly, the scaled increments $\Delta N(t)/m = [N(t) - N(t-1)]/m$, $t = 1, 2, \ldots$, are virtually identical to the smooth function $\Delta F(t)$.

Insert Figure 1 about here

In practical applications, the cumulative number of adopters, and in particular its increments, are not completely described by a smooth curve as in Figure 1, even for very large m. This may be caused by individual-specific effects on the hazard rates, which could be observable or which could amount to unobserved heterogeneity, as well as by aggregate effects caused by marketing-mix variables or environmental factors (like the macro economy). These effects imply that one should either extend the model with individual and/or aggregate covariates, or allow for more general random deviations from (3). In this paper we follow the latter approach.

2.2 New representation

Upon translating the Bass diffusion theory to a model that can be confronted with the data to deliver estimates of the parameters, one needs to allow for random, and possibly serially dependent, deviations from the path of the continuous process given by (3)–(4). Furthermore, the continuous-time model has to be translated to discrete-time observations $N_i = N(t_i), i = 0, 1, \ldots, T$, with corresponding incremental number of adopters over the interval $(t_{i-1}, t_i]$ given by $X_i = N_i - N_{i-1}$. In what follows, we shall refer to X_i as "sales".

Our starting point is to generalize the differential equation (2) to a (continuous-time) *stochastic* differential equation (SDE) for $N(t)$ and $n(t) = dN(t)/dt$. This SDE may then be discretized into a stochastic difference equation, which in turn may be used for least-squares or maximum likelihood estimation. We require the process for $n(t)$ and $N(t)$ to satisfy the following properties. First, $n(t)$ is non-negative, so that $N(t) = \int_0^t n(s)ds$ is non-decreasing. Second, and related to the first, we require the volatility of $n(t)$ to depend on the level of $n(t)$, in such a way that the random variation converges to zero as $n(t) \downarrow 0$ (which, for example, will occur as $N(t) \uparrow m$). Third, we allow the random deviations in $n(t)$ from its mean path to display mean-reverting serial correlation. To our knowledge, no studies in the relevant literature have addressed the second and third property.

For any path $N(t), t \geq 0$, let

$$n^*(t) = p[m - N(t)] + \frac{q}{m} N(t)[m - N(t)]. \tag{5}$$

Note that this corresponds to (2), but with the mean number of adopters $\bar{N}(t)$ replaced by the actual number. We propose the following two alternative specifications

$$\begin{aligned} dn(t) &= \alpha\left[\bar{n}(t) - n(t)\right] dt + \sigma n(t)^\gamma dW(t), && \text{(6a)} \\ dn(t) &= \alpha\left[n^*(t) - n(t)\right] dt + \sigma n(t)^\gamma dW(t), && \text{(6b)} \end{aligned}$$

where $\alpha > 0$, $\sigma > 0$ and $\gamma \geq \frac{1}{2}$, and where $W(t)$ is a standard Brownian motion.

The model (6) generalizes the basic Bass (1969) representation in two directions. First, it amounts to a stochastic generalization, as it incorporates a continuous-time error term with standard deviation proportional to $\sigma n(t)^{\gamma}$. Hence, with $\gamma \neq 0$ this implies a level-dependent heteroskedastic error term. Secondly, the model amounts to a dynamic generalization. The drift term implies that $n(t)$ "mean-reverts" to either $\bar{n}(t)$ or $n^*(t)$, which is its value under the non-stochastic differential equation (2) or its empirical counterpart (5). This mean reversion occurs with speed determined by α. When α is close to zero, there is almost no mean reversion, and when $\alpha \to \infty$ this speed is infinitely fast. Note that the basic deterministic Bass model is obtained when $\alpha \to \infty$ and $\sigma \to 0$.

The volatility $\sigma^{\gamma} n(t)$ guarantees that the random variation vanishes as $n(t) \to 0$, and hence that the process $n(t)$ remains non-negative. Here γ will not be treated as an unknown parameter to be estimated, but as a choice variable that is set by the researcher. In the remainder of our theoretical analysis, we will focus on the case $\gamma = 1$, but we should stress here that the choice for $\gamma = \frac{1}{2}$ or $\gamma = 2$ can be considered too, as long as $\gamma \geq \frac{1}{2}$, which is needed to guarantee non-negativity (as discussed further below). In practice, one may decide to use diagnostic tests for heteroskedasticity, to see which variant suits the data best. For example, in the illustration below, we will find that a volatility specification of $\sigma\sqrt{n(t)}$ is a good choice.

The specification (6a) is inspired by a class of stochastic processes that is often used in the mathematical finance literature for interest rates, see for example Hull and White (1990). In particular, with $\gamma = 0$ the model corresponds to the Vasicek (1977) model with time-varying mean $\bar{n}(t)$ (which does not exclude negative realizations), whereas for $\gamma = \frac{1}{2}$ and $\gamma = 1$ it leads to the Cox-Ingersoll-Ross (Cox *et al.*, 1985) and Brennan and Schwartz (1982) models, respectively. It has been shown in this literature that when $\gamma = 1$, the process is strictly positive. For $\gamma = \frac{1}{2}$ it is non-negative, and strictly positive under the condition $\alpha\bar{n}(t) \geq \frac{1}{2}\sigma^2$; since $\bar{n}(t)$ converges to zero as t increases, this condition will eventually be violated, which means that $n(t)$ eventually reaches zero for this choice of γ.

The alternative specification (6b) leads, upon substitution of (5) in (6b), and under the restriction $\gamma = 1$, to the following process:

$$dn(t) = \alpha \left(p[m - N(t)] + \frac{q}{m} N(t)[m - N(t)] - n(t) \right) dt + \sigma n(t) dW(t). \tag{7}$$

Observing that $N(t) = \int_0^t n(s)ds$, we see that the model (7) is actually a stochastic differential-integral equation. We prefer to work with this specification, because it replaces the exogenously given mean adoption rate function $\bar{n}(t)$ by the endogenous target adoption rate $n^*(t)$. Moreover, its discretization, discussed below, leads to a more convenient estimation procedure than the model with $\bar{n}(t)$. We do not address the questions of existence and uniqueness of a solution to (7) explicitly in this paper, but simulation evidence reveals that realizations from the two specifications closely resemble each other, with the $n^*(t)$ specification leading to somewhat more random variation for the same value of σ. From these simulations, we expect most of the properties of (6a) to carry over to (6b), and in particular we expect the process (6b) to be non-negative for $\gamma \geq \frac{1}{2}$.

One aspect that can be immediately derived from both (6a) and (6b) is that $N(t)$ in both cases increases monotonically from $N(0) = 0$ to $\lim_{t\to\infty} N(t) = m$. That it increases monotonically follows from non-negativity of $n(t) = dN(t)/dt$. Furthermore, that it eventually converges to m follows because $n(t)$ eventually has to converge to zero, and this can only occur when the drift in (6a) or (6b) is zero, so that in particular $\lim_{t\to\infty} n^*(t) = 0$, which requires $\lim_{t\to\infty} N(t) = m$.

To give an impression of what data from (7) look like, consider the graphs in Figure 2. It contains two realizations of the process with $\alpha = 5$, $\sigma = 0.5$, and the remaining parameters the same as in Figure 1. These figures are obtained by simulating a discrete Euler-type approximation of the continuous-time model (7), discussed below, using a very fine partition. We observe that the cumulative number of adoptions $N(t)$ follows the familiar smooth S-shaped curve, which however deviates randomly from its mean $\bar{N}(t) = mF(t)$, with the largest deviations occuring around the inflection point. Correspondingly, the sales data X_i clearly show mean-reversion to $\bar{n}(t) = mf(t)$, with the largest random variation around the peak sales.

Insert Figure 2 about here

The discretize the model, we use the standard Euler approach, which for (7) results in

$$\begin{aligned} n(t_i) - n(t_{i-1}) &\approx \alpha\left(p[m - N(t_{i-1})] + \frac{q}{m}N(t_{i-1})[m - N(t_{i-1})] - n(t_{i-1})\right)[t_i - t_{i-1}] \\ &\quad + \sigma n(t_{i-1})\left[W(t_i) - W(t_{i-1})\right] \\ &= \alpha\left(p(m - N_{i-1}) + \frac{q}{m}N_{i-1}(m - N_{i-1}) - n(t_{i-1})\right)\delta + n(t_{i-1})\varepsilon_i, \end{aligned}$$

where $\delta = t_i - t_{i-1}$, and

$$\varepsilon_i = \sigma\left[W(t_i) - W(t_{i-1})\right] \sim \text{i.i.d. } N(0, \sigma^2\delta), \tag{8}$$

as the increments $W(t_i) - W(t_{i-1})$ are independently and identically distributed with mean zero and variance $(t_i - t_{i-1}) = \delta$.

The approximation of the sales X_i over the interval $(t_{i-1}, t_i]$, is given by

$$\begin{aligned} X_i = N(t_i) - N(t_{i-1}) &= \int_{t_{i-1}}^{t_i} n(t)dt \\ &\approx n(t_i)(t_i - t_{i-1}) = n(t_i)\delta. \end{aligned}$$

This implies that

$$\begin{aligned} X_i - X_{i-1} &\approx \delta[n(t_i) - n(t_{i-1})] \\ &\approx \delta\alpha\left(p(m - N_{i-1}) + \frac{q}{m}N_{i-1}(m - N_{i-1}) - \frac{X_{i-1}}{\delta}\right)\delta + \delta\frac{X_{i-1}}{\delta}\varepsilon_i \\ &= \alpha\delta^2 p(m - N_{i-1}) + \alpha\delta^2\frac{q}{m}N_{i-1}(m - N_{i-1}) - \alpha\delta X_{i-1} + X_{i-1}\varepsilon_i. \end{aligned}$$

In summary, the discretization gives

$$\Delta X_i = \beta_1 + \beta_2 N_{i-1} + \beta_3 N_{i-1}^2 + \beta_4 X_{i-1} + X_{i-1}\varepsilon_i, \tag{9}$$

with $\varepsilon_i \sim$ i.i.d. $N(0, \sigma^2\delta)$, and where $\beta_4 = -\alpha\delta$, and

$$\begin{aligned} \beta_1 &= pm\alpha\delta^2 = -\beta_4\delta pm, \\ \beta_2 &= \alpha\delta^2(q - p) = -\beta_4\delta(q - p), \\ \beta_3 &= \frac{-q\alpha\delta^2}{m} = \beta_4\delta\frac{q}{m}. \end{aligned}$$

The discrete-time model (9) immediately suggests a maximum likelihood estimation procedure. Note that this reparametrized Euler approximation encompasses the case with $\beta_4 = -1$, corresponding to immediate adjustment of X_i to its equilibrium path. This submodel is a straightforward extension of the Bass linear regression considered below, using a particular heteroskedastic specification of the error term. On the other hand, in the underlying continuous model (7), immediate adjustment is only possible in the limiting case $\alpha \to \infty$. The fact that $\beta_4 = -\alpha\delta$ suggests that β_4 can be less than or equal to -1 must be seen as an approximation error. However, we expect this approximation to mainly affect inference on the adjustment parameter α, and much less on the parameters of interest (m, p, q).

We now should say a few words about identification of the parameters. As is clear from (9), there are four β parameters, corresponding to the four underlying parameters (m, p, q, α), given a choice for δ. Note that δ represents the observation frequency. Assuming that t is measured in years, annually observed data correspond to $\delta = 1$, quarterly data to $\delta = \frac{1}{4}$, and monthly data to $\delta = \frac{1}{12}$. This means that δ is not an unknown parameter, but a known property of the data. This allows recovering the parameters of interest (m, p, q) together with the speed of adjustment α, from the four β parameters, which may be estimated by maximum likelihood as discussed in the next section.

2.3 Alternative representations

The most commonly considered empirical representations of the Bass diffusion theory are proposed in Bass (1969) and Srinivasan and Mason (1986).

In Bass (1969) it is proposed to consider ordinary least-squares estimation of the difference equation corresponding to (2), that is,

$$\begin{aligned} X_i &= p(m - N_{i-1}) + \frac{q}{m} N_{i-1}(m - N_{i-1}) + \varepsilon_i \\ &= \alpha_1 + \alpha_2 N_{i-1} + \alpha_2 N_{i-1}^2 + \varepsilon_i, \end{aligned} \tag{10}$$

where ε_i is assumed to be an independent and identically distributed (i.i.d.) error term with mean zero. Note that (p, q, m) must be obtained from $(\alpha_1, \alpha_2, \alpha_3)$. An alternative approach is to apply non-linear least squares directly. Note also that these values of p and q correspond to those of the continuous-time model (2) only if $\delta = t_i - t_{i-1} = 1$. In other instances, these parameters have to be rescaled, see also Putsis (1996). Clearly, our model adds a regressor on the right-hand side and also modifies the error term.

Srinivasan and Mason (1986) recognize that the Bass (1969) formulation above may introduce aggregation bias, as X_i is simply taken as the discrete representative of $n(t)$. Therefore, these authors propose to apply nonlinear least-squares to their representation, that is,

$$X_i = m[F(t_i;\theta) - F(t_{i-1};\theta)] + u_i, \tag{11}$$

with u_i assumed to be i.i.d error term. This method is quite elegant, although parameter estimation might not be easy as it involves estimating p and q jointly from ratios like $\frac{q}{p}$ and sums $p+q$, and this might make the estimation process unstable. Furthermore, again it is assumed that there are homoskedastic errors.

Below we will compare these two alternative estimation procedures with the estimators resulting from our model, which are analyzed next.

3 Estimation and inference

In this section we discuss the estimation of the parameters in our new representation. As it will turn out that standard asymptotic properties does not hold, nor would it for the standard Bass (1969) model, we resort to simulations to see how large the differences are between standard normal variates and our estimators divided by their standard errors.

3.1 Estimation

Suppose that we have observations $\{N_i = N(\delta i), i = 0, \ldots, T\}$, and its increments X_i. The (quasi-) maximum likelihood estimators of the parameters (β_i, σ^2) of the discretization (9) can be obtained simply by weighted least squares (WLS), with $1/X_{i-1}$

acting as weights. Thus, letting $Y_i = \Delta X_i / X_{i-1}$, $Z_i = (1, N_{i-1}, N_{i-1}^2, X_{i-1})/X_{i-1}$, and $\beta = (\beta_1, \beta_2, \beta_3, \beta_4)'$, we find

$$\hat{\beta} = \left(\sum_{i=1}^{T} Z_i Z_i' \right)^{-1} \sum_{i=1}^{T} Z_i Y_i, \qquad \hat{\sigma}^2 = \frac{1}{\delta T} \sum_{i=1}^{T} (Y_i - \hat{\beta}' Z_i)^2. \tag{12}$$

From β, we may obtain estimates of the parameters (m, p, q) and the rate of adjustment α from

$$\hat{m} = \frac{-\hat{\beta}_2 - \sqrt{\hat{\beta}_2^2 - 4\hat{\beta}_1\hat{\beta}_3}}{2\hat{\beta}_3}, \qquad \hat{p} = -\frac{\hat{\beta}_1}{\delta \hat{\beta}_4 \hat{m}}, \qquad \hat{q} = \frac{\hat{m}\hat{\beta}_3}{\delta \hat{\beta}_4}, \qquad \hat{\alpha} = -\frac{\hat{\beta}_4}{\delta}. \tag{13}$$

Note that the estimates of p, q and α involve the sampling frequency δ, which is a known constant as discussed in the previous section. Observe also that the estimate of m does not depend on δ, nor on the parameter β_4 characterizing the speed of adjustment. Asymptotic standard errors of the parameters could be obtained, using the delta method, from $\hat{\mathrm{V}}[\hat{\beta}] = \delta \hat{\sigma}^2 \left(\sum_{i=1}^{T} Z_i Z_i' \right)^{-1}$. Whether these standard errors may be used to obtain asymptotic t-tests and confidence intervals in the usual fashion depends on the asymptotic properties of the estimators, is considered next.

3.2 Asymptotic properties

As T increases and δ is kept fixed, we have seen in the previous section that $N_T \to m$, so that $X_T \to 0$. This implies that as $T \to \infty$, the parameter m may be estimated without error. On the other hand, when N_T has essentially reached m, it is clear that the information on the other parameters p, q and α will no longer increase, and this implies that the estimators of p and q in (12)–(13) will not be consistent. However large T is, different realizations of $\{N_i\}_{i=0}^{T}$ will lead to different realizations of the estimators. This in turn means that we cannot validate the conventional use of asymptotic standard errors for $\hat{p}$ and $\hat{q}$, since the proof of asymptotic normality builds on consistency of the relevant estimators.

As an alternative approach, suppose that we fix the time span $[0, S]$ for the continuous time parameter t, and we let the time interval $\delta = S/T$ vary, such that $\delta \to 0$ as $T \to \infty$.

Then we may characterize the limiting behavior of sample moments such as $\sum_{i=1}^{T} Z_i Z_i'$ and $\sum_{i=1}^{T} Z_i \varepsilon_i$ by the corresponding integrals of the continuous time processes. This leads to the following results for the estimator $\hat{\beta}$ and its estimated covariance matrix.

Theorem 1 *Let $Z(t) = n(t)^{-1}[1, N(t), N(t)^2, n(t)]'$ and $D_\delta = diag(\delta^2, \delta^2, \delta^2, \delta)$. Then, as $T \to \infty$ and $\delta \to 0$, with S fixed,*

$$D_\delta^{-1}(\hat{\beta} - \beta) = \begin{pmatrix} \delta^{-2} \begin{pmatrix} \hat{\beta}_1 - \beta_1 \\ \hat{\beta}_2 - \beta_2 \\ \hat{\beta}_3 - \beta_3 \end{pmatrix} \\ \delta^{-1}(\hat{\beta}_4 - \beta_4) \end{pmatrix} \xrightarrow{d} \sigma \left(\int_0^S Z(t)Z(t)'dt \right)^{-1} \int_0^S Z(t)dW(t), \quad (14)$$

and

$$D_\delta^{-1}\hat{\mathrm{V}}[\hat{\beta}]D_\delta^{-1} \xrightarrow{d} \sigma^2 \left(\int_0^S Z(t)Z(t)'dt \right)^{-1}. \quad (15)$$

The proof of this theorem is given in the Appendix. The limiting distribution obtained here is expressed as a functional of continuous-time processes, which themselves are functionals of a Brownian motion process. Such expressions are very common in the econometrics literature on unit roots and cointegration, see for example Hamilton (1994). An important result from this literature is that if $Z(t)$ would be independent of $W(t)$, then the limiting expression in (14) would have a so-called *mixed normal* distribution, which would enable the construction of asymptotic standard normal t-ratios on β. However, this independence is violated here, which means that such t-ratios can have a non-normal asymptotic distribution. In principle, this distribution will depend on all parameter (p, q, m, α) as well as S. In practice we can set S to 1 by redefining the units of time measurement, but then the dependence on the unknown parameters remains.

The limiting behavior of $(\hat{m}, \hat{p}, \hat{q}, \hat{\alpha})$ may be obtained from (13). Here, one has to take into account that the true value of β is really a sequence satisfying $(\beta_1, \beta_2, \beta_2) = \mathrm{O}(\delta^2)$ and $\beta_4 = \mathrm{O}(\delta)$. Together with (14), this implies that $\delta^{-2}(\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3) = \mathrm{O}_p(1)$ and $\delta^{-1}\hat{\beta}_4 = \mathrm{O}_p(1)$, which in combination with (13) implies that $(\hat{m}, \hat{p}, \hat{q}, \hat{\alpha})$ are all $\mathrm{O}_p(1)$, and hence do not converge in probability to their true values for fixed S. In other words, the parameters of the Bass model cannot be estimated consistently based on a fixed time span S, even if the sampling frequency δ goes to zero. This is a common phenomenon when estimating the parameters of a continuous-time model.

In general, one expects consistency and asymptotic normality as $S \to \infty$. However, in this particular case one may definitely expect $\hat{m}$ to be consistent as $S \to \infty$, but for the other parameters this might not be possible for the same reasons given earlier. At some point, $n(t)$ has converged to zero, and extending the sample period will not lead to any new information. An important consideration is that as $S \to \infty$, $n(S) \to 0$ and hence $n(S)^{-1}$ explodes. This means that as the process reaches its saturation point, the information on the parameters should increase considerably, since it involves a factor $n(t)^{-1}$. In the next subsection, we will use Monte Carlo simulation methods to investigate the asymptotic behavior of the estimators and their standard errors as S increases.

The result that the parameters p and q in the Bass model cannot be estimated consistently, even as the sample period increases, may shed some new lights on the simulation results reported in Van den Bulte and Lilien (1997). These authors study the effect of extending the sample period on the bias of the Srinivasan-Mason (SM) estimator. The interpretation of their simulation results is based on the premise that the SM estimator, as any non-linear least squares estimator under conventional circumstances, is consistent as the sample size increases or as the error variance decreases to zero. Although consistency will indeed be obtained as $\sigma \to 0$, the results here indicate that consistency as the sample size increases is very doubtful, because the information no longer increases as the sample size is extended beyond the saturation point.

3.3 Monte Carlo simulation of asymptotic properties

The theoretical results in the previous subsection indicate that both consistency and asymptotic normality of the estimators of the parameters m, p and q cannot be proved analytically, even as both the sampling frequency δ goes to zero and the sample period

S goes to infinity, although the latter is sufficient for consistency of $\hat{m}$. On the other hand, we have obtained a characterization of the limiting distribution of the reduced-form parameter estimator $\hat{\beta}$ and its estimated covariance matrix. In this subsection, we study the properties of the implied limiting distribution of $(\hat{m}, \hat{p}, \hat{q}, \hat{\alpha})$, for various values of S, using Monte Carlo simulation of the limiting integral expressions in Theorem 1. That is, we simulate realizations from this limiting distribution, and then transform these to realizations of $(\hat{m}, \hat{p}, \hat{q}, \hat{\alpha})$ and its estimated standard error. Using these simulations, we hope to answer the following questions:

1. Do the estimator biases and variances decrease to zero as S increases?

2. Do the standardized estimators (that is, t-ratios) have an approximate standard normal distribution as S increases?

In all cases we use a relatively fine discretization of $\delta = 0.01$ to simulate an approximation of the continuous-time processes and stochastic integrals. All numerical results in this section and the next have been obtained using Ox version 3.1, see Doornik (2001).

In order to address the question of (lack of) consistency as S increases, we consider first the behaviour of the asymptotic ($\delta \to 0$) standard errors of $\hat{m}$, $\hat{p}$, $\hat{q}$ and $\hat{\alpha}$, as S increases. Figure 3 plots the average over 1000 Monte Carlo replications of these asymptotic standard errors, expressed as fraction of the true value of the corresponding parameter. Here the true values are chosen as $(p, q, \alpha, \sigma) = (0.01, 0.5, 5, 0.5)$. The results are invariant to m, which is here set at 1, and other parameter combinations lead to (unreported) qualitatively similar patterns. For this parameter configuration, the inflection point lies around 8, and the process has reached its maturity level at about $t = 25$. This leads us to depict the behavior of the relative standard errors for S ranging from 8 to 32.

Insert Figure 3 about here

From Figure 3 we clearly see that the standard error of $\hat{m}$ has essentially converged to zero for $S \geq 20$. However, the standard errors of the other three parameters indicate that these parameters are still estimated with some uncertainty when $S = 32$, after which no new information on these parameters can be expected. This is most evident for the parameter p. After about 15 years the standard error of $\hat{p}$ virtually remains the same at about 10% of its true value, indicating that no additional information on the parameter is obtained from larger samples. For the other two parameters q and α, the standard errors seem to decrease somewhat at $S = 32$, but only very slowly. Bearing in mind that the process has already reached its maturity level at $S = 32$, this again indicates that these parameters cannot be estimated without error, even when the full diffusion process has been observed.

These properties are confirmed by the simulated distributions of the parameter estimates in deviation from their true values, depicted in Figure 4. Here we use the same parameter combination as before, with the estimated densities based on 10000 Monte Carlo replications. As expected, the density of $\hat{m}$ quickly concentrates around the true value as S increases, to such an extent that the densities for $S = 16$ and $S = 32$ could not be conveniently displayed. In contrast, the distribution of $\hat{p}$ hardly changes as S increases beyond 10, although the densities seem to become slightly better centered around the true value. The densities of $\hat{q}$ and $\hat{\alpha}$ do become more concentrated as S increases, but it is important to note that for $S = 32$, both estimates still display substantial variation, even though the process has converged to its maturity level.

Insert Figure 4 about here

Figure 5 considers the distribution of t-statistics corresponding to four parameters, for the same parameter values. With the possible exception of α, the densities of the t-statistics seem to be fairly close for S large enough. Therefore, although the theoretical analysis in the previous sub-section does not explicitly indicate asymptotic normality, the simulation results in this subsection indicate that the standard normal distribution may be used as an approximation for t-statistics based on $\hat{m}$, $\hat{p}$ and $\hat{q}$. Although we have only demonstrated this for one particular parameter value, unreported simulations indicate that this result holds more generally, at least for the cases where the inflection point is within the estimation sample.

Insert Figure 5 about here

4 Comparing model representations

In this section we evaluate the empirical performance of the representations of Bass (1969) and of Srinivasan and Mason (1986) [abbreviated as SM], when the data generating process is our newly developed model, which we label here as BF1. We also incorporate the new model without the autoregressive variable, and label this as BF2. We examine the consequences of assuming immediate adjustment and of neglecting heteroskedasticity. As the two alternative empirical models lack autoregressive dynamics, we may expect that this omission leads to biased estimates of all parameters. It is however not immediately clear which parameter would be affected most. Next, we would expect that the erroneous assumption of homoskedasticity would lead to empirical distributions which are not close to the normal distribution.

Insert Figures 6, 7, 8 and 9 about here

The DGP is again the same as before, as well as the Monte Carlo setup, where we now consider $S = 12$ and $S = 18$, and we vary over $\delta = 0.1$ and $\delta = 1$. The cases with $\delta = 1$ can be viewed as concerning aggregated data. It is unclear to what extent the exclusion of a relevant variable hurts the models with also homoskedastic errors, but one might expect that aggregation makes such an exclusion less harmful. This can seen by comparing the graphs for BF1 and BF2 when δ increases from 0.1 to 1.

The results in Figure 6 concern the finite-sample distributions of the BF1, BF2, Bass and SM in case $S = 12$ and $\delta = 0.1$. In words, this case would concern rather disaggregated data, where the inflection point has been passed just recently. The first conclusion from these graphs is that the parameter estimates obtained from the BF2, Bass and SM models show substantial bias, whereas the estimates from the BF1 model, as expected given the simulation results in the previous section, do rather well. For example, the mean estimated value for p from the BF1 model would be around 0.01, whereas that of the SM model equals about 0.0075. The second conclusion (from the right-hand side panel of Figure 6) is that the empirical distributions of the non-DGP models are, effectively speaking, all over the place. Hence, standard confidence intervals of the BF2, Bass and SM models can not be trusted at all, at least when the data correspond with the BF1 model.

When we compare the graphs in Figure 6 with those in Figure 7, where the only change is that we go from $\delta = 0.1$ to $\delta = 1$, we observe that the earlier seen overwhelming differences between the various models becomes smaller. There still is substantial bias in the key parameters for the non-DGP models, and also do the distributions of the $t-$ratios become a little more closer to the standard normal distribution, but differences seem to be smaller. This suggests that neglecting a variable and a time-varying error process might be less harmful for more aggregated data. The same type of conclusion can be drawn from comparing Figure 8 with Figure 9.

Upon comparing Figure 6 with Figure 8, where $S = 18$ and $\delta = 0.1$, we see that the differences between the empirical distributions for the estimated parameters become smaller. This suggests that, as the diffusion process proceeds, the point estimates of the various empirical models get closer to each other, where it should be mentioned that the largest differences across models persist for the parameter p. Given the discussion in the previous section about consistency, we should not expect that the empirical distribution of the t-ratios would change much with more data. This is indeed confirmed by the graphs in the right-hand panel of Figure 8.

A final remark about the empirical models would concern the performance of the SM model. It is not easy to explicitly indicate how the omitted variable and the neglected heteroskedasticity emerge in this highly non-linear model, but the simulation results do suggest that this method is rather unstable, see for example Figure 9 and in particular for m. Perhaps these simulation results shed additional light on similar kinds of findings in Van den Bulte and Lilien (1997).

5 Illustration

We illustrate various issues discussed above for a series of 74 monthly observations concerning the number of ATMs in the Netherlands. The main focus is on how one can make decisions in practice. The data range from March 1990 to April 1996. We decide to use this series, which contains more observations than one typically has when considering Bass-type diffusion models, as the graphs of this series obviously display the heteroskedastic features we emphasized before. The graph of the monthly number of newly installed machines is given in Figure 10, and the graph of the cumulative number appears in Figure 11. It seems that the inflection point has just been passed at the end of the observation sample. Also, it is clear that the uncertainty around the installed series is not constant over time, being much less in the beginning of the sample and rather high towards the end of the sample.

Insert Figures 10 and 11 about here

We start with fitting the BF1 model, where we use EViews version 4.1, and then we examine the estimation results for the other models, also considered in the simulations before. We fix δ as 1. We first need to decide on the type of heteroskedasticity. To this end, we fit the BF1 model with homoskedastic errors, and focus on the properties of the estimated residuals. When we regress the squares of these residuals on X_i and X_i^2, we obtain the highest and significant t-ratio for X_i, and hence we take the error term as $\sqrt{X_i}\varepsilon_i$. The estimation results appear in the top panel and the first column of Table 1.

Insert Table 1 about here

These estimation results suggest that β_4 is not equal to -1, and hence the autoregressive terms does matter. If this is truly the case, one might expect that the BF2, Bass and

SM models, which all neglect this term, have estimated residuals with serial correlation. Such correlation could be detected by the familiar Durbin-Watson [DW] test. And indeed, the results in the fifth row of Table 1 show that the DW test obtains a comforting value for the BF1 model, whereas it suggests neglected autoregressive dynamics in the other three models. Such neglected dynamics also have consequences for the estimated parameters p, q, and to a lesser extent for m, as we have seen in the previous section. When we compare the estimation results for these parameters across models, we can observe substantial differences across models. Note by the way, that these differences also imply differences across the estimated locations of peak sales, or inflection point.

When the BF1 model would indeed provide the most accurate description of the data, then the neglecting of heteroskedasticity by the Bass and SM model could perhaps be diagnosed by tests for heteroskedasticity. Examples of those tests are the LM type test of White and the test for ARCH. Application of these farily standard tests gives the outcomes in the rows underneath the DW test results in Table 1. Clearly, the test statistics for the Bass and SM models are highly significant, and clearly indicate that these models would be inadequate for this dataset.

Finally, if the BF1 model would indeed provide the best fit to the data at hand, one might expect model selection criteria to favor this model. When we consider the values of the familiar Akaike Information Criterion [AIC] in the last row of Table 1, which by the way should be minimized, we see that the BF1 model is also preferred using this criterion.

6 Concluding remarks

In this paper we have introduced a new model representation for the Bass diffusion theory, which in a sense comes closest to the basic premises of this theory. Taking this theory along, we ended up with a representation which extends the original Bass model with an autoregressive variable. Additionally, we argued that the theory should involve an empirical model which allows uncertainty about the diffusion path to be small in the beginning as well as at the end of the process. We analyzed the econometric properties of this new model, and we found that its parameters cannot be consistently estimated. As our model encompasses the original Bass model, this finding also holds for this model. Through simulation experiments we aimed to see what this inconsistency means for sample data typically encountered in practice. The main conclusion was that, provided that the inflection point is in the sample period, one might rely on the standard normal distribution for the $t-$ratios as an approximation. Again using simulations, we also examined the consequences of neglecting the autoregressive variable and the heteroskedasticity. We found that these consequences can be very large for the point estimates of the key parameters as well as for the distributions of their $t-$ratios. In other words, if our new model fits the data well, the estimates and their standard errors obtained from the Bass (1969) and Srinivasan and Mason (1986), are far from reliable, and in some cases nonsensical. On the positive side, a reassuring result of our simulations is that the differences across models for highly aggregated data (think of, 10 years of annual data) would not be large, but again provided the inflection point is in the sample.

There are various avenues for further research. The key parameters in the Bass theory determine the location of the inflection point and the size of peak sales. It would be interesting to examine how the various models perform for these parameters. Further, we plan to derive the one- and multi-step ahead forecasts for the new model, as well as their associated confidence bounds. As the model is rather complicated in terms of serial dependence, we expect that simulation methods will be needed. Finally, we plan to illustrate the new model for many real-life data series, in order to see whether it really also provides an empirical improvement over currently fashionable models.

Appendix

Proof of Theorem 1. To prove the theorem, we need to obtain the limiting behavior of the sample moments $\sum_{i=1}^{T} Z_i Z_i'$ and $\sum_{i=1}^{T} Z_i \varepsilon_i$. It will be convenient to express these sums, after appropriate standardization, as integrals. Let $\lfloor x \rfloor$ is the integer part of x, and define

$$Z_\delta(t) = \delta^{-1} D_\delta Z_{\lfloor t/\delta \rfloor + 1} = \frac{1}{\delta^{-1} X_{\lfloor t/\delta \rfloor + 1}} \begin{pmatrix} 1 \\ N_{\lfloor t/\delta \rfloor + 1} \\ N^2_{\lfloor t/\delta \rfloor + 1} \\ \delta^{-1} X_{\lfloor t/\delta \rfloor + 1} \end{pmatrix},$$

and

$$W_\delta(t) = \sigma^{-1} \sum_{i=1}^{\lfloor t/\delta \rfloor} \varepsilon_i = W(\delta \lfloor t/\delta \rfloor).$$

Then we have

$$\begin{aligned} \delta^{-1} D_\delta \sum_{i=1}^{T} Z_i Z_i' D_\delta &= \sum_{i=1}^{T} Z_\delta(t_{i-1}) Z_\delta(t_{i-1})' \delta \\ &= \sum_{i=1}^{T} \int_{t_{i-1}}^{t_i} Z_\delta(t) Z_\delta(t)' dt = \int_0^S Z_\delta(t) Z_\delta(t)' dt, \end{aligned}$$

and similarly

$$\delta^{-1} D_\delta \sum_{i=1}^{T} Z_i \varepsilon_i = \sigma \sum_{i=1}^{T} Z_\delta(t_{i-1}) [W_\delta(t_i) - W_\delta(t_{i-1})] = \sigma \int_0^S Z_\delta(t) dW_\delta(t).$$

Consider

$$\frac{X_{\lfloor t/\delta \rfloor + 1}}{\delta} = \frac{N(\delta \lfloor t/\delta \rfloor + \delta) - N(\delta \lfloor t/\delta \rfloor)}{\delta}.$$

It is clear that as $\delta \to 0$, $\delta \lfloor t/\delta \rfloor \to t$, and hence $X_{\lfloor t/\delta \rfloor + 1}/\delta \to n(t)$. Similarly, with $N_{\lfloor t/\delta \rfloor + 1} = N(\delta \lfloor t/\delta \rfloor + \delta)$ we have $N_\delta(t) \to N(t)$ as $\delta \to 0$. Combining these two results leads to

$$Z_\delta(t) \xrightarrow{d} Z(t),$$

where the convergence in distribution applies in $D[0,S]$, the space of processes on $[0,S]$ that are right-continuous and have finite left limits. Analogously, it follows that $W_\delta(t) \xrightarrow{d} W(t)$, jointly with the convergence of $Z_\delta(t)$ to $Z(t)$. Applying the continuous mapping theorem, this gives

$$\delta^{-1} D_\delta \sum_{i=1}^{T} Z_i Z_i' D_\delta = \int_0^S Z_\delta(t) Z_\delta(t)' dt \xrightarrow{d} \int_0^S Z(t) Z(t)' dt, \tag{16}$$

$$\delta^{-1} D_\delta \sum_{i=1}^{T} Z_i \varepsilon_i = \sigma \int_0^S Z_\delta(t) dW_\delta(t) \xrightarrow{d} \sigma \int_0^S Z_\delta(t) dW(t). \tag{17}$$

The second result actually cannot be proved by the continuous mapping theorem only, but needs additional results about convergence to the solution of a stochastic differential equation, see Hansen (1992).

From (16)–(17) we find

$$\begin{aligned} D_\delta^{-1}(\hat{\beta} - \beta) &= \left(\delta^{-1} D_\delta \sum_{i=1}^{T} Z_i Z_i' D_\delta \right)^{-1} \delta^{-1} D_\delta \sum_{i=1}^{T} Z_i \varepsilon_i \\ &\xrightarrow{d} \sigma \left(\int_0^S Z(t) Z(t)' dt \right)^{-1} \int_0^S Z_\delta(t) dW(t). \end{aligned}$$

Note that $\hat{\beta}$ is a consistent estimator. The convergence rate is $\mathrm{O}_p(\delta^2) = \mathrm{O}_p(T^{-2})$ for the first three components, and $\mathrm{O}_p(\delta) = \mathrm{O}_p(T^{-1})$ for the fourth component (describing the autoregressive dynamics). This in turn may be used to show that $\hat{\sigma}^2 \xrightarrow{p} \sigma^2$, and hence

$$D_\delta^{-1} \hat{\mathrm{V}}[\hat{\beta}] D_\delta^{-1} = \hat{\sigma}^2 \left(\delta^{-1} D_\delta \sum_{i=1}^{T} Z_i Z_i' D_\delta \right)^{-1} \xrightarrow{d} \sigma^2 \left(\int_0^S Z(t) Z(t)' dt \right)^{-1}.$$

References

Bass, F.M. (1969), "A new product growth model for consumer durables", *Management Science*, 15, 215–227.

Brennan, M.J. and E.S. Schwartz (1982), "An equilibrium model of bond pricing and a test of market efficiency", *Journal of Financial and Quantitative Analysis*, 17, 301–329.

Cox, J.C., J.E. Ingersoll and S.A. Ross (1985), "A theory of the term structure of interest rates", *Econometrica*, 53, 385–467.

Doornik, J.A. (2001), *Ox: An Object-Oriented Matrix Programming Language.* London: Timberlake Consultants Ltd. and http://www.nuff.ox.ac.uk/Users/Doornik.

Hamilton, J.D. (1994), *Time Series Analysis.* Princeton: Princeton University Press.

Hansen, B.E. (1992), "Convergence to stochastic integrals for dependent heterogeneous processes", *Econometric Theory*, 8, 489–500.

Hull, J. and A. White (1990), "Pricing interest-rate derivative securities", *The Review of Financial Studies*, 3, 573–592.

Putsis, W.P. (1996), "Temporal aggregation in diffusion models of first-time purchase: Does choice of frequency matter?", *Technological Forecasting and Social Change*, 51, 265–279.

Putsis, W.P. and V. Srinivasan (2000), "Estimation techniques for macro diffusion models", Chapter 11 in V. Mahajan, E. Muller and Y. Wind (eds.) *New-Product Diffusion*

Models, Boston: Kluwer, 263–291.

Srinivasan, V. and C.H. Mason (1986), "Nonlinear least squares estimation of new product diffusion models", *Marketing Science*, 5, 169–178.

Van den Bulte, C. and G.L. Lilien (1997), "Bias and systematic change in the parameter estimates of macro-level diffusion models", *Marketing Science*, 16, 338–353.

Vasicek, O.A. (1977), "An equilibrium characterization of the term structure", *Journal of Financial Economics*, 5, 177–188.

Table 1: Estimation results for the new Bass model (BF1), the same model with α set equal to 1 (BF2), the original Bass (1969) model, and the representation of Srinivasan and Mason (1986)(SM). The sample for all models ranges from May 1990 to April 1996. Standard errors are given in parentheses.

	BF1	BF2	Bass	SM
m	142064.3 (45502.99)	142052.5 (23644.90)	142060.8 (18048.95)	142060.1 (18806.93)
p	0.000406 (0.000350)	0.0000978 (0.000148)	0.001081 (0.000559)	0.000730 (0.000170)
q	0.072660 (0.015874)	0.066406 (0.007353)	0.064948 (0.007221)	0.064831 (0.007761)
$\beta_4 + 1$	0.425568 (0.137421)			
DW	2.117	1.278	1.488	1.513
$LM - ARCH(1)$	0.001	< 0.001	7.300	7.951
$LM - White$	2.122	4.097	22.638	11.791
AIC	13.872	14.074	15.007	14.975

Figure 1: Simulation from the Bass hazard rate model, with $m = 100000$, $p = 0.01$, and $q = 0.15$.

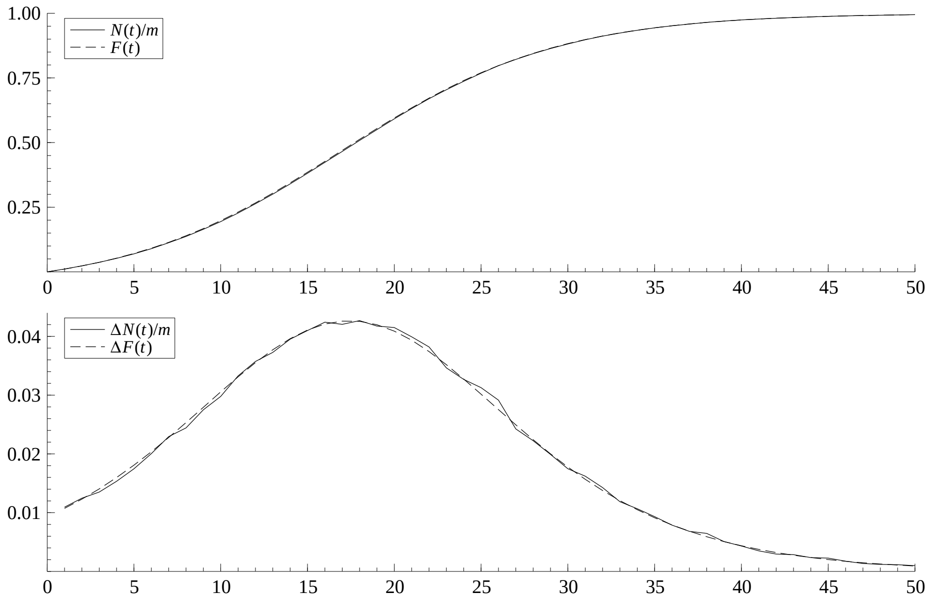


Figure 2: Simulations from (7), with $p = 0.01$, $q = 0.15$, $\alpha = 5$, and $\sigma = 0.5$.

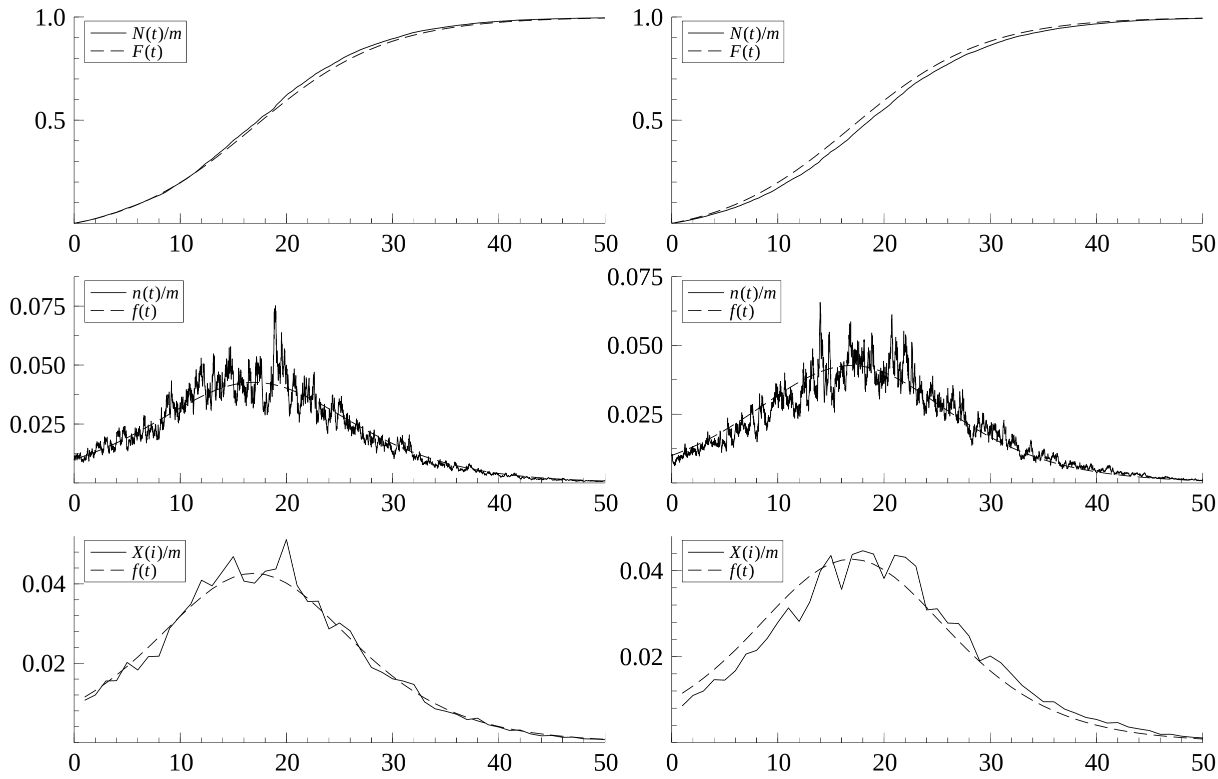


Figure 3: Asymptotic relative standard errors of $\hat{m}$,$\hat{p}$, $\hat{q}$ and $\hat{\alpha}$, as a function of the sample period S. Key parameters in the DGP are $(p, q, \alpha, \sigma) = (0.01, 0.5, 5, 0.5)$.

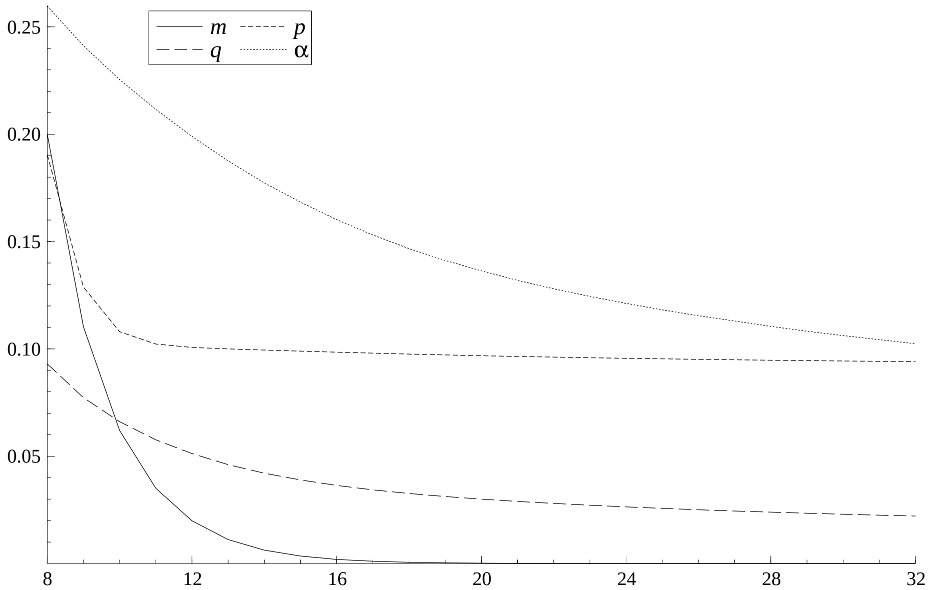


Figure 4: Asymptotic distributions of $\hat{m} - m$, $\hat{p} - p$, $\hat{q} - q$ and $\hat{\alpha} - \alpha$, for various values of S. Key parameters in the DGP are $(p, q, \alpha, \sigma) = (0.01, 0.5, 5, 0.5)$.

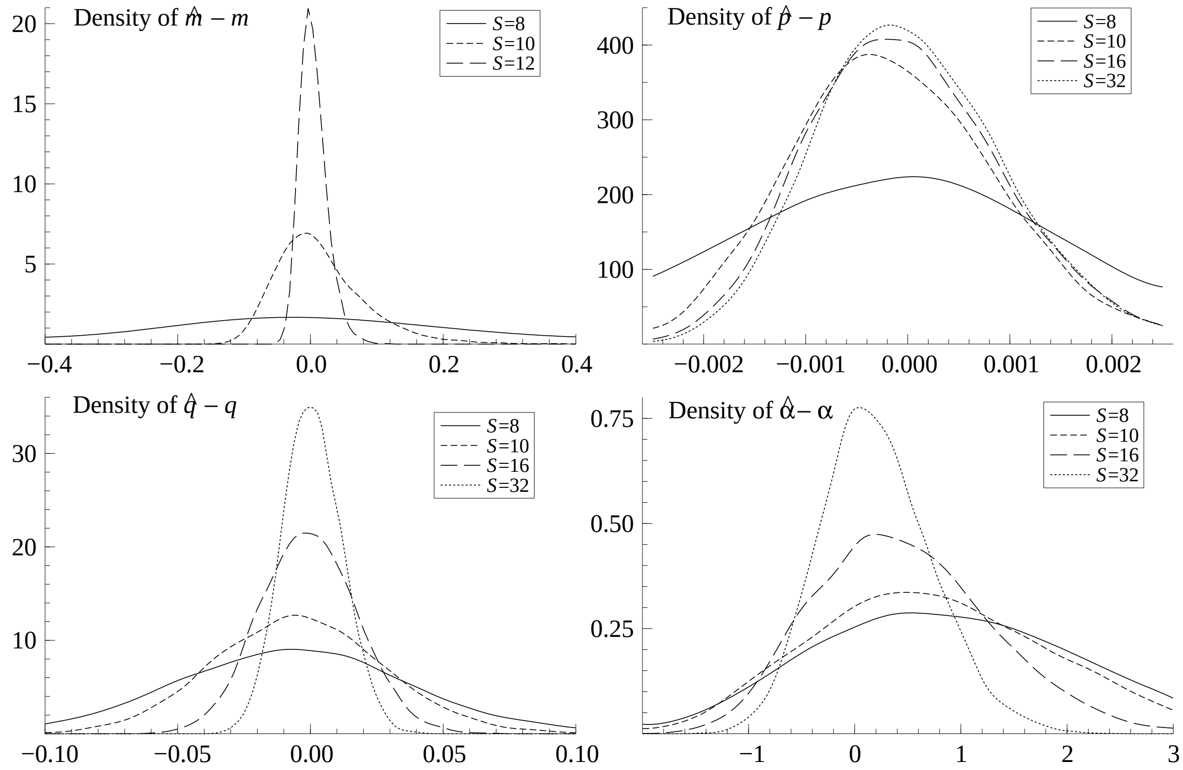


Figure 5: Asymptotic distributions of t-statistics of $\hat{m}$, $\hat{p}$, $\hat{q}$ and $\hat{\alpha}$, for various values of S. Key parameters in the DGP are $(p, q, \alpha, \sigma) = (0.01, 0.5, 5, 0.5)$.

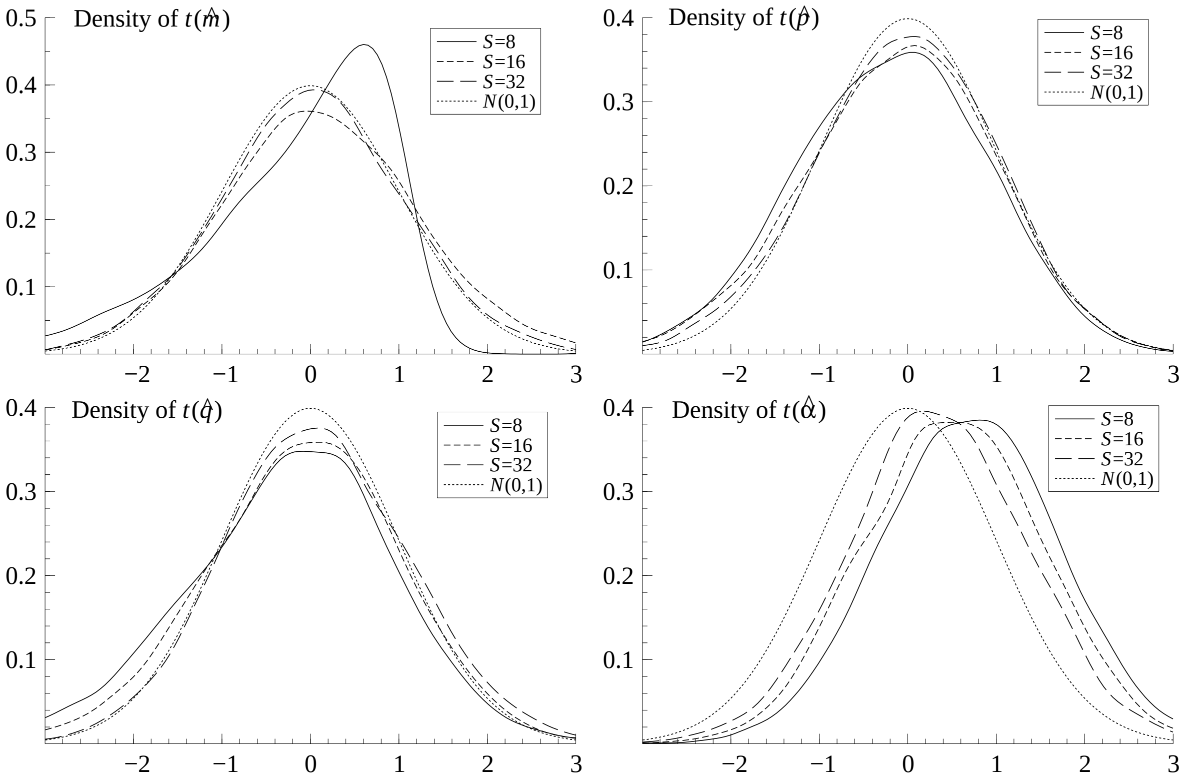


Figure 6: Finite-sample distributions of estimators (left) and t-statistics (right) of m, p, and q, with sample period and frequency given by $S = 12, \delta = 0.1$. Key parameters in the DGP are $(p, q, \alpha, \sigma) = (0.01, 0.5, 5, 0.5)$.

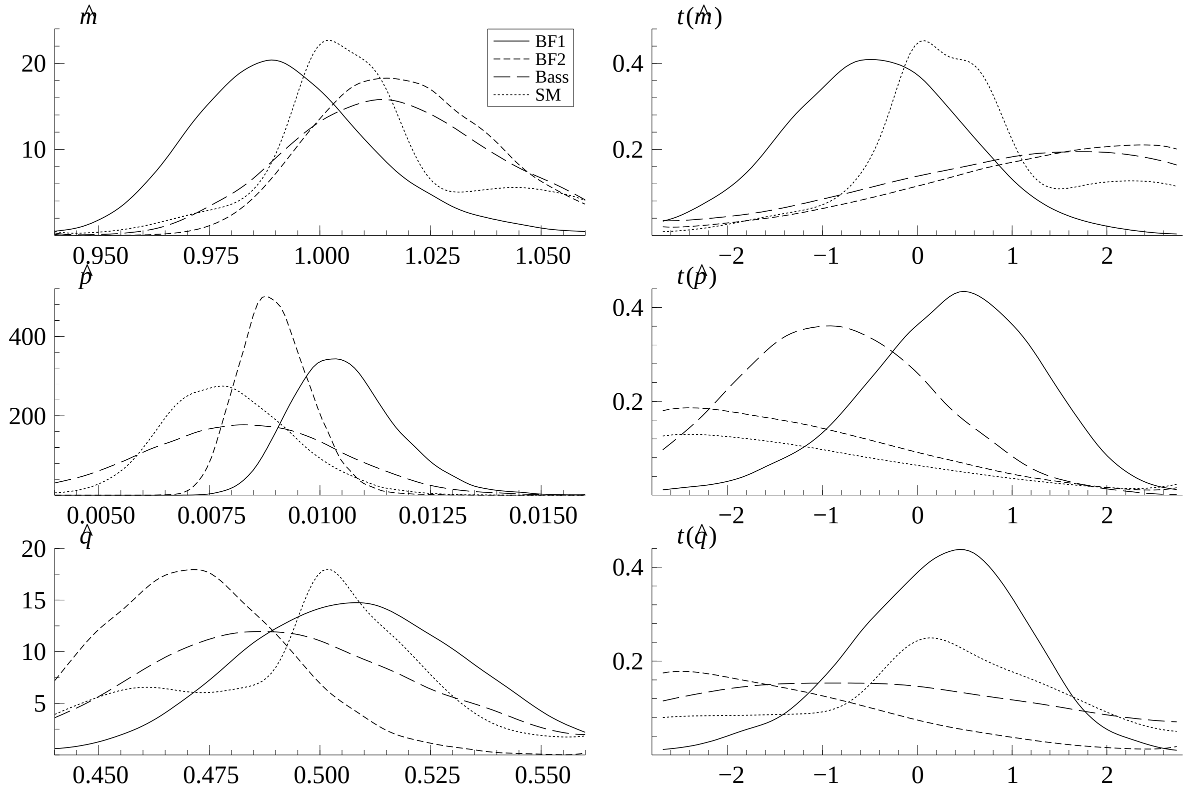


Figure 7: Finite-sample densities of estimators (left) and t-statistics (right) of m, p, and q, with sample period and frequency given by $S = 12, \delta = 1$. Key parameters in the DGP are $(p, q, \alpha, \sigma) = (0.01, 0.5, 5, 0.5)$.

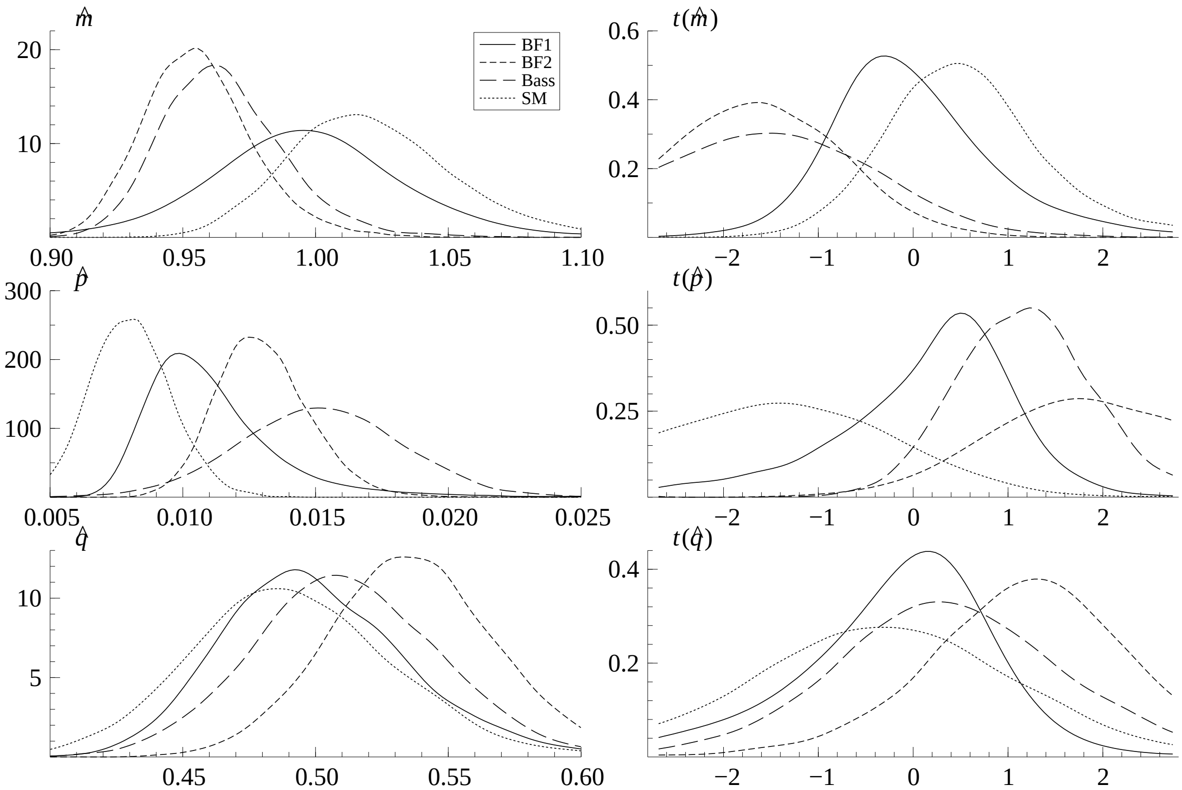


Figure 8: Finite-sample densities of estimators (left) and t-statistics (right) of m, p, and q, with sample period and frequency given by $S = 18, \delta = 0.1$. Key parameters in the DGP are $(p, q, \alpha, \sigma) = (0.01, 0.5, 5, 0.5)$.

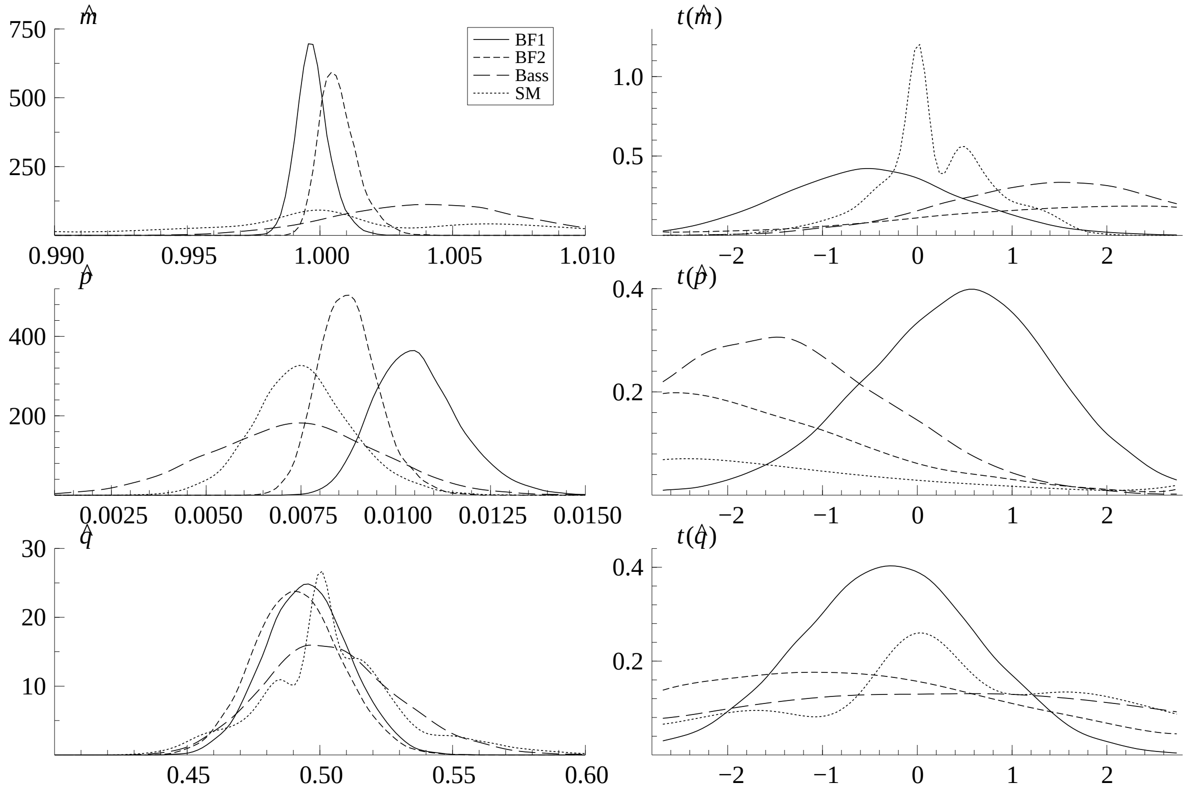


Figure 9: Finite-sample densities of estimators (left) and t-statistics (right) of m, p, and q, with sample period and frequency given by $S = 18, \delta = 1$. Key parameters in the DGP are $(p, q, \alpha, \sigma) = (0.01, 0.5, 5, 0.5)$.

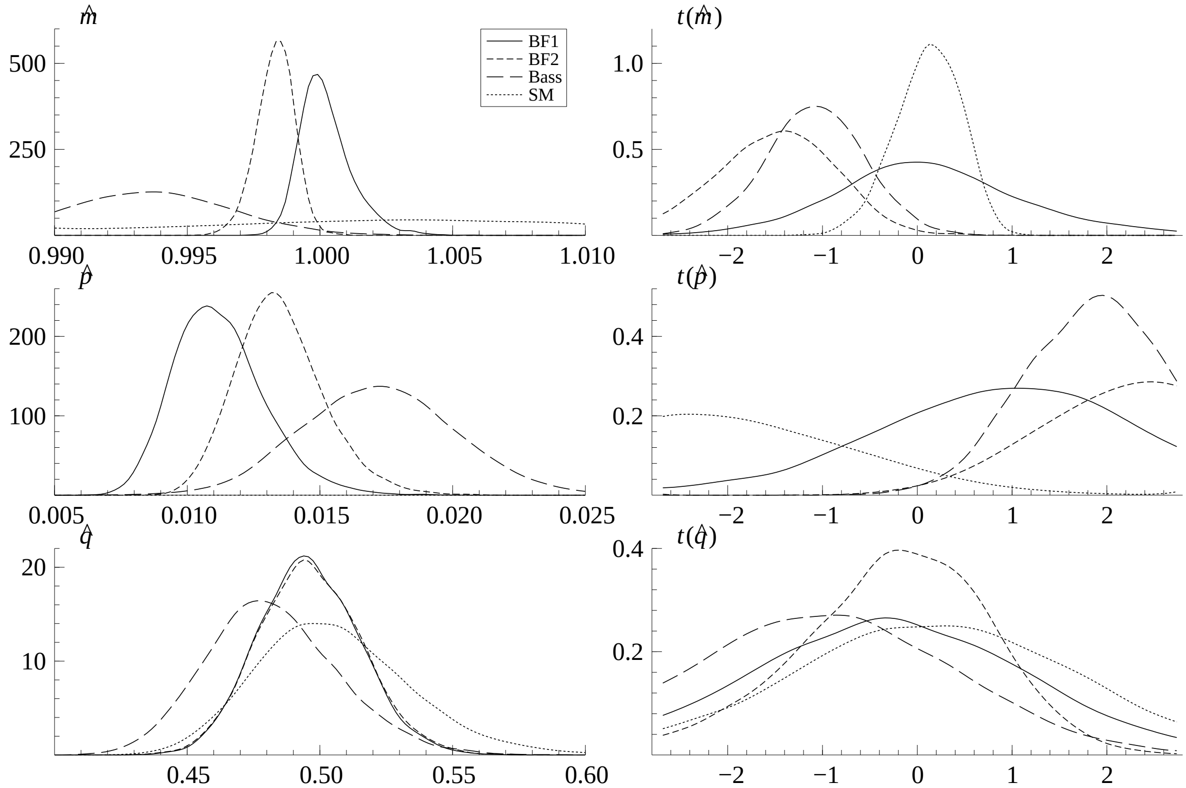


Figure 10: Number of newly installed ATMs, March 1990 - April 1996

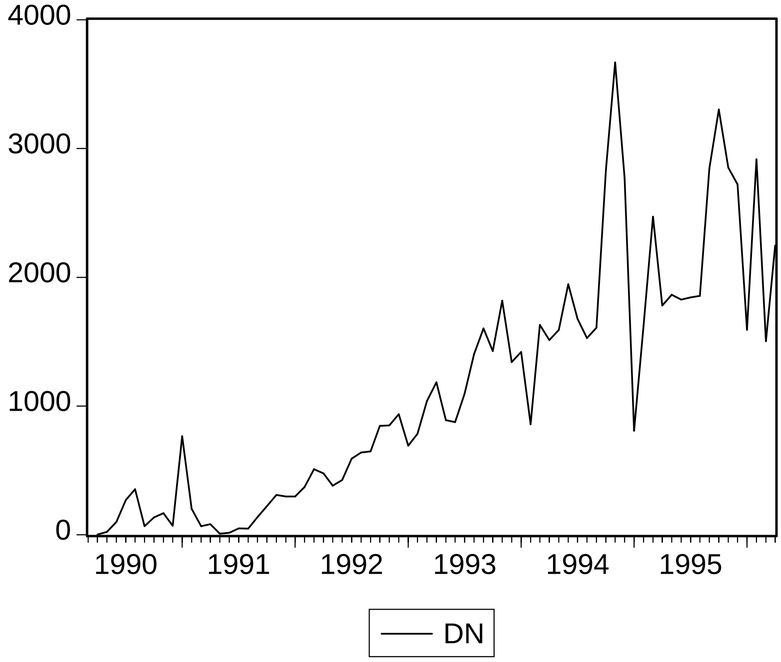
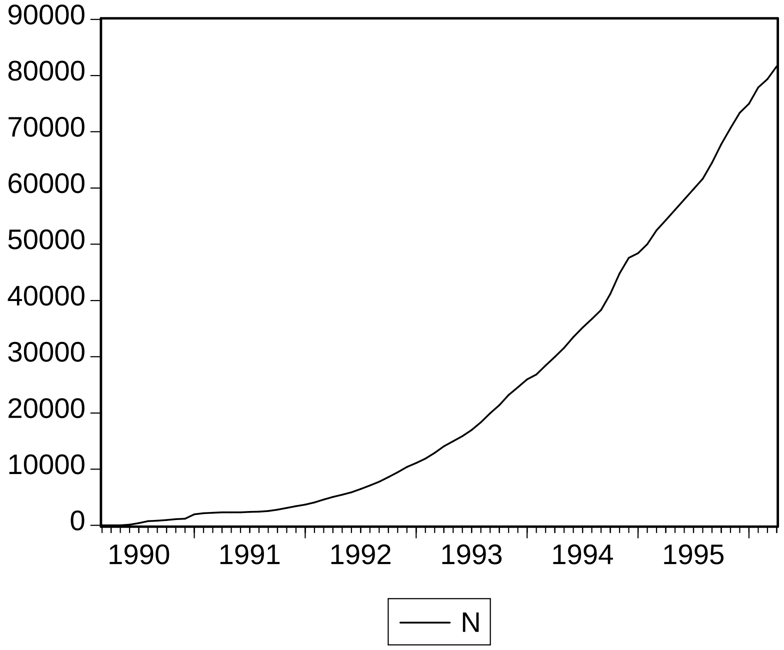


Figure 11: Cumulative number of installed ATMs



Publications in the Report Series Research* in Management

ERIM Research Program: "Marketing"

2002

Suboptimality of Sales Promotions and Improvement through Channel Coordination
Berend Wierenga & Han Soethoudt
ERS-2002-10-MKT

The Role of Schema Salience in Ad Processing and Evaluation
Joost Loef, Gerrit Antonides & W. Fred van Raaij
ERS-2002-15-MKT

The Shape of Utility Functions and Organizational Behavior
Joost M.E. Pennings & Ale Smidts
ERS-2002-18-MKT

Competitive Reactions and the Cross-Sales Effects of Advertising and Promotion
Jan-Benedict E.M. Steenkamp, Vincent R. Nijs, Dominique M. Hanssens & Marnik G. Dekimpe
ERS-2002-20-MKT

Do promotions benefit manufacturers, retailers or both?
Shuba Srinivasan, Koen Pauwels, Dominique M. Hanssens & Marnik G. Dekimpe
ERS-2002-21-MKT

How cannibalistic is the internet channel?
Barbara Deleersnyder, Inge Geyskens, Katrijn Gielens & Marnik G. Dekimpe
ERS-2002-22-MKT

Evaluating Direct Marketing Campaigns; Recent Findings and Future Research Topics
Jedid-Jah Jonker, Philip Hans Franses & Nanda Piersma
ERS-2002-26-MKT

The Joint Effect of Relationship Perceptions, Loyalty Program and Direct Mailings on Customer Share Development
Peter C. Verhoef
ERS-2002-27-MKT

Estimated parameters do not get the "wrong sign" due to collinearity across included variables
Philip Hans Franses & Christiaan Hey
ERS-2002-31-MKT

Dynamic Effects of Trust and Cognitive Social Structures on Information Transfer Relationships
David Dekker, David Krackhardt & Philip Hans Franses
ERS-2002-33-MKT

Means-end relations: hierarchies or networks? An inquiry into the (a)symmetry of means-end relations.
Johan van Rekom & Berend Wierenga
ERS-2002-36-MKT

---

* A complete overview of the ERIM Report Series Research in Management:
http://www.ers.erim.eur.nl

ERIM Research Programs:
LIS Business Processes, Logistics and Information Systems
ORG Organizing for Performance
MKT Marketing
F&A Finance and Accounting
STR Strategy and Entrepreneurship

Cognitive and Affective Consequences of Two Types of Incongruent Advertising
Joost Loef & Peeter W.J. Verlegh
ERS-2002-42-MKT

The Effects of Self-Reinforcing Mechanisms on Firm Performance
Erik den Hartigh, Fred Langerak & Harry R. Commandeur
ERS-2002-46-MKT

Modeling Generational Transitions from Aggregate Data
Philip Hans Franses & Stefan Stremersch
ERS-2002-49-MKT

Sales Models For Many Items Using Attribute Data
Erjen Nierop, Dennis Fok, Philip Hans Franses
ERS-2002-65-MKT

The Econometrics Of The Bass Diffusion Model
H. Peter Boswijk, Philip Hans Franses
ERS-2002-66-MKT

How the Impact of Integration of Marketing and R&D Differs Depending on a Firm's Resources and its Strategic Scope
Mark A.A.M. Leenders, Berend Wierenga
ERS-2002-68-MKT