Characterization of Demand for Short Life-Cycle Technology Products

Berrin Aytac
S. David Wu
Lehigh University

Report No. 08T-001

CHARACTERIZATION OF DEMAND FOR SHORT LIFE-CYCLE TECHNOLOGY PRODUCTS

Berrin Aytac and S. David Wu
Department of Industrial and Systems Engineering, Lehigh University, Bethlehem, PA 18015
bea4@lehigh.edu, david.wu@lehigh.edu

Abstract

Most technology companies are experiencing highly volatile markets with increasingly shortening product life cycles due to rapid technological innovation and market competition. Current supply-demand planning systems remain ineffective in capturing the short life-cycle nature of the products and the high volatility in the markets. In this study, we propose a demand characterization approach that combines life-cycle modeling with advanced demand signals from leading-indicator products through a Bayesian update. The proposed approach describes life-cycle demand in scenarios and provides a means to reduce the variability in demand scenarios via leading-indicator products. Computational testing on real-world data sets from three semiconductor manufacturing companies suggests that the proposed approach is effective in capturing the short life-cycle nature of the products and early demand signals and is capable of reducing the uncertainty in the demand forecasts by more than 20%.

1 Introduction

In the mid-to-late 1990s, high-tech companies, such as consumer electronics, telecommunications equipment, and semiconductors, grew rapidly. To reduce costs and cycle times, many firms developed and deployed supply-chain-management systems, but continued to rely on traditional demand planning in which the basis of a demand forecast is either an internal marketing judgement based on customer projections or traditional time-series forecasting methods that rely heavily on historical data and require stable demand trends. These approaches are ineffective in characterizing the volatile, non-cyclic high-tech product demands and in responding to the fast changing conditions of the market-place, such as the industry decline that succeeded in 2001.

In the challenging technology-driven market conditions, firms rapidly innovate technology and introduce new products to maintain their competitive position. This leads to shortening technology life cycles, which are known to follow a general demand life cycle that starts with an initial growth (ramp up) followed by a period of stability and then a decline in sales when a new generation of products is introduced. In this study, we develop a demand modeling approach that combines the insights from leading-indicator products and growth models in order to capture advanced market signals and demand life-cycles in an insightful manner.

The study has been motivated by our extensive analysis of demand data from a major U.S. semiconductor manufacturing company (Meixell et al. 2001, Wu et al., 2006). The analysis starts by grouping the products with similar attributes such as market segment, type of resource, and fab technology code. Within each product group, a search is performed to identify some potential leading indicator products. We define a "leading indicator product" as a single product or a group of products that are introduced earlier into the market and are likely to give information about the market a certain period ahead of time. We utilize the information from these products to extend the available data set of the products that they lead.

Besides leading indicators, the second key ingredient of our demand characterization approach is the use of life-cycle growth models. Combining the concepts of product life cycles and forecasting, growth models provide attractive alternatives to the traditional time-series forecasting methods for short life-cycle (high-technology) products (Kurawarwala and Matsuo, 1996). Technological forecasting literature, which studies the diffusion of innovations in a population, suggests several growth models that we can use to describe a wide variety of life-cycle patterns observed in practice. Mahajan, Muller and Bass (1990) provide an extensive review of these models and study their applications to marketing.

Availability of several models, however, makes it difficult to select the model that best describes the historical data of a given set of products and that generates accurate forecasts. As opposed to the technological forecasting literature, which mostly focuses on the performances of individual models, we use a large collection of models to characterize and predict the demand growth patterns of the products. We use the additional data from several leading indicators to extend the available data set of the products and to Bayesian update the estimates of the models obtained from the available historical data. The update provides a means to incorporate the future demand information, and reduces the uncertainty in the characterization of demand. In addition, one problem associated with the short life-cycle product environment is the scarcity of available data. A byproduct of this approach is that the increase in the number of data points can result in parameter estimates with higher significance.

To investigate the impact of the additional data on the projections of the models, we first study the relative change in the projections of the growth models. As a measure, we use the range of the projections. Second, we consider the combined estimate of the models and describe demand in terms of the scenarios that are sampled from the probability distribution of the combined estimate, rather than from the probability distribution of the estimate of the single best model. We show that the update streamlines the characterization of the future demand patterns of the products and generates demand scenarios with smaller variation.

In the next section, we review the relevant literature. We introduce our demand characterization approach in Section 3 and the analysis of the variability in demand characterization in Section 4. The implementation of the approach on real-life data sets is presented in Section 5, and concluding remarks follow in Section 6.

2 Literature Review

The emphasis of the technological forecasting literature has been on obtaining point forecasts via trend extrapolation and improving the accuracy of these forecasts. The focus of the studies in the literature in terms of improving forecast accuracy has been to select the right model(s) and to determine the best parameters. Our study further seeks to improve the performances of the best models, by incorporating future demand signals. Furthermore, we aim to obtain demand forecasts with smaller variation, so that their estimates can be given more confidence.

Algebraic estimation and fitting procedures are the two basic approaches in the literature for the estimation of the best parameters and generation of demand forecasts. Mahajan and Sharma (1986) propose an algebraic estimation procedure that uses the estimates of some key information about the life-cycle of a product, such as the timing of peak sales and the market potential. Fitting procedures determine the parameters that best fit the available data in each period. Various fitting procedures have been suggested, including the ordinary least squares estimation (Bass, 1969), the maximum likelihood estimation (Schmittlein and Mahajan, 1982), and the nonlinear least squares estimation (Srinivasan and Mason, 1986). The ordinary least squares procedure requires that the model be linear in its parameters, and therefore uses a linearized version of the original growth model; however, linearization results in poor parameter estimates (Mahajan et al., 1990). Both the maximum likelihood and nonlinear least squares estimation procedures provide superior results to the ordinary least squares estimation procedure; however, a disadvantage of the maximum likelihood estimation procedure is that it underestimates the standard errors of the parameters (Srinivasan and Mason, 1986). Therefore, among these approaches, the most widely used is the nonlinear least squares procedure.

Extensions of the algebraic estimation and fitting methods are the use of feedback filters and Bayesian estimation procedures that allow the model parameters to be updated as additional data become available. Feedback filters adjust the estimates of the parameters from the fitting methods based on the error between the actual and the predicted values of the most recently observed demand data (e.g., Meade, 1985). Bayesian approaches combine the data-based estimates from the fitting methods with the estimates from the algebraic estimation procedure or with the information from other sources such as from similar products that were launched earlier into the market. Among the studies that use Bayesian updating to forecast demand over the product life cycles, one well-known study is Sultan et al. (1990), in which a meta-analysis was performed using 213 applications of diffusion models from 15 articles. The impact of factors such as the type of innovation, geographic effect, and the marketing variables on the value of parameters is determined. The estimates from the meta-analysis are used before a product is launched. As new data become available, the prior estimates from the meta-analysis are Bayesian updated with the data-based estimates. It is shown that the Bayesian scheme produces more robust results, particularly early in the product history. Zhu et al (2004) is another study that proposes a Bayesian forecasting algorithm for products with short life cycles. The algorithm combines the knowledge of the prior products with the actual demand data of the product to generate and update demand forecasts. The empirical analysis on a data set from a PC manufacturer shows that the Bayesian algorithm produces better forecasting performance than double exponential smoothing (assumes linear demand trend), algebraic estimation (no update with actual data), and fitting methods. All these studies use some past information for the update and do not take into account the future changes in the conditions of the market-place and the product.

A study that is similar to ours in incorporating future information into demand analysis is by Islam et al. (2002). They consider the diffusion of an innovation in several countries, which have different starting times but similar dynamics of the diffusion process. Data from several countries are pooled using linearized growth curve models. However, the empirical analysis by Takada et al. (1991) on the diffusion rates of several consumer durable goods in four Pacific Rim countries shows that cultural and communications systems in a country and whether a product is introduced earlier or later do have an impact on the adoption rates of the products; therefore, the use of data from other countries is not necessarily valid. Meixell and Wu (2001) is another study that considers future market information. As in our study, leading-indicator products are used to obtain advanced demand information, but in a different manner. The available data set of the leading indicator product defines an expected nominal growth curve. The parametric deviations from the nominal curve such as shift in volume and skewness form alternative demand scenarios. The probability of each parameter state changes as more data from the leading indicator product becomes available.

Another stream of related research is the quantification of uncertainty in growth-curve forecasting, which is not a well-studied subject. Chatfield (1993) states the difficulty in modeling the uncertainty in the estimates with linear models and describes several underlying reasons such as incorrect model identification, uncertainty in parameter estimates, and changes in the data-generating process. The intrinsic nonlinearity of the growth models contributes to this difficulty and makes it a less appealing area to study. One of the noteworthy studies is by Meade and Islam (1995). They suggest three approaches to construct prediction intervals for growth-curve forecasts and compare their performances on both simulated data sets and real data sets. The empirical results indicate that as the number of data points available for parameter estimation increases, the performances of the prediction intervals improve. However, no attempt is made to reduce the variance of the forecasts and hence improve the width of the prediction intervals. Heeler and Hustad (1981), Tigert et al. (1981), Srinivasan et al. (1986), and Islam et al. (2002) are among the studies that empirically show that as the number of actual data points used for estimation increases, estimates of the parameters stabilize (i.e., their estimation variance becomes close to zero) and accuracy of the forecasts increases. Other than this line of approach that studies the impact of the increase in the number of historical data points, to the best of our knowledge, there is no study in the technological forecasting literature that is concerned with reducing the uncertainty in growth-curve forecasting.

An attempt in the general forecasting literature to reduce forecast variance is to combine several forecasts. There are several empirical studies that show that with combination, forecasts with smaller variances are obtained (Dickinson 1973). In the light of this information, we consider the combined estimate of several growth models, and characterize the probability distribution of this estimate. We then use the continuous probability distribution by discretizing and generating random samples as demand scenarios.

The resulting demand scenarios can be used in the decision-making activities of the firms. For example, stochastic programming models use distributional forecasts rather than point forecasts as in deterministic optimization models. Since it is difficult to solve a stochastic programming model exactly with the characterization of the random variables as continuous distribution, the model can be approximated using random samples generated from the continuous probability distribution. A related study is by Morton and Popova (2004). They use Bayesian forecasting to model uncertain machine up-times and production rates in a production planning problem. Data from the past months is used to update the distributional forecasts of these parameters, and Monte Carlo random sampling is used to generate observations of the random parameters for a stochastic decision problem.

We propose a demand characterization approach that aims to improve the accuracy and reduce the variance of the forecasts. The additional data from leading indicators that carry information about the future demand is used to achieve this. If used together with the variance reduction techniques in sampling, the proposed approach has the potential to increase the statistical efficiency and time efficiency of the decision models that will use the resulting demand scenarios.

3 Life-Cycle Demand Characterization: A Methodological Approach

The projections of the sales of the products with short life cycles typically go through one phase of rapid growth, maturity, and decline and are subject to high uncertainty. The central concept of our approach is to characterize the life-cycle projections of these products using a large number of single-modal growth models and to update the projections with the advanced demand information from leading-indicator products. We then study the change in the accuracy and the variability of the projections with the update.

In this section, we first introduce the concept leading indicator products and propose a search procedure to identify potential leading indicator products for a given group of products. We next describe the modeling procedure with life-cycle growth curves. We then discuss how to combine the life-cycle projections obtained from the available historical data of the products with the information from leading indicator products through Bayesian approach. The output of the Bayesian approach is the characterization of the estimate of each growth model with a probability distribution that includes advanced demand information.

3.1 Leading-Indicator Analysis

We define a leading-indicator product as a product that demonstrates a demand pattern that is similar to the demand pattern of a larger product group a certain period ahead of time. With the use of a leading-indicator product, we aim to obtain advanced indications about the demand pattern of the product group. A similar concept is used in the fashion goods industry, where long production lead times and short selling seasons are common as in the high-tech industry. For example, Fisher and Raman (1996), Eppen and Iyer (1997), and Kim (2003) use the sales records in some representative fashion stores in the pre-season test periods and the actual sales data in the early season as demand indicators for the actual demand in the selling season. These studies also apply the Bayesian approach to forecast demand, though without any consideration of life-cycle modeling.

There are several ways to form the product groups, for which we perform the analysis. In Meixell and Wu (2001), after the study of demand data for about 3,500 products, it was found that these products follow a few (approximately six) life-cycle patterns and can be grouped into these patterns using statistical cluster analysis. In Wu et al. (2006), we focus on exogenously-defined product groups, in which products that share the same resource, technology group, or market segment are grouped together. This type of grouping is helpful in modeling production planning problems, where there is typically a set of capacity-related constraints that link usage of a resource to its available capacity.

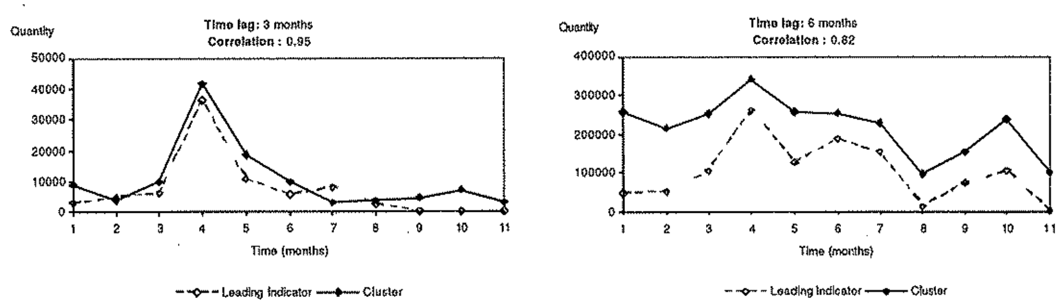


Figure 1: Examples of Leading Indicator Products

Within each product group, we search for potential leading-indicator products. Given a group of products C, we treat each product i in the group as a potential leading-indicator product and check the time lag L ahead of which the product exhibits a similar demand pattern to the rest of the product group. Let $x_{i,t}$ and x_t denote the demand data of the potential leading-indicator product i and the rest of the product group in period t, respectively. At any time t, $\{x_{i,1}, ..., x_{i,t}\}$ and $\{x_1, ..., x_t\}$ are known while $\{x_{i,t+1}, x_{i,t+2}, ...\}$ and $\{x_{t+1}, x_{t+2}, ...\}$ are unknown. Given the data available up to time T, we determine a potential leading indicator product based on the absolute value of the correlation coefficient between the demand time series of the product over $[1, T-L]$ and that of the product group shifted by the time lag of the leading indicator, that is, over $[L+1, T]$, using the formula:

$$\rho_{iL} = \frac{\sum_{t=L+1}^{T}(x_{i,t-L} - \bar{x}_i)(x_t - \bar{x})}{\sqrt{\sum_{t=L+1}^{T}(x_{i,t-L} - \bar{x}_i)^2 \sum_{t=L+1}^{T}(x_t - \bar{x})^2}} \tag{1}$$

where $\bar{x}_i$ and $\bar{x}$ are the average demand of the potential leading-indicator product i and the rest of the group over the time interval for which the value of correlation is calculated. Table 1 gives the details of the procedure to find candidate leading-indicator products.

Table 1: The Leading-Indicator Search Procedure

---

(1) Identify a product group of interest and set the threshold minimum time lag ($\underline{L}$), maximum time lag ($\bar{L}$), and minimum correlation. Initialize the procedure by grouping all the products into one group.

(2) To find all the leading indicators above the required threshold, for each product i in a given group C,

- (a) Set time lag $L = \underline{L}$.
- (b) Compute the correlation between the demand time series associated with product i where the time series is offset by L and the demand time series associated with the product group C as in Equation 1.
- (c) Record the correlation ρ_{iL} computed for product i and time lag L.
- (d) Set $L = L + 1$. If $L \leq \bar{L}$, repeat Step (b).

(3) Examine all the recorded correlation numbers ρ_{iL}. If at least one of the correlation values ρ_{iL} and its corresponding time lag L satisfy the specified threshold, go to Step 4. Otherwise, recluster as follows:

- (a) Using statistical cluster analysis, subdivide the product group based on statistical patterns in each product's historical demand; some attributes that can be used for clustering are mean shipment quantity, shipment frequency, volatility, or skewness.
- (b) Repeat Steps 2 and 3 for each subgroup.

(4) Return the leading indicator(s) and the corresponding product group(s).

For a given group of products, if a leading-indicator product that satisfies the minimum requirements on the time-lag and correlation value cannot be found, the search procedure is performed within subgroups of the product group that are obtained using statistical cluster analysis. Once we determine a leading-indicator product with a time lag L and a value of coefficient of correlation ρ_{iL}, we construct a forecast for the product group based on the time series of the leading indicator. By regressing the time series of the leading indicator over $[1, T-L]$ against the time series of the product group over $[L+1, T]$, we determine the formula that transforms the leading-indicator-based data into the data of the product group. Using this formula, we obtain the leading-indicator-based forecasts over $[T+1, T+L]$, which is used to extend the available data of the product group beyond the available data $[1, T]$. The details of the procedures follow in Table 2.

Table 2: The Leading-Indicator-Based Forecast

(1) Regress the time series of product group C over $[1+L, T]$ against the time series of the leading indicator over $[1, T-L]$. Determine the corresponding regression parameters $\hat{\beta}_0$ and $\hat{\beta}_1$.

(2) For a given month t, generate the forecast for the product group, $\tilde{x}_t$, using L-month earlier time-series data of leading indicator i as follows:

$$\tilde{x}_t = \hat{\beta}_0 + \hat{\beta}_1 x_{i,t-L}$$

Figure 1 illustrates two examples for the monthly time-series data of a product group (solid line) against that of a leading-indicator product (dashed line). The time series of the group has been shifted ahead by the corresponding time lag to show the mapping between the two series. The first chart shows a leading indicator that predicts the demand pattern of a larger group three months ahead of time with a correlation of 0.95; the second chart shows a leading indicator with a six-month time lag and correlation of 0.82. For more examples, see Wu et al. (2006).

3.2 Life-Cycle Growth Modeling

The projections of the sales of the products with short life cycles typically go through one phase of rapid growth, maturity, and decline, and follow a single-modal curve. The use of non-cumulative growth models requires the precise prediction of the patterns of a given set of time-series data. In order to lessen the impact of the short-term fluctuations on the prediction quality of the growth models, we use cumulative life-cycle models. Hence, an S-shaped curve characterizes the projection of the life-cycle demand that occurs over time (Figure 2).

With this new characterization of demand, forecasting the demand in each time period

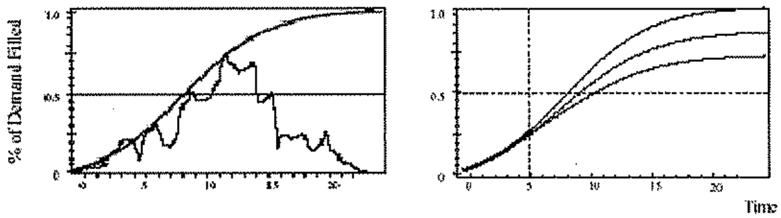


Figure 2: Cumulative Demand and Projections of Different Curves

requires the projection of the S-shaped cumulative life-cycle demand patterns of the products into the future. We use several cumulative life-cycle growth models, each of which exhibits different features of a projection (Figure 2). For the selection of these models, we refer to the study by Meade and Islam (1998). They survey the 29 models in the technological forecasting literature and group these models into different classes. The classification is based on either the way the models incorporate randomness or their shape. The shape of a model is characterized by the timing of its point of inflection, which corresponds to the point on the growth curve with the maximum rate of sales.

Given the projection of the life-cycle demand of the product group up to any time T, we forecast the percentage of the projected demand that will be met by time $T+M$ $(M > 0)$. Let $X(t)$ be the percentage of the life-cycle demand that has been met by time t $(0 \leq X(t) \leq 1, t > 0)$, and let $\widehat{X}_k(t|\Theta_T)$ represent the estimate of the cumulative growth model k $(k = 1, ..., K)$, given data up to time T, i.e., $\Theta_T = \{X(1), ..., X(T)\}$, such that:

$$X(t) = \widehat{X}_k(t|\Theta_T) + \epsilon_k(t|\Theta_T), \tag{2}$$

where $\epsilon_k(t|\Theta_T)$ represents the estimation error. We assume that the estimate of a cumulative life-cycle growth model for time $T+M$ generates an unbiased forecast for the actual percentage of the total life-cycle demand that will be met by time $T+M$. We further assume that the estimation error is normally distributed with mean zero and constant known variance $\sigma^2_{\epsilon_k}$. In the next section, we discuss the Bayesian demand analysis that combines the information from leading-indicator products and life-cycle growth modeling from the historical data.

3.3 Bayesian Demand Analysis

Given the importance of associating a measure of uncertainty with a point forecast, we express the uncertainty in the estimates of growth models using Bayesian approach, which treats the estimate of each growth model as a random variable and describes each random variable with a probability distribution. This allows a complete characterization of the uncertainty associated with the estimates.

The Bayesian approach starts with a prior distribution that summarizes the data available. With the observation of new data, the prior distribution is updated, resulting in the posterior distribution. In light of this explanation, we obtain the prior projection of a growth model from the historical data, which is available up to time T, and derive the prior distribution of the estimate of the model (Figure 3a). We then use several leading indicators with a time lag of at least L periods to forecast the actual data over $[T+1, T+L]$, and use the leading-indicator-based forecast to extend the historical data. The estimate of a growth model over the data set extended with each leading indicator provides a sample life-cycle projection for the true projection of the model that would be obtained with the actual data over $[1, T+L]$ (Figure 3b). Bayesian updating the prior distribution with the sampling information from leading indicators gives the distributions that characterize the posterior life-cycle projections as the new estimates (Figure 3c).

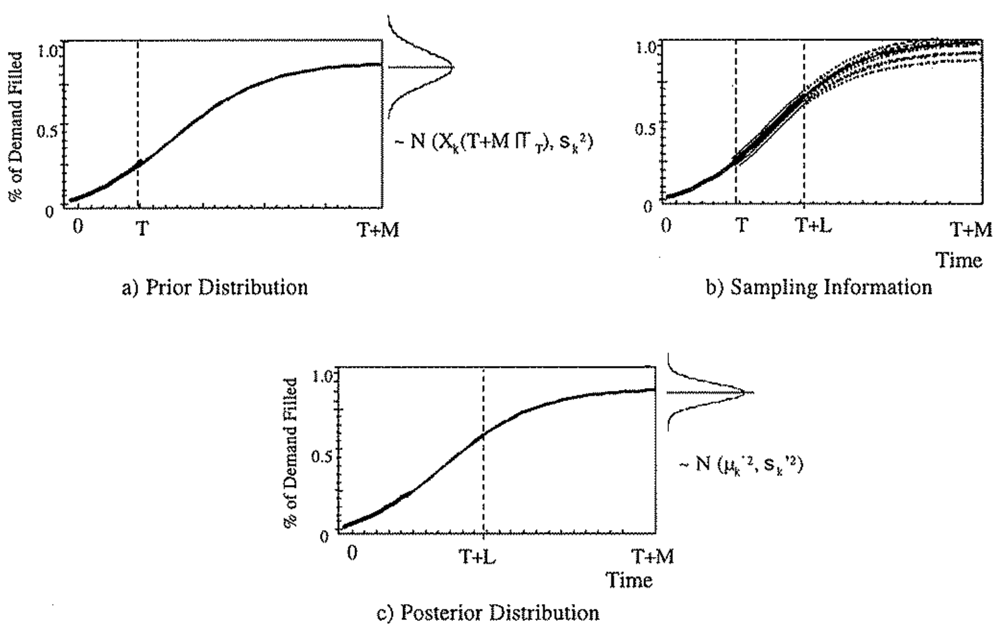


Figure 3: Probability Distribution of the Estimate of a Growth Curve

3.3.1 Prior Life-Cycle Projections

We assume that the estimate of a model with the data available up to time T, i.e., $\widehat{X}_k(T + M|\Theta_T)$ $(k = 1, ..., K)$, is an unbiased estimator of the actual data at time $T + M$, i.e., $X(T + M)$. We use $\widehat{X}_k(T + M|\Theta_T)$ to obtain prior information for $X(T + M)$ from Equation 2. From the assumed normality of the estimation error, it follows that the estimate of the actual data is normally distributed around the prior mean with variance σ_k^2:

$$\widetilde{X}_k(T + M) \sim N\left(\widehat{X}_k(T + M|\Theta_T), \sigma_k^2\right)$$

where $\widetilde{X}_k(T + M)$ denotes the random variable that corresponds to the actual data at time $T + M$ for model k.

We quantify the uncertainty in $\widetilde{X}_k(T + M)$ by its variance, σ_k^2. Under the assumption that the estimate of a growth model and the error between the actual data and the estimate are independent (Equation 2), σ_k^2 can be expressed as the sum of the variance of the estimate of the growth model, ω_k^2, and the variance of the estimation error, $\sigma_{\epsilon_k}^2$:

$$\begin{aligned} \widehat{var}\left(\widetilde{X}_k(T + M)|\Theta_T\right) &= \widehat{var}(\hat{X}_k(T + M|\Theta_T)) + \widehat{var}(\epsilon_k(T + M|\Theta_T)) \\ \sigma_k^2 &= \omega_k^2 + \sigma_{\epsilon_k}^2. \end{aligned}$$

The main sources of the uncertainty in the estimate of a growth model are known to be the error in the estimation of its parameters and the inadequacy of the selected model. To estimate the variance of this quantity, we refer to Meade and Islam (1995), in which three approaches to construct prediction intervals for the estimate of growth models are proposed. The first approach, the approximated variance approach, uses the variance of the linearized Taylor series expansion of the model at the final estimates of the parameters and the (asymptotic) covariance matrix of the parameters. The second approach, the explicit density approach, derives the estimate of the variance by explicitly modeling the uncertainties in the parameter estimates. The third approach, bootstrapping, is a non-parametric approach that develops the distribution of the estimate of the growth curve by resampling the residuals. Meade and Islam's analysis with simulated data sets shows that the explicit density approach performs best if the number of data points available is small (less than 20 observations). As a measure of performance, differences between the percentage of predictions that fall within the intervals and the significance level that is used to construct the interval are considered.

3.3.2 Samples of Life-Cycle Projections

At time period T, we use the L-period leading-indicator data over $[T - L + 1, T]$ to obtain the estimate of the actual data of the product group over $[T + 1, T + L]$, which is unknown at time T. As illustrated in Figure 3b, this advances the position on the projected curve from time T to time $T + L$. We consider several leading-indicator products. With the data from each leading-indicator product, we obtain the projection of the models, each of which gives a sample life-cycle projection for the true projection of the model that would be obtained if actual data was used over $[1, T + L]$.

We assume that the data from leading indicators provide unbiased samples for the actual data over $[T + 1, T + L]$; therefore, the estimate of a model with a leading indicator is assumed to provide an unbiased sample for the true estimate of the model that would be obtained with actual data over $[1, T + L]$. We use a sample of m leading indicators to derive the sampling information about the actual data at time $T + M$.

The estimate of model k with leading indicator j ($j = 1, ..., m$), is obtained over the estimation period extended with the leading-indicator-j-based data, i.e., Θ^j_{T+L}, the first $T-$periods from the actual data, and the next $L-$periods from the leading-indicator-j-based data. We can express our assumption that the estimate of a model with each leading indicator provides an unbiased sample for the true estimate of the model with actual data, i.e. over Θ_{T+L}, as:

$$\widehat{X}_{kj}(T + M|\Theta^j_{T+L}) = \widehat{X}_k(T + M|\Theta_{T+L}) + \varepsilon_{kj} \tag{3}$$

where ε_{kj} represents sampling noise. We assume that ε_{kj} is normally distributed with mean zero and constant known variance φ^2_k. Integrating our assumption that the estimates of the models are unbiased estimators of the actual data gives that the estimates of the models with the leading indicators are unbiased estimators of the actual data. As a result, under the assumption that the leading indicators are independent, it is obtained that $\{\widehat{X}_{kj}(T + M|\Theta^j_{T+L}), j = 1, ..., m\}$ are independently and identically distributed observations from a normal distribution with mean $X(T + M)$ and variance $\tau^2_k = \sigma'^2_{\epsilon_k} + \varphi^2_k$, where $\sigma'^2_{\epsilon_k}$ represents the variance of the estimation error at time $T + L$. We can summarize the sampling information as:

$$\widetilde{X}_k(T + M) \sim N\left(\frac{1}{m}\sum_{j=1}^{m}\widehat{X}_{kj}(T + M|\Theta^j_{T+L}), \frac{\tau^2_k}{m}\right)$$

The confidence in the sample mean increases with the number of leading indicators used. By the strong law of large numbers, with probability one the sample mean provides an unbiased estimate for the actual data at time $T+M$ when m goes to infinity.

Biased Leading-Indicator-Based Data

There is a possibility that the data from a leading indicator is not an unbiased estimator of the actual data over $[T+1, T+L]$. Thus, the estimate of a model using this leading indicator will not be an unbiased estimator of the true estimate of the model with actual data over $[1, T+L]$, either. We model the impact of the bias as follows: We describe the bias of a model with a leading indicator as a random variable and characterize it as a normally distributed perturbation.

Proposition 1. *Under the assumption that the bias of the estimate of a model with a leading indicator, β_{kj}, is a normally distributed random variable with mean b_{kj} and variance θ_{kj}^2, the variance of the sampling estimate increases with the bias; the mean of the estimate may shift depending on the relative bias of the leading indicators.*

Proof. By definition, the bias of the estimate of a growth model with a leading indicator implies that

$$\begin{aligned} E[\widehat{X}_{kj}(T+M|\Theta_{T+L}^{j})] &= E[\widehat{X}_k(T+M|\Theta_{T+L}) + \beta_{kj}] \\ &= X(T+M) + b_{kj}. \end{aligned}$$

The mean and the variance of the sampling distribution change to:

$$E\left[\frac{1}{m}\sum_{j=1}^{m}\widehat{X}_{kj}(T+M|\Theta_{T+L}^{j})\right] = X(T+M) + \frac{1}{m}\sum_{j=1}^{m} b_{kj}$$

$$Var\left[\frac{1}{m}\sum_{j=1}^{m}\widehat{X}_{kj}(T+M|\Theta_{T+L}^{j})\right] = \frac{\tau_k^2}{m} + \frac{\sum_{j=1}^{m}\theta_{kj}^2}{m^2}.$$

The results indicate that bias increases the variance of the sampling estimate; the mean of the estimate may shift depending on the relative bias of the leading indicators. $\square$

3.3.3 Posterior Life-Cycle Projections

The Bayesian update of the prior distribution with the sampling information gives the posterior distribution of the actual data at time $T+M$:

Proposition 2. *If the m leading-indicator-based sampling estimates follow a normal distribution with mean $X(T+M)$ and variance τ_k^2, and $X(T+M)$ has prior normal distribution with mean*

$\widehat{X}_k(T+M|\Theta_T)$ *and variance* σ_k^2*, the posterior estimate is normally distributed:*

$$\widetilde{X}_k(T+M) \sim N(\mu_k', \sigma_k'^2)$$

where

$$\begin{aligned}
\mu_k' &= \frac{1/\sigma_k^2}{1/\sigma_k^2 + m/\tau_k^2}\,\widehat{X}_k(T+M|\Theta_T) + \frac{m/\tau_k^2}{1/\sigma_k^2 + m/\tau_k^2}\,\frac{1}{m}\sum_{j=1}^{m}\widehat{X}_{kj}(T+M|\Theta_{T+L}^j) \\
\sigma_k'^2 &= \frac{\sigma_k^2\tau_k^2}{m\sigma_k^2 + \tau_k^2}
\end{aligned}$$

Proof. The proof is provided in the Appendix.

The updated life-cycle demand projections are the means of the posterior distributions, which are the weighted averages of the prior projections and the sampling means. The weights are inversely proportional to the variances of the prior distribution and the sampling distribution, respectively. The use of additional data from leading indicators leads to the following conclusions:

Corrolary 1. *Given the variances of the projections, the following are true:*

(i) Bayesian update of the projections of the models with the additional data from leading indicators results in estimates with smaller variance, i.e., $\sigma_k'^2 \le \sigma_k^2$ $(k = 1, ..., K)$.

(ii) If all the leading indicators are unbiased, the variance of the estimates asymptotically approaches zero as the number of leading indicators increases.

(iii) Variance of the projections is increasing with the bias of leading indicators, and the following ordering holds: $\sigma_k'^2 \le \sigma_k''^2 \le \sigma_k^2$*, where* $\sigma_k''^2$ *is the variance with some leading indicators biased.*

The assumption of normal distribution is not restrictive. We can express our reason for making this assumption as follows: Normal distribution provides a reasonable characterization of the sampling and estimation errors. In our case, this results in a normal sampling distribution. The corresponding natural conjugate prior distribution, which is selected for mathematical convenience, is also normal distribution. The use of natural conjugate prior distributions has the advantage that the induced posterior distribution is in the same family as the prior distribution, and therefore allows for closed-form expressions. The analysis is still valid if this assumption is released. In particular, when the variance is unknown, sample variance can be used, in which case t-distribution is a more accurate representation of the uncertainty in the sampling estimates.

We defined the change in the projections of the individual growth models with the addition of data from leading indicators. In the next section, we investigate the impact of the additional data on the variability in the characterization of demand with several growth models.

4 Variability in Life-Cycle Demand Characterization

Given the uncertainty associated with any forecasting method, this study emphasizes the importance of the degree of the variability of a forecast as well as its accuracy. In this section, we study the variability in the life-cycle demand characterization and the impact of the additional data on the variability. In particular, we question whether the additional data results in estimates with smaller variability. We consider two measures. The first one is the change in the relative estimates of the projections of the models, and the second one is the variance of the combined estimate of the projections.

4.1 Range of the Life-Cycle Demand Projections

We use the distributional characterization of the estimates of the models to investigate the change in the variation of the projections with the additional data. As a measure of variation, we consider the *range* of the estimates at time $T+M$, which is the difference between the estimates with the maximum and the minimum values at time $T+M$, and make a probabilistic conclusion regarding the change in the variation of the projections.

The methodology used here follows a similar logic to *ordinal comparison.* Given a set of alternative choices, ordinal comparison finds the best choice based on their relative performances, rather than based on the exact estimates of their performances. Each decision is assigned a probability measuring that the alternative choices are ranked correctly and at least one good choice is identified. Among the several studies, Chen (1996) and Chen et al. (2000) are concerned with the selection of the best design in a manufacturing and communication system, Chen et al. (1999) select among heuristic algorithms for combinatorial problems, and Bonser et al. (2001) consider several supply contracts for fuel procurement in electrical utilities using ordinal comparison. These studies approximate the *exact performances* of the alternatives using a finite number of *sample problem instances.* Similarly, we approximate the *true estimates* of the growth models with actual data using data from a finite number of *leading indicators* that sample the actual data.

Let $\triangle^0_{T+M}$ denote the range at time $T+M$ obtained using the actual data over $[1,T]$, that is, prior to using any additional data, and let $\triangle_{T+M}$ denote the random variable that characterizes the range at time $T+M$ obtained using the posterior projections of the models. We determine the exact probability that the variation in the projections (in terms of range) decreases with the additional data, that is,

$$Pr(VR) = Pr\{\triangle_{T+M} < \triangle^0_{T+M}\}$$

where $\triangle^0_{T+M}$ is a constant, and $\triangle_{T+M}$ is a random variable:

Proposition 3. *Let $f_k(\cdot)$ be the probability density function for the random variable $\widetilde{X}_k(T+M)$ that denote the estimate of model k, $k=1,...,K$, for the actual data at time period $T+M$. Define two new random variables as $Z_1 = \max\{\widetilde{X}_1(T+M),...,\widetilde{X}_K(T+M)\}$ and $Z_2 = \min\{\widetilde{X}_1(T+M),...,\widetilde{X}_K(T+M)\}$. Under the assumption that the estimates of the growth models are independent, the joint probability density function that $Z_1 = z_1$ and $Z_2 = z_2$, $f_{Z_1Z_2}(z_1,z_2)$, is:*

$$f_{Z_1Z_2}(z_1,z_2) = \sum_{i=1}^{K} \sum_{j=1, j\neq i}^{K} f_i(z_1) f_j(z_2) \prod_{k=1,k\neq i,k\neq j}^{K} Pr\{z_2 < \widetilde{X}_k(T+M) < z_1\}$$

if $z_1 \geq z_2$, 0 otherwise. Then, the exact probability that range decreases is:

$$\begin{aligned} Pr(VR) = Pr\{\triangle_{T+M} < \triangle^0_{T+M}\} &= Pr\{z_1 - z_2 < \triangle^0_{T+M}\} \\ &= \int_{-\infty}^{\infty} \int_{z_2}^{z_2+\triangle^0_{T+M}} f_{Z_1Z_2}(z_1,z_2) dz_1 dz_2 \end{aligned}$$

For computational purposes, we develop a lower bound for the exact probability of variance reduction, which requires the summation of pairwise comparison probabilities:

Proposition 4. *Given that the range of the prior projections for time $T+M$ is $\triangle^0_{T+M}$, the probability that the range of the estimates reduces with the additional data from leading indicators is bounded below as:*

$$Pr\{VR\} \geq 1 - \sum_{(i,j):j>i} \left[1 - Pr\{-\triangle^0_{T+M} < \widetilde{X}_i(T+M) - \widetilde{X}_j(T+M) < \triangle^0_{T+M}\}\right]$$

Proof. By definition, the probability that the variation of the estimates reduces with the additional data is the probability that the maximum difference between the estimates of the growth curves is less than $\triangle^0_{T+M}$:

$$\begin{aligned}
Pr\{VR\} &= Pr\{\triangle_{T+M} < \triangle^0_{T+M}\} \\
&= \Pr\left\{\max_{(i,j):j>i} \{\mid \widetilde{X}_i(T+M) - \widetilde{X}_j(T+M) \mid\} < \triangle^0_{T+M}\right\}
\end{aligned}$$

which we can write more explicitly as:

$$\begin{aligned}
Pr\{VR\} &= \Pr\left\{\mid \widetilde{X}_i(T+M) - \widetilde{X}_j(T+M) \mid < \triangle^0_{T+M},\ \forall (i,j) : j > i\right\} \\
&= \Pr\left\{-\triangle^0_{T+M} < \widetilde{X}_i(T+M) - \widetilde{X}_j(T+M) < \triangle^0_{T+M},\ \forall (i,j) : j > i\right\}
\end{aligned}$$

To obtain a lower bound, we use the Bonferroni inequality, which states that $P\left(\bigcap_{i=1}^{n} A_i\right) \geq 1 - \sum_{i=1}^{n}[1 - P(A_i)]$, where A_i represents any condition. This inequality allows us to write the probability of a set of conditions in terms of the probabilities of individual conditions:

$$Pr\{VR\} \geq 1 - \sum_{(i,j):j>i} \left[1 - Pr\{-\triangle^0_{T+M} < \widetilde{X}_i(T+M) - \widetilde{X}_j(T+M) < \triangle^0_{T+M}\}\right].$$

□

The use of the normal probability distribution results in a specific case, which states that if the range of the mean estimates is smaller than the prior range, the variation reduces with probability equal to one when the number of leading indicators is infinitely large.

Corrolary 2. *For* $\triangle^0_{T+M} > max\{\mu_i' - \mu_j', \forall (i,j) : j > i\}, \quad lim_{m\to\infty} Pr(VR) = 1.$

Proof. For the normal posterior distributions, the lower bound is:

$$Pr\{VR\} \geq 1 - \sum_{(i,j):j>i} \left[1 - \left(\Phi\left(\frac{\triangle^0_{T+M} - [\mu_i' - \mu_j']}{\sqrt{\sigma_i'^2 + \sigma_j'^2}}\right) - \Phi\left(\frac{-\triangle^0_{T+M} - [\mu_i' - \mu_j']}{\sqrt{\sigma_i'^2 + \sigma_j'^2}}\right)\right)\right]$$

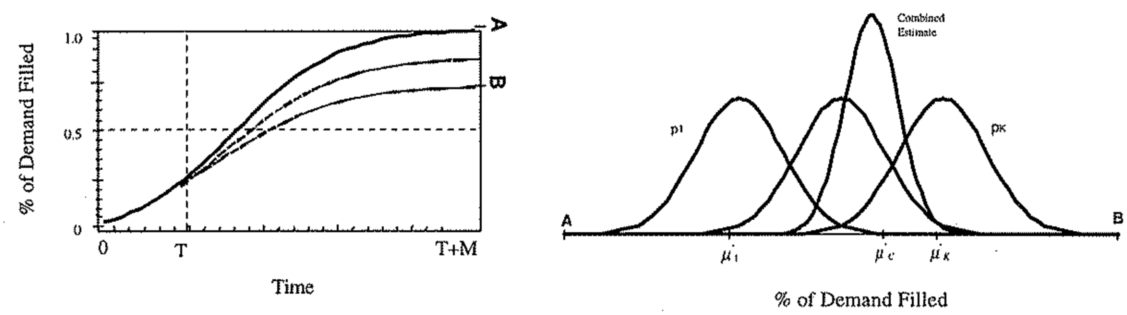
where $\Phi(\cdot)$ is the standard normal cumulative distribution.

As the number of leading indicators increases, $\sigma_i'^2 \to 0$ for $i = 1, ..., K$ (i.e., more confidence is given to the sample means). As a result, the lower bound and therefore the exact probability converge to one. □

4.2 Combined Life-Cycle Projection and Life-Cycle Demand Scenarios

In this section, we study the change in the variance of the combined estimate of the projections and describe demand in scenarios that are sampled from the probability distribution of the combined estimate.

The topic of forecast combination has been widely studied since the seminal work of Bates and Granger (1969). The motivation for combining forecasts has been to avoid the risk associated with the choice of a single "best" forecasting model and to aggregate information from different models. Several empirical studies show that combining forecasts improves the forecasting accuracy and reduces the variance of the forecasting errors (Timmermann, 2005). In light of these, we consider the combined forecast from the projections of the growth models and show that the variance of the combined forecast reduces with the additional data from leading indicators. Figure 4 illustrates the distribution of the estimates of individual models and combined forecast for any future period $T + M$.



(a) Prediction with growth models

(b) The probability distribution of the estimate of the growth models and the combined forecast at time T+M.

Figure 4: Distribution of the Estimates of the Growth Models and the Combined Forecast for time T+M

Among the many different combination methods, we consider the one that minimizes the variance of the combined forecast (Dickinson 1973). This method determines the weights of the individual model forecasts based on the covariance matrix of their estimates and takes into consideration the relative accuracy of the models. Since the covariance matrix is not known, the optimal weights must be estimated. There are several studies that propose different approaches to estimate the weights and compare the performances of these estimates. A common conclusion from these studies is that the combination methods that assume independence between individual forecast errors perform considerably better than those attempting to estimate the full covariance matrix and include the correlation of the estimates. Newbold and Granger (1974) and Winkler and Makridakis (1983) are among these studies and reach this conclusion from the analysis of 80 and 1001 time series, respectively. Under the assumption of independence, this method assigns each model a weight, p_k, that is inversely proportional to the variance of its estimation error.

Proposition 5. *The variance of the combined forecast, which is a weighted average of the models with weights being inversely proportional to the variance of the estimation errors, is smaller than the variance of its component model forecasts, and reduces with the use of additional data from leading indicators.*

Proof. Since the combined forecast is a linear combination of the independent normal random variables, from the Reproductive Property of Normal Distribution, the combined forecast is normally distributed. For the prior distribution, the probability distribution of the combined forecast is

$$\sum_{k=1}^{K} p_k \widetilde{X}_k(T+M) \sim N\left(\sum_{k=1}^{K} p_k \hat{X}_k(T+M|\Theta_T), \sum_{k=1}^{K} p_k^2 \sigma_k^2\right),$$

with the variance:

$$\sigma_c^2 = \sum_{k=1}^{K} p_k^2 \sigma_k^2 = \frac{1}{\sum_{k=1}^{K} 1/\sigma_k^2}$$

in which the variance of the combined forecast is no greater than the variance of its component model forecasts. The same holds for the combined forecast of the posterior projections.

Since the update with the additional data from leading indicators results in a demand distribution with smaller variance for the individual models (i.e., $\sigma_k^{\prime 2} \leq \sigma_k^2$, for $\forall k = 1, ..., K$) , the variance of the combined forecast is also smaller when additional data is used (i.e., $\sigma_c^{\prime 2} \leq \sigma_c^2$). $\square$

Given the Bayesian-updated probability distribution of the combined forecast, the results in Corollary 1 also hold for the combined forecast.

Generation of Life-Cycle Demand Scenarios and Variation in the Scenarios

We utilize the characterization of demand with continuous probability distributions using discrete demand scenarios. We use sampling to generate demand scenarios from the probability distribution of the combined forecast. The main thesis of this paper is that the posterior probability distribution of the combined forecast with smaller variance produces demand scenarios with smaller variation, which also include future demand signals.

One way to obtain life-cycle demand scenarios is to generate random samples from the probability distribution of the combined forecast. The variance of the sample average of these scenarios is σ_c^2/n, where n is the number of scenarios. There are several other sampling techniques that are suggested in the literature as appropriate for generating demand scenarios more

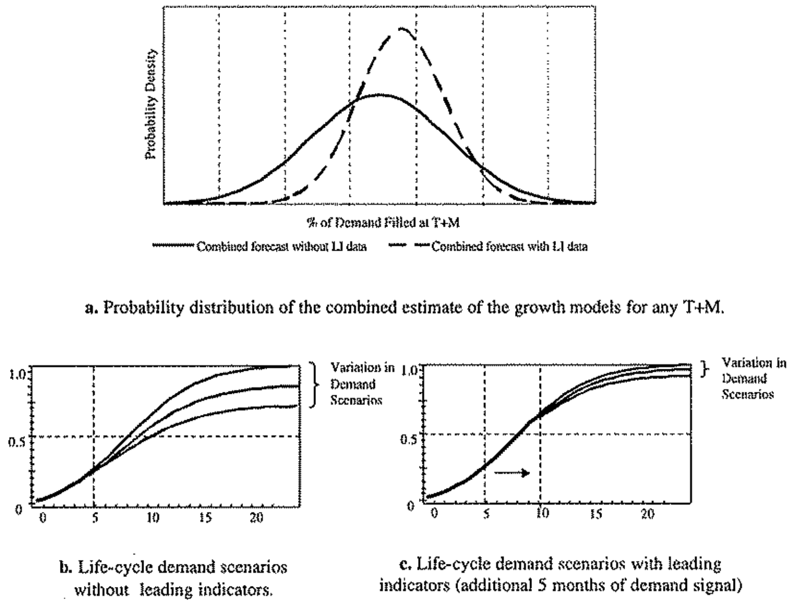


Figure 5: Sampling Life-Cycle Demand Scenarios from the Distributions of the Combined Forecasts

systematically and with smaller variation than those generated with random sampling. One such technique is a specific case of stratified sampling, which provides considerable intuitive appeal. The range of probable demand values are divided into segments of equal probability (Figure 5a), and a random observation is generated from each range as a possible demand scenario (Figure 5b, c). The variance of the sample average of the demand scenarios generated with this approach is $\sigma_c^2/n - \sum_{i=1}^{n}(\mu_i - \mu_c)^2/n^2$, where μ_i is the mean estimate of the i^{th} probability region, and μ_c is the mean estimate of the combined forecast (McKay et al., 1979). Given that the variance of the posterior combined forecast is smaller than the variance of the prior combined forecast, the demand scenarios from the posterior distribution are likely to have smaller variation than those from the prior distribution.

Through a Bayesian update, the variation in the life-cycle scenarios can be reduced by a significant margin. One of the reasons is that a small shift on the time axis might correspond to a drastic change on the curve, especially when the point of inflection is within this time shift.

5 Empirical Analysis

5.1 Experimental Data

For the purpose of empirically demonstrating our approach to characterize life-cycle demand patterns of short life-cycle products and the impact of additional information from leading-indicator products, we use disguised real data sets from three semiconductor manufacturing companies, which we will refer to as company A, company B, and company C. The latter two companies are members of the Semiconductor Research Corporation.

The data set provided by Company A is the weekly sales data of approximately 3000 products from January 2002 to December 2003. Some of the attributes of the products are the strategic business unit, business entity, fab process group, tester group, and package type. The data set from company B is the weekly sales data of 228 products from April 2004 to March 2006. All the products are in the same strategic business unit. The data set from company C is the weekly sales data of over 2000 products from February 2005 to March 2006. Some of the attributes of the products are the business unit, wafer-fab process group, package group, and technology code of each product. In order to lessen the impact of the short-term fluctuations that carry less information, we transformed the data into monthly figures. We then normalized the monthly data by taking into consideration that some fiscal months are 4 weeks long and others are 5 weeks long.

For each company, we demonstrate the characterization of the life-cycle demand pattern of an exogenously determined product group, as in Wu et al. (2006). The product group we consider for company A consists of 105 products that share the same fab capacity and belong to the same market segment. For company B, we consider all 228 products in the given strategic business unit. For company C, we study a group of 969 products in one strategic business unit.

5.2 Experimental Design

Meade and Islam (1998) studied 29 models from the technological forecasting literature using simulated and real data sets, and determined seven well-performing models in terms of their fitting and forecasting performances. In order to practically perform the analysis, we also use these models, namely, simple logistic, Mansfield, Gompertz, Floyd, Weibull, extended logistic, and cumulative log-normal. Simple logistic and Mansfield models are symmetric about the point of inflection, which is fixed at the value of cumulative proportion of 0.5. Gompertz and Floyd models are asymmetric with a point of inflection fixed at a value of cumulative proportion less than 0.5. The points of inflection for Weibull, extended logistic, and cumulative log-normal models are flexible and fall within a range of values that includes 0.5 depending on the parameter values. Different shapes of these models allow the characterization of a wide range of life-cycle patterns observed in practice.

We use growth models to describe the projections of the cumulative proportion of the total life-cycle demand that is met over time. Hence, we need to transform the given time-series data of the products into the scale that represents the cumulative proportion. The transformation requires an estimate of the expected total life-cycle sales, also known as the market potential. This quantity can be estimated as a parameter from the given data set, or its market estimate, which can be determined through market surveys or management judgement, can be provided as an input to the model. Heeler and Hustad (1980) and Tigert and Farivar (1981) are among the empirical studies that compare these two approaches and report that the accuracy of the forecasts increases significantly when the estimate of this quantity is an input to the estimation procedure since the estimation procedure results in unrealistic estimates of this parameter in order to increase the fitting performances of the models. In our study, we also use the market estimates of the expected total life-cycle sales.

We proceed as follows: There are 24 months of data for company A and company B, and 14 months of data for company C. We first use the initial 9-month period ($T = 9$) as the *estimation period*, in which the prior life-cycle projections of the individual models and the combined estimate are obtained. Next, we use the L-period demand signals from m leading-indicator products. This extends the estimation period to month $T + L$, the first T months from the actual data and the following L months from the leading-indicator-based data. The prior projections of the models and their combined forecasts are updated with this additional information. We use the remainder of the data, i.e., $[T + L + 1, 24]$ as the *validation period* to compare the performances of the estimates with and without additional data.

We first compare the range of the projections. We then study if the variance of the combined forecast reduces with the additional data. In order to verify if the reduction in variance is significant, we construct a paired-comparison experimental design, which compares the variance of the combined forecast with and without additional information over the same sets of observations. The variances of the combined forecasts for each period in the validation horizon over the data sets from all the companies provides a set of observations. The null hypothesis that the variance remains the same $H_0 : \sigma_C^2 = \sigma_C'^2$ is tested against the one-sided alternative hypothesis $H_1 : \sigma_C^2 > \sigma_C'^2$. The hypothesis is evaluated using a t-test statistic based on the standardized average difference between the variance of the prior estimates and the variance of the posterior estimates over all sets of observations.

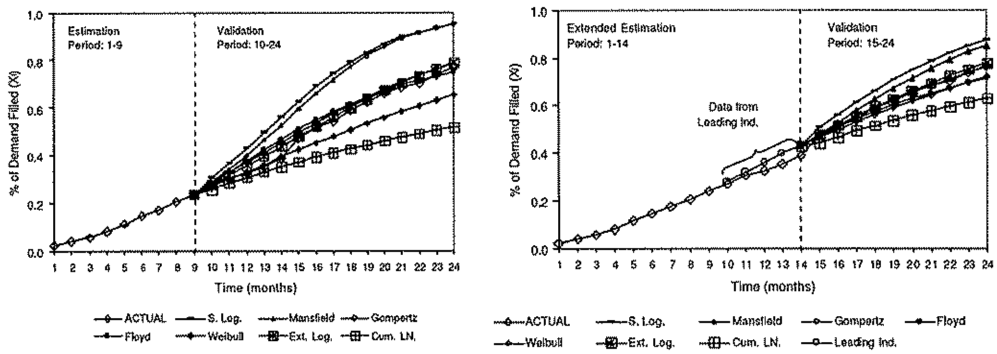
We identify three factors that can affect the results: (i) number of leading-indicator products, (ii) time lag of the leading-indicator products, and (iii) the bias of leading-indicator products. From Corrolary 1, it follows that variances of the estimates decrease at a diminishing rate with the number of leading indicators if all the leading indicators are unbiased. However, it is likely that data from some of the leading indicators are biased, and therefore the results are distorted. We want to test if there are biased leading indicators, and if so, how significantly they affect the results. In addition, the time-lag of a leading indicator may have an impact on its estimation quality, which in turn has an impact on the estimation quality of the growth models. To analyze the impact of these factors, we consider different number of leading indicators ($m = 5, 10, 15, 20$) that provide data up to a time lag of 5 months ($L = 1, ..., 5$). We present the experimental results in the next section.

5.3 Experimental Results

We start our analysis by illustrating the impact of a single leading indicator on the estimates of the models. Figure 6a plots the prior projections of the models obtained over the estimation period with 9 months of actual data for the product group of company A. The demand signals from a leading indicator with a time lag 5 months are used to extend the estimation period to month 14, and the projections of the models change as in Figure 6*b*.

The updated projection of each model with this leading indicator provides a sampling instance for the true projection of the model that would have been obtained over the estimation period with 14 months of actual data. Similarly projecting with several leading indicators gives the sampling distribution of the estimate of each model. Combining the sampling distribution with the prior distribution determined at month 9 gives the posterior distribution.

We perform the computational analysis for different combinations of m and L. The goals of the analysis are to study the following: (i) How closely do the sampling estimates with several leading indicators approximate the true estimates with actual data? (ii) Does the additional data from leading indicators result in projections with a smaller range? (iii) How significant is the



a. Prediction over the first 9 months of actual data (without a leading indicator)

b. Prediction over the 14 months of data (with 5-month demand signal from a leading indicator)

Figure 6: Projection of Cumulative Demand with Life-Cycle Growth Models

reduction in the variance of the combined forecast with respect to the number of leading indicators used? (iv) Do the time-lags of leading indicators have a notable effect on the results? (v) Are there any biased leading indicators? If so, what is the impact of bias on the mean and variance of the combined forecast? (vi) How does the forecasting performance change with the additional data?

Approximation Quality of the Estimates with Leading Indicators

Our objective in using Bayesian forecasting is to combine the future demand data in the highly volatile market environment with limited historical data. We use leading-indicator products as a means to capture this information. It is inevitable that the true prediction of leading indicators is critical to the success of our approach. In order to assess the impact of the prediction power of the leading indicators on the given data sets, we compare the mean estimates of the sampling distributions that are the averages of the estimates with several leading indicators against the true estimates that would have been obtained with the actual data instead of the leading-indicator-based data.

Figure 7 depicts the mean estimate of the sampling distribution under different combinations of m and L against the true estimates with actual data across the data sets from all the companies. The value for each combination of (m, L) is the average of the sampling means of all the seven models over all the periods in the part of the validation period over $[T+6, 24]$. Here, no one model or forecast horizon is given particular interest, so averages over all the models and horizons are taken. The sampling estimates are obtained using T-period actual data and L-period leading-indicator-based data as the estimation period. The true estimates are the averages of the

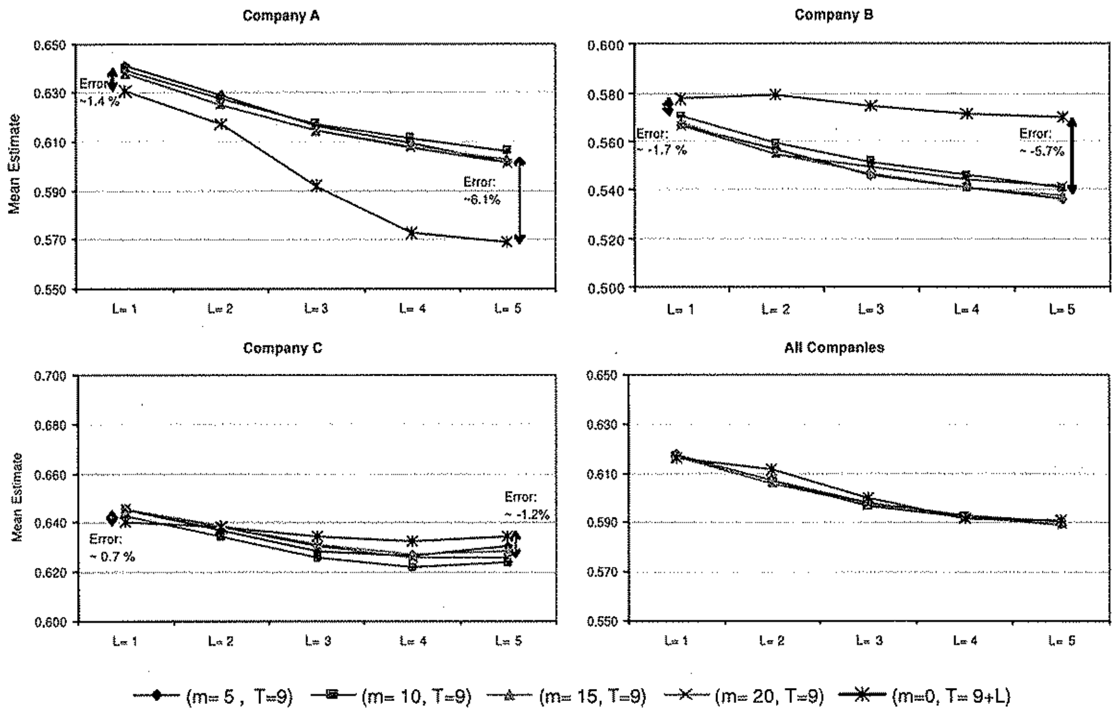


Figure 7: Comparison of sampling means against the true estimates with actual data.

estimates of all the models with the actual data used over $[1, T+L]$ as the estimation period, and shown with the lines that correspond to the tag $(m = 0, T = 9 + L)$.

As shown in the figure, the mean sampling estimates are within 6% of the true estimates, and the approximation quality has a tendency to degrade with increasing L. The changes are most notable for companies A and B. One possible reason is that data from leading indicators replaces the actual data over $[T+1, T+L]$. As L increases, the number of data points the leading-indicator-based data replaces increases, and approximation quality of the leading indicators decreases. This, in turn, has a negative impact on the approximation quality of the models. The average over all companies indicates the overall quality of the estimates. The results also suggest that the number of leading indicators has no significant impact on the mean sampling estimates. That is, the use of the top few leading indicators is sufficient to obtain mean estimates with approximately same accuracy.

Change in the Range of the Projections with Leading Indicators

We compare the expected value of the range of the posterior projections against the range of the prior projections. The results indicate that the range with the posterior projections is smaller for all combinations of m and L. That is, the estimates of the models become more similar with the additional data. Therefore, the condition for Corrollary 2 holds. Furthermore, the value of the lower bound on the probability that the range of the projections reduces with the additional data from leading indicators is equal to 1 for all combinations of m and L. This results in the range of the estimates decreasing for all the examples under consideration.

Value of Increasing the Number of Leading Indicators

The empirical analysis that compares the variances of the combined forecast over all the time horizons in the validation period across the data sets from the three companies rejects the null hypothesis that the variances of combined forecasts obtained from prior and posterior projections are equal ($p < 0.01$). That is, the reduction in variance is significant.

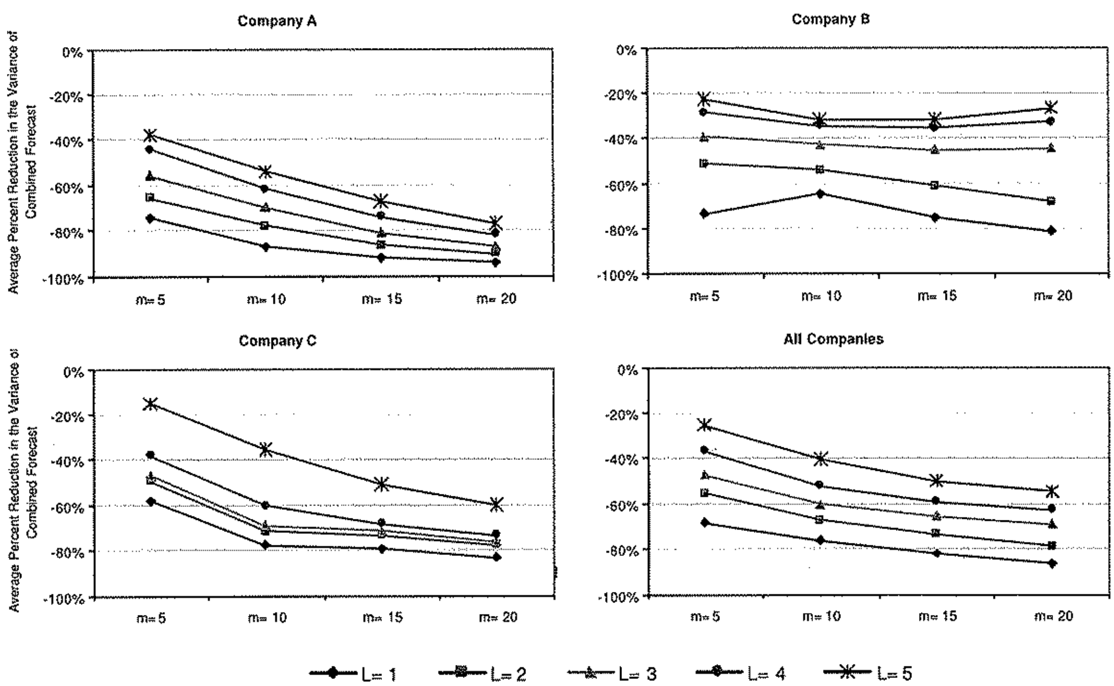


Figure 8: Percent reduction in the variance of the combined forecast with m leading indicators, each of which provides L periods of data.

Figure 8 displays the percent reduction in the variance of prior combined forecast under different combinations of (m, L) across the data from all the companies. The value shown for each combination of (m, L) is the average reduction over all the time horizons in the validation period. When all the leading indicators are unbiased, from Corrolary 1, it follows that the marginal value of the additional data (as measured by the reduction in the variance) diminishes when more leading indicators are considered. This is most apparent in the results for company A. However, the results for company B follow a conflicting pattern. This may be an indication of bias in the estimates of leading indicators, which we will discuss later in greater detail.

Another notable effect that is apparent in the results is that the rate of change in the reduction of variance tends to be less sensitive to m for small values of L when all leading indicators are unbiased. One possible reason is that as L decreases the approximation quality of the leading indicators are likely to be better and the variation between their estimates be smaller. As a result, the impact of m is smaller. Overall, the results indicate that it is possible to obtain $20-80\%$ reduction with the additional data across the data sets of all the companies.

Impact of Time-Lag of the Additional Information

Figure 8 reveals that reduction in variance tends to be less for larger values of L due to the relatively poor estimation quality of the leading indicators, as mentioned earlier. In addition, in case of unbiased leading indicators, for large values of m, the rate of change is less sensitive to L. This suggests the use of as many leading indicators as possible; however, this increases the possibility of having biased leading indicators.

Impact of Bias of Leading Indicators

Leading indicators are unbiased estimators of the actual data over the estimation period $[1, T]$. However, it is possible that the data they provide over $[T+1, T+L]$ is biased. To measure if there is any systematic bias in the L-period data provided by a leading indicator, we use a test that evaluates the null hypothesis that the errors have zero mean, for which a t-test is appropriate. The test results indicate that all of the 20 leading indicators for company A and company C provide unbiased data at 0.01 level of significance. However, 8 of the 20 leading indicators for company B provide biased data at 0.01 level of significance. This suggests that smaller reduction in variance for company B with m may be due to the biased leading indicators. This reveals that there is a possible trade-off in increasing m. Leading indicators are selected based on their relative ranking in terms of the absolute value of correlation coefficient between their demand time-series data and the shifted time-series data of the product group. As m increases, leading indicators with low absolute value of correlation coefficients are included. This increases the probability of inclusion of biased leading indicators, in which case increasing m may have a negative impact on the reduction in variance unlike the case where all leading indicators are unbiased.

The theoretical results imply that with bias, the variance of the estimates increases and the mean of the estimates may shift depending on the relative bias of the leading indicators. To empirically support these claims, we replaced the 8 leading indicators diagnosed as biased with unbiased ones. The average of the changes over all the periods in the validation period and over all the companies are summarized in Table 3 for all combinations of (m, L). For instance, the variance with biased leading indicators is 0.43% greater than the unbiased case when 20 leading indicators with a time lag 5 are considered. Also, the mean estimate of the combined forecast with biased leading indicators is 0.003 less than the mean estimate with all the leading indicators unbiased. This corresponds to an approximately 0.5% change in the mean estimate. Overall, the change in percent variance reduction is less than 5%, and the change in the mean is less than 0.012, which is about 2.5% of the mean with unbiased leading indicators.

	Change in % Variance Reduction				Change in Mean			
	m=5	m=10	m=15	m=20	m=5	m=10	m=15	m=20
L= 1	-0.15%	0.42%	0.16%	-0.02%	-0.005	-0.002	-0.003	-0.003
L= 2	2.31%	2.68%	0.79%	0.58%	-0.008	-0.010	-0.005	-0.005
L= 3	-1.32%	1.87%	0.39%	0.67%	-0.006	-0.007	-0.004	-0.004
L= 4	4.19%	3.68%	1.16%	0.47%	-0.012	-0.009	-0.005	-0.003
L= 5	1.34%	2.91%	1.04%	0.43%	-0.006	-0.008	-0.005	-0.003

Table 3: Impact of biasedness of leading indicators on the percent reduction in variance and on the mean estimate of the combined forecast (in terms of deviation from the estimates with unbiased leading indicators at 0.01 level of significance).

There are a few negative signs in the change in variance reduction which indicate the contradictory result that there is smaller reduction in variance with unbiased leading indicators. Since there is no exact method to diagnose the bias, we think that one possible reason is related to the method we use to identify unbiased leading indicators.

Forecasting Performance with Leading Indicators

An immediately raising question is how the forecasting performance is affected with the additional data. We measure the forecasting performance of the distributional estimates as the average performance of the resulting demand scenarios. This measurement provides a two-dimensional evaluation of the forecasts, in terms of accuracy and uncertainty, and allows the acceptance of a series of demand scenarios with slightly less accuracy but smaller variation over a series of demand scenarios with greater accuracy but larger variation.

We generate demand scenarios based on stratified sampling. In particular, we discretize the probability distribution into regions with equal probabilities, and take the midpoint of each region as a demand scenario. Table 4 compares the forecasting performances of the prior distribution and the posterior distribution with six scenarios, and gives the forecasting performance of the probability distribution obtained with actual data replacing the leading-indicator-based data, as a reference. The results are in terms of mean absolute percentage errors of the scenarios over the validation horizon $[15, 24]$. Due to the lack of data for company C over this time horizon, we present results only for companies A and B.

		Prior	Posterior				With Actual Data
		$m = 0$	$m = 5$	$m = 10$	$m = 15$	$m = 20$	$m = 0$
	L	$T = 9$	$T = 9$	$T = 9$	$T = 9$	$T = 9$	$T = 9 + L$
Company A	1	1.33%	2.42%	3.30%	4.04%	4.24%	4.22%
	2	1.33%	1.89%	2.58%	3.52%	3.83%	6.65%
	3	1.33%	2.20%	2.68%	3.78%	4.21%	11.58%
	4	1.33%	1.96%	2.57%	3.71%	4.38%	14.54%
	5	1.33%	1.93%	2.45%	3.55%	4.39%	14.58%
Company B	1	17.87%	17.30%	15.60%	15.91%	16.17%	14.75%
	2	17.87%	16.85%	15.57%	16.32%	16.16%	10.95%
	3	17.87%	16.94%	16.17%	16.82%	16.40%	10.08%
	4	17.87%	17.11%	16.44%	16.91%	16.59%	9.45%
	5	17.87%	17.23%	16.61%	16.87%	16.70%	8.97%

Table 4: Average forecasting performances of the combined forecasts with prior distribution, posterior distribution, and distribution with the actual data in terms of mean absolute percentage error over the forecast validation horizon across data sets from companies A and B.

The results indicate that the posterior distribution does not always perform better than the prior distribution, rather it displays an average performance between the prior distribution and the distribution with the actual data. The posterior distribution performs worse when the scenarios with actual data performs worse, since the leading-indicator-based data replace the actual data. Given the results in Table 4 and Figure 8, we assess the change in the forecast accuracy against the value of reduction in variance. For company A, around 2% deterioration in forecasting performance is compensated with approximately 70% reduction in variance. For company B, about 1.5% improvement in forecast accuracy is accompanied by 45% reduction in variance. The significant reduction in variance with a small change in the accuracy of the forecasts for these examples indicates the value of the proposed approach.

Companies can use the proposed demand-modeling approach in their operational decision-making processes. When the forecasting performance is not significantly worse, the reduction in forecast variation has implications on both the computational efforts required in the decision-making processes and the value of the expected operational costs. Smaller variation in demand scenarios either requires smaller number of scenarios to obtain solutions at a given significance level or returns solutions at a higher significance level with the same number of scenarios. This, in turn, has an impact on the computational requirements. In addition, the smaller variation has an impact on the expected operational costs due to the reduced uncertainty that needs to be hedged against.

6 Conclusions

In this paper, we propose a demand characterization approach for technology products with short life-cycle patterns and high demand volatility based on Bayesian forecasting. Several technological growth models characterize the life-cycle projections of the products. Leading-indicator products provide a means to learn from the uncertainty in future. Bayesian forecasting combines the life-cycle projections obtained using historical data with the information from leading-indicator products, and produces distributional estimates for the uncertain demand. Discrete demand scenarios describe the distributional estimates of the demand.

Given the observed data of the products and several growth models that characterize the technology life cycles, this approach allows us to develop a streamlined way to model demand scenarios for a particular market segment in a technology-driven market. The scenarios can be integrated into the supply-demand planning decision systems of the companies. Inclusion of the leading-indicator-based data reduces the variability in future demand scenarios. This, in turn, has a potential to improve the decision-making activities of the companies in terms of both computational time and expected total operational costs.

One drawback of this approach is, however, that leading indicators might provide biased information. In order to alleviate this affect, we use several leading indicators, the biases of which may cancel out, but at the expense of increased variance of the estimates.

Computational testing on real-world data provided by three semiconductor manufacturing companies suggests that the proposed approach is effective in capturing the short life-cycle nature of the products and early demand signals, and capable of reducing the uncertainty in the demand forecasts by more than 20%.

Acknowledgement

Our research is supported by the Semiconductor Research Corporation (SRC) grant 2004-OJ-1223.

Appendix

Proof of Proposition 2 We simplify the notation for notational convenience as:

$$\begin{aligned} X_k(T+M) &\rightarrow X_k & \widehat{X}_k(T+M|\Theta_{T+L}) &\rightarrow \widehat{X}_k \\ \widehat{X}_{kj}(T+M|\Theta^j_{T+L}) &\rightarrow \widehat{X}_{kj} & \widehat{X}_k(T+M|\Theta_T) &\rightarrow \widehat{X}_{k0} \end{aligned}$$

The posterior distribution is obtained by updating the prior distribution with the sampling distribution (J. Johnston 1984, S. J. Press 2003):

- The sampling distribution dictates that $\{\widehat{X}_{kj}, j = 1, ..., m\}$ are independently and identically distributed observations from $N(X_k, \tau_k^2)$. Their probability density function conditional on X_k is:

  $$p\left(\widehat{X}_{k1}, ..., \widehat{X}_{km}|X_k\right) = \left(2\pi\tau_k^2\right)^{-m/2} exp\left[-\frac{1}{2\tau_k^2}\sum_{j=1}^{m}(\widehat{X}_{kj} - X_k)^2\right]$$

  where X_k is unknown.

- The prior distribution of X_k is $N\left(\widehat{X}_{k0}, \sigma_k^2\right)$ with probability density function:

  $$p(X_k) = \left(2\pi\sigma_k^2\right)^{-1/2} exp\left[-\frac{1}{2\sigma_k^2}(X_k - \widehat{X}_{k0})^2\right]$$

- Posterior probability density function of X_k from Bayes' theorem is :

  $$p\left(X_k|\widehat{X}_{k1}, ..., \widehat{X}_{km}\right) = \frac{p\left(\widehat{X}_{k1}, ..., \widehat{X}_{km}|X_k\right) p(X_k)}{\int p\left(\widehat{X}_{k1}, ..., \widehat{X}_{km}|X_k\right) p(X_k)\, dX_k}$$

  where the denominator is a constant for a given sample data, σ_k^2, τ_k^2, and $\widehat{X}_{k0}$. Thus, the posterior probability density function can be written as:

  $$p\left(X_k|\widehat{X}_{k1}, ..., \widehat{X}_{km}\right) \propto p\left(\widehat{X}_{k1}, ..., \widehat{X}_{km}|X_k\right) p(X_k).$$

  When the probability density functions of the prior and sampling distributions are substituted and the constant multipliers are eliminated:

  $$p\left(X_k|\widehat{X}_{k1}, ..., \widehat{X}_{km}\right) \propto exp\left[-\frac{1}{2}\left\{\frac{\sum_{j=1}^{m}(\widehat{X}_{kj} - X_k)^2}{\tau_k^2} + \frac{(X_k - \widehat{X}_{k0})^2}{\sigma_k^2}\right\}\right].$$

Rearranging the terms in the exponent gives:

$$p\left(X_k|\widehat{X}_{k1},...,\widehat{X}_{km}\right) \propto exp\left[-\frac{1}{2}\frac{(X_k-\mu_k^{'})^2}{\sigma_k^{'2}}\right]$$

where

$$\begin{aligned}\mu_k^{'} &= \frac{1/\sigma_k^2}{1/\sigma_k^2+m/\tau_k^2}\,\widehat{X}_{k0}+\frac{m/\tau_k^2}{1/\sigma_k^2+m/\tau_k^2}\,\frac{1}{m}\sum_{j=1}^{m}\widehat{X}_{kj} \\ \frac{1}{\sigma_k^{'2}} &= \frac{1}{\sigma_k^2}+\frac{1}{\tau_k^2/m}\end{aligned}$$

Thus, the posterior distribution of the estimate of the life-cycle growth curve k is normal with mean $\mu_k^{'}$ and variance $\sigma_k^{'2}$.

References

F. M. Bass. A new product growth for model consumer durables. *Management Science*, 15(5): 215–226, 1969.

J. M. Bates and W. J. Granger. The combination of forecasts. *Operational Research Quarterly*, 20 (4):451–468, 1969.

J. S. Bonser and S. D. Wu. Procurement planning to maintain both short-term adaptiveness and long-term perspective. *Management Science*, 47(6):769–786, 2001.

C. Chatfield. Calculating interval forecasts. *Journal of Business and Economic Statistics*, 11(2): 121–133, 1993.

C. H. Chen. A lower bound for the correct subset-selection probability and its application to discrete-event system simulations. *IEEE Transactions on Automatic Control*, 41(8):1227–1231, 1996.

C. H. Chen, S. D. Wu, and L. Dai. Ordinal comparison of heuristic algorithms using stochastic optimization. *IEEE Transactions on Robotics and Automation*, 15(1):44–56, 1999.

H. C. Chen, C.H. Chen, and E. Yucesan. Computing efforts allocation for ordinal optimization and discrete event simulation. *IEEE Transactions on Automatic Control*, 45(5):960–964, 2000.

J. P. Dickinson. Some statistical results in the combination of forecasts. *Operational Research Quarterly*, 24(2):253–260, 1973.

G. D. Eppen and A. V. Iyer. Back-up agreements in fashion buying- the value of upstream flexibility. *Management Science*, 43(11):1469–1484, 1997a.

G. D. Eppen and A. V. Iyer. Improved fashion buying with Bayesian updates. *Operations Research*, 45(6):805–819, 1997b.

R. Fisher and A. Raman. Reducing the cost of demand uncertainty through accurate response to early sales. *Operations Research*, 44(1):87–99, 1996.

R. M. Heeler and T. P. Hustad. Problems in predicting new product growth for consumer durables. *Management Science*, 26(10):1007–1020, 1980.

T. Islam, D. G. Fiebig, and N. Meade. Modelling multinational telecommunications demand with limited data. *International Journal of Forecasting*, 18:605–624, 2002.

H. Kim. A Bayesian analysis on the effect of multiple supply options in a quick response environment. *Naval Research Logistics*, 50:937–952, 2003.

A. A. Kurawarwala and H. Matsuo. Product growth models for medium-term forecasting of short life cycle products. *Technological Forecasting and Social Change*, 57:169–196, 1998.

V. Mahajan and S. Sharma. A simple algebraic estimation procedure for innovation diffusion models of new product acceptance. *Technological Forecasting and Social Change*, 30:331–345, 1986.

V. Mahajan, E. Muller, and F.M. Bass. New product diffusion models in marketing: A review and directions for future research. *Journal of Marketing*, 54:1–26, 1990.

M. D. McKay, R. J. Beckman, and W. J. Conover. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21: 239–244, 1979.

N. Meade. Forecasting using growth curves- an adaptive approach. *Journal of Operational Research Society*, 36(12):1103–1115, 1985.

N. Meade and T. Islam. Prediction intervals for growth curve forecasts. *Journal of Forecasting*, 14: 413–430, 1995.

N. Meade and T. Islam. Technological forecasting– model selection, model stability, and combining models. *Management Science*, 44(8):1115–1130, 1998.

M. J. Meixell and S.D. Wu. Scenario analysis of demand in a technology market using leading indicators. *IEEE Transactions on Semiconductor Manufacturing*, 14:65–75, 2001.

D. P. Morton and E. Popova. A Bayesian stochastic programming approach to an employee scheduling problem. *IIE Transactions*, 36:155–167, 2004.

P. Newbold and C. W. J. Granger. Experience wih forecasting univariate time series and the combination of forecasts. *Journal of the Royal Statistics A*, 137:131–146, 1974.

S. J. Press. *Subjective and Objective Bayesian Statistics: Principles, Models, and Applications*. J. Wiley and Sons, Inc., New Jersey, 2003.

D. Schmittlein and V. Mahajan. Maximum likelihood estimation for an innovation diffusion model of new product acceptance. *Marketing Science*, 1(1):57–78, 1982.

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

F. Sultan, J. U. Farley, and D. R. Lehmann. A meta-analysis of diffusion models. *Journal of Marketing Research*, 27:70–77, 1990.

H. Takada and D. Jain. Cross-national analysis of diffusion of consumer durable goods in pacific rim countries. *Journal of Marketing*, 55:48–54, 1991.

D. Tigert and B. Farivar. The Bass new product growth model: A sensitivity analysis for a high technology product. *Journal of Marketing*, 45:81–90, 1981.

A. Timmermann. *Handbook of Economic Forecasting*, chapter Forecast Combination. North Holland, 2005.

R. L. Winkler and S. Makridakis. The combination of forecasts. *Journal of the Royal Statistical Society Series A*, 146(2):150–157, 1983.

S. D. Wu, B. Aytac, R. T. Berger, and C. A. Armbruster. Managing short life-cycle technology products for Agere systems. *Interfaces*, 36:234–247, 2006.

K. Zhu and U.W. Thonemann. An adaptive forecasting algorithm and inventory policy for products with short life cycles. *Naval Research Logistics*, 51(5):633–653, 2004.