Manufacturing & Service Operations Management

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

Why Are Forecast Updates Often Disappointing?

Kyle Cattani, Warren Hausman,

Why Are Forecast Updates Often Disappointing?

Kyle Cattani • Warren Hausman
The Kenan-Flagler Business School, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-3490
Management Science and Engineering, Stanford University, Stanford, California 94305-4026
kyle_cattani@unc.edu • hausman@stanford.edu

Demand forecasts do not become consistently more accurate as they are updated. We present examples demonstrating this counterintuitive phenomenon and some theoretical results to explain its occurrence. Specifically, we analyze the effect of demand randomness on forecast-update performance. A surprising result is that under various theoretical models involving demand randomness alone, updated forecasts will be less accurate between 30% and 50% of the time.
(*Forecast Performance; Forecast Updates; Exponential Smoothing*)

1. Introduction

Managers are often surprised to observe that many of their item-demand forecasts become less accurate as they are updated from period to period, especially in the periods immediately preceding the actual demand event. Many manufacturers and business managers express frustration that the latest forecast for the current month's demand is no better than the forecast for the same month created two or three months earlier. In this article we study the performance of demand forecast updates—the extent that updates improve or fail to improve the resulting forecast. We provide data from two companies to illustrate the phenomenon and then demonstrate theoretically how such behavior can result from applying exponential smoothing to random demands in both stationary and nonstationary environments. Our research shows that it may be unrealistic to expect an item-demand forecast made one period in advance (the $t - 1$ forecast, e.g., the forecast of December demand as of December 1) to be more accurate than one made three periods in advance (the $t - 3$ forecast, e.g., the forecast of December demand as of October 1).

This paper is organized as follows. Section 2 reviews related research. Section 3 describes various forecasting metrics, including our own item-level metric, and presents data from two companies illustrating the phenomenon. Section 4 contains theoretical results, and the remaining sections discuss the managerial implications of our results.

2. Previous Research

Forecast updating has been studied by various researchers. Hausman (1969) showed that forecast updates can be modeled as Markovian updates to existing forecasts, especially for items with a single realization per year, such as total seasonal demand for fashion goods or total supply of annual food crops. Heath and Jackson (1994) extended the Hausman model, creating a more general model for quantifying the effects of forecast error on production and distribution costs. Many other researchers have studied the use of Bayesian updating—the use of current data points to update the parameters characterizing the demand distribution (for example, see Azoury 1985).

For time-series forecasting with trend and/or seasonality, Winters (1960) developed a three-parameter

1523-4614/00/0202/0119$05.00
1526-5498 electronic ISSN

exponential-smoothing algorithm that is straightforward to implement. Brown (1963) published a thorough treatment of time-series forecasting and analyzed in depth the simple exponential-smoothing model.

In our research, we analyze the performance of forecast updates in a general demand environment discussed by Muth (1960), composed of two kinds of random components: one lasting a single time period (transitory) and the other lasting through all subsequent periods (permanent). Such a time series may be regarded as a random walk with "noise" superimposed. Muth proved that the exponential-smoothing forecast minimizes the forecast error variance for this demand pattern. We explore the performance characteristics of forecast updates in such an environment.

Various researchers have studied demand forecasting issues related to supply chains. Sterman (1989) considered nonlinear effects in a simple supply chain and provided approximations to model demand transmissions through the supply chain. Lee et al. (1997) considered the distortion of demand as it is transmitted through the supply chain. Gardner (1990) analyzed the impact of forecasting on inventory decisions and demonstrated that the choice of forecasting model is an important factor in determining the amount of investment needed to support any target level of customer service. For an ARMA process, Badinelli (1990) compared the steady-state sum of holding cost and stockout cost per unit time resulting from using the correct forecasting model with that resulting from using exponential smoothing (inappropriately) or from using the assumption that the probability distribution of demand is i.i.d. He concluded that correctly identifying the demand process is warranted.

To our knowledge, no one has directly analyzed the performance of forecast updates.

3. Forecast Performance Measures and Data

Most companies use a combination of forecasting methods. Many rely largely on judgmental methods. Others use the easy-to-calculate moving average or exponential-smoothing algorithms and massage the resulting forecast based on other available information (Sanders and Mandrodt 1994). It is common for companies to consider forecasts as a single number (e.g., the mean) and overlook the statistical properties of random demand.

Time-series methods require no information other than the past values of the variable being predicted. This assumes that information can be inferred from the pattern of past observations. A time-series analysis may attempt to isolate the patterns that arise most often, including trends and/or seasonality. The theoretical development of time-series methods over the last 40 years has been mainly for those cases of autocorrelated and nonstationary demand where time-series methods are optimal. For example, if a nonstationary demand is a random walk composed of temporary and permanent components, the time-series method of exponential smoothing has been shown to minimize the error variance (see Muth 1960).

Typical metrics for aggregate forecast performance include mean absolute deviation (MAD), mean squared error (MSE), and mean absolute percentage error (MAPE). It is possible to use these metrics to assess forecast-update performance, for example, by comparing the aggregate metric for $t - 3$ forecasts versus $t - 2$ or $t - 1$ forecasts. However, there is a difference between *aggregate* analysis of forecasts in general and the specific improvement of a forecast for demand for a *specific SKU* during a *specific time period* (e.g., forecasts of demand as of February and then as of March for SKU #123 during the month of April 2000). In particular, aggregate measures of forecast performance usually improve over time (see, e.g., Hausman 1969, Heath and Jackson 1994, Azoury 1985), but aggregate improvements can mask significant occurrences of individual item-level forecasts that become worse.

This is a significant concern for managers faced with demand-related decisions. For tactical production-planning and inventory-control purposes, there are important decisions that relate solely to individual forecasts of demand for individual SKUs during specific time periods. In this study, we investigate the performance of individual forecast updates by comparing point estimates (forecasts) of demand for a specific SKU as the updates get closer to the demand event.

Table 1 Summary of Forecast-Update Data from Two Companies

	Company One		Company Two	
	Number of Observations	Percentage of Observations	Number of Observations	Percentage of Observations
Strict Forecast Improvement	45	44.6%	97	46.6%
Strict Forecast Degradation	23	22.7%	49	23.6%
Unchanged Forecast	33	32.7%	62	29.8%
Total Number of Observations	101	100%	208	100%
Percent Same or Improved Forecast		77.2%		76.4%
Percent of Changed Forecasts which Degraded		33.8%		33.6%

Our item-level metric of forecast-update performance is the percentage of times that the updated forecast is more accurate than the earlier forecast, as determined by the absolute values of the two forecast errors for each SKU. The higher the percentage of forecast updates with smaller absolute errors, the better the forecast-update performance. For each individual SKU, the metric uses the following information: the forecast made at time $t - k$ for demand in period t, for $k = 1, 2, 3, \ldots$; and the actual demand in period t. Note that all of the standard aggregate measures of forecast error (MAD, MSE, MAPE, etc.) reduce to our item-level metric when applied to a single SKU and a single target period's demand.

As a secondary metric for forecast-update performance we also consider whether the i-step-ahead absolute forecast error is stochastically smaller than the j-step-ahead absolute forecast error. Successful forecast updates are defined as those that lead to stochastically smaller absolute forecast errors. While useful, this metric will fail to indicate significant occurrences of individual forecasts that become worse after the update, even if the updated forecast error is stochastically smaller.

We now turn to the application of our item-level metric to actual forecast updates in two high-technology firms. For our first example, we examine data from a company that provided forecasts and actual demand data from five products over a number of months, for a total of 101 forecast updates. Of these updates, 45 (44.6%) showed "strict" improvement, 23 (22.7%) were less accurate, and 33 (32.7%) were the same as the earlier forecast. Thus, performance was the same or better in 78 (77.2%) of the updated demand forecasts, but 33.8% of the updates where the forecast was changed (23/68) resulted in worse forecasts.

In our second example, we analyzed individual demand forecasts and actual demand data for 21 products over 10 months, for a total of 208 product-months. Comparing in particular the $t - 3$ versus $t - 1$ absolute forecast error, we found that 97 (46.6%) of the updates resulted in strictly improved forecasts, 49 (23.6%) of the forecasts became less accurate, and 62 (29.8%) were unchanged. Thus, performance improved or stayed the same in 76.4% of the updates. Of the 146 demand forecasts that were changed, 33.6% were less accurate than before. Table 1 presents these results in summary form. Note that the percentages in Table 1 across the two independent sets of data are remarkably similar.

Figure 1 shows the forecasts and actual demand for one of the products from Company Two, for which improved demand forecasts ($t - 3$ vs. $t - 1$) were obtained in only 4 of 10 months: January, February, April, and May.

The MAD, MSE, and MAPE for the $t - 3$, $t - 2$, and $t - 1$ forecasts for this individual product's data are shown in Table 2 (aggregating across all 10 months). Note that updated forecasts are better under each of these aggregate measures except for MSE from $t - 3$ to $t - 2$, whereas the update performance for these 10 individual updates ($t - 3$ vs. $t - 1$) under our item-level metric is only 40%.

Figure 1 **Product *Y* Forecasts and Actual Demand**

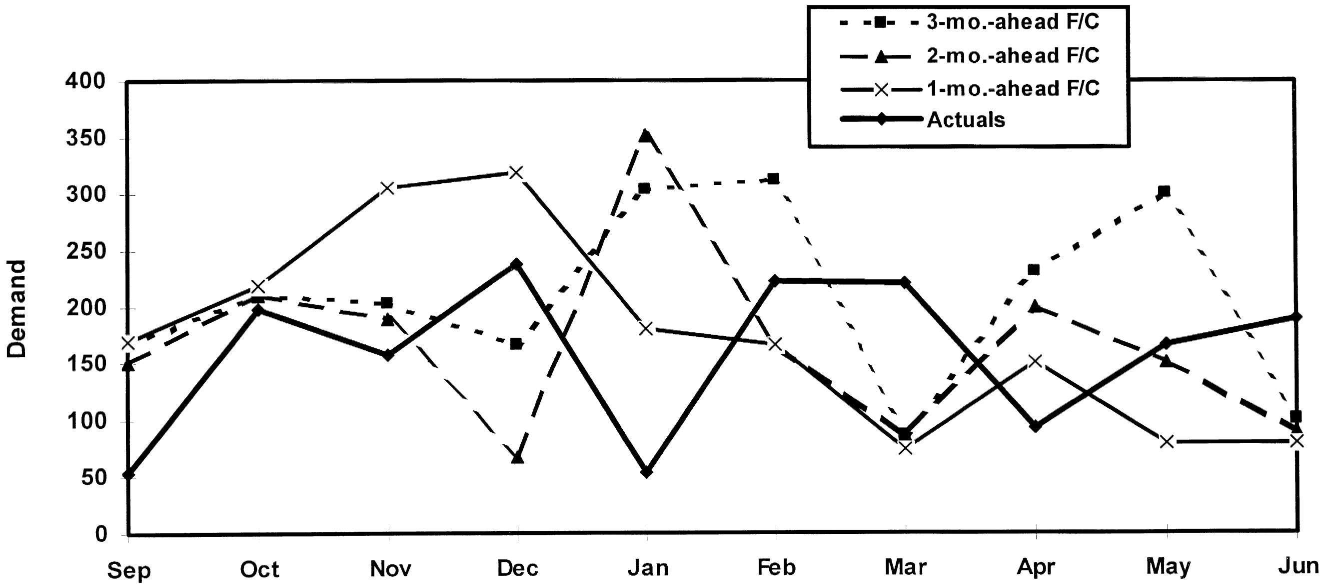


Table 2 **Aggregate Forecast Performance for Product *Y***

	$t-3$	$t-2$	$t-1$
MAD	108	102	95
MSE	15,333	17,096	15,508
MAPE	113%	111%	86%

4. Insights from Theory: Exponential-Smoothing Forecasts

Why do individual forecast updates show such poor improvement rates when aggregate demand forecasts for the same products over the same periods on average improve? Insights into this question can be found in a theoretical analysis of forecast-update performance in models where exponential smoothing is used. In this section we analyze exponential-smoothing models in both stationary and nonstationary demand environments.

Define $F_{t,t+\tau}$ as the forecast made after period t for demand in period $t+\tau$. Let $F_{t,t+1}$ be defined as F_{t+1}. The simple exponential-smoothing forecast after time t with smoothing constant α is expressed as:

$$F_{t+1} = \alpha x_t + (1-\alpha)F_t, \tag{1}$$

where x_t is the actual demand in period t and α is the smoothing constant, with $0 < \alpha \leq 1$. For the environments we consider, the forecast after time t, F_{t+1}, is for demand in period $t+1$ and in all subsequent periods as well.

4.1. Stationary Demands

We first review the theoretical limitations of exponential smoothing for forecasting stationary demand and then determine forecast-update performance under such conditions. We use Nahmias's definition of a stationary time series as one in which each observation can be represented by a constant plus a random fluctuation (Nahmias 1997, p. 71); that is, $x_t = \mu + \epsilon_t$, where μ is an unknown constant corresponding to the mean of the series and ϵ_t is an i.i.d. random error with mean zero and variance σ^2.

Assuming that the demand process is stationary, Brown (1963) showed that exponential-smoothing forecasts are unbiased, i.e., $E[F_{t+1}] = \mu$ as $t \to \infty$. Despite this, when demand is stationary, exponential smoothing is not necessarily the most appropriate choice in terms of standard aggregate forecast measures such as MSE or MAD. For example when demand is given as above, by $d_t = \mu + \epsilon_t$ with ϵ_t an i.i.d. random variable, then a simple (i.e., equally-weighted) average of all past demands is superior to exponential smoothing, since it will have a smaller expectation for both MSE and MAD.

For normally-distributed demands we state the following theorem (see Appendix A for proof).

THEOREM 1. *Assume demand is from a stationary time series $x_t = \mu + \epsilon_t$, where μ is an unknown constant and ϵ_t is i.i.d. random error $\sim N(0, \sigma^2)$. Let $F_{t+1} = \alpha x_t + (1 - \alpha) F_t$ be the exponential-smoothing forecast at time t and let $Y_k = x_t - F_{t-k,t}$ be the k-step-ahead forecast error. Then, as $t \to \infty$, $Pr[|Y_i| \le |Y_j|] = 0.5, \forall i \ne j$.*

Theorem 1 states, for example, that given stationary demand, the $t - 3$ exponential-smoothing forecast will be at least as accurate as the $t - 1$ forecast 50% of the time. It is also easy to show that if there is known trend and/or seasonality in the demand data and exponential-smoothing forecasts are applied to detrended and deseasonalized data, Theorem 1 applies directly and the same result carries over to this more general setting.

It is also informative to consider forecast-update performance over all three forecasts preceding the demand event.

LEMMA 1. *Given stationary demands with i.i.d. normal errors under exponential smoothing,* $1/6 \le \Pr[|Y_1| \le |Y_2| \le |Y_3|] < 1/4$.

PROOF. See Appendix A.

Lemma 1 states that forecast performance for a specific SKU over the three forecasts before the demand event will improve over *both* updates ($t - 3$ to $t - 2$ and $t - 2$ to $t - 1$) less than 25% of the time under the stated assumptions.

The following theorem considers an alternate definition of forecast-update improvement: that the i-step-ahead absolute forecast error is stochastically larger than the j-step-ahead absolute forecast error when $i < j$.

THEOREM 2. *Let $Y_k = x_t - F_{t-k,t}$ be the k-step-ahead forecast error. For an exponential-smoothing forecast of stationary demands with i.i.d. normal errors, $|Y_i|$ is stochastically equivalent to $|Y_j|$, $\forall i \ne j$, as $t \to \infty$.*

PROOF. This follows directly from the fact that Y_i and Y_j have the same distribution, as shown in the proof of Theorem 1. □

Theorem 2 considers our secondary metric and demonstrates that under exponential smoothing and stationary demands, the distribution of absolute forecast errors remains stochastically the same, even after updates.

To summarize, in the stationary case using exponential smoothing, forecast updates are equally likely to create larger forecast error as smaller forecast error (Theorem 1) and the absolute forecast error from the updated forecast is stochastically equivalent to the preupdate absolute forecast error (Theorem 2).

4.2. Nonstationary Demands (Random Walk)

We now analyze forecast-update performance for a general random-walk process with both permanent and temporary components, as described in Muth (1960). In particular, suppose that x_t is generated as follows:

$$x_t = u_t + \epsilon_t \text{ with } \epsilon_t \sim N(0, \sigma_N^2), \text{ and} \tag{2}$$

$$u_t = u_{t-1} + \xi_t \text{ with } \xi_t \sim N(0, \sigma_{RW}^2). \tag{3}$$

ϵ_t is the temporary component of demand (e.g., noise), while u_t is the permanent component of demand. The ϵs are serially independent, the ξs are serially independent, and the ϵs and the ξs are mutually independent.

Using Equations (2) and (3), demand may also be expressed as:

$$x_t = \epsilon_t + \xi_t + \xi_{t-1} + \xi_{t-2} + \dots \tag{4}$$

Forecasts are generated once again using Equation (1).

LEMMA 2. *Given a random-walk demand process, $0.5 \le Pr[|Y_1| \le |Y_2|] \le 0.6476$ and $0.5 \le Pr[|Y_1| \le |Y_3|] \le 0.6959$.*

PROOF. See Appendix B.

In Figure 2 we show how the probability of item-level forecast improvement depends on the ratio of the variance parameters of the data-generating process. As expected, forecast-update performance is worse when the standard deviation of the permanent component of demand (σ_{RW}) is small relative to the temporary component of the demand (σ_N).

We note that with this nonstationary demand, updated forecasts have stochastically smaller absolute forecast errors than the earlier forecasts. Hence, under our secondary metric, forecasts improve with each update.

4.3 Reason for Counterintuitive Behavior of Exponential-Smoothing Models

The forecast-update performance demonstrated in § 4.1 and 4.2 is due to the inherent demand randomness of the problem and is not due to the "incorrect" forecasting model. In particular, although exponential smoothing is an incorrect model when demand is stationary, Theorem 1 also holds even if the correct model is used. The forecasting model that minimizes error variance for the stationary-demand environment of §4.1 is an unweighted average of all past data points. Theorem 3 proves that even using an unweighted average of all past data points, the Theorem 1 results are obtained (see Appendix C). For a nonstationary demand (random walk), the results of §4.2 were developed using the correct forecasting model.

5. Managerial Implications

5.1. "Churning" Forecasts

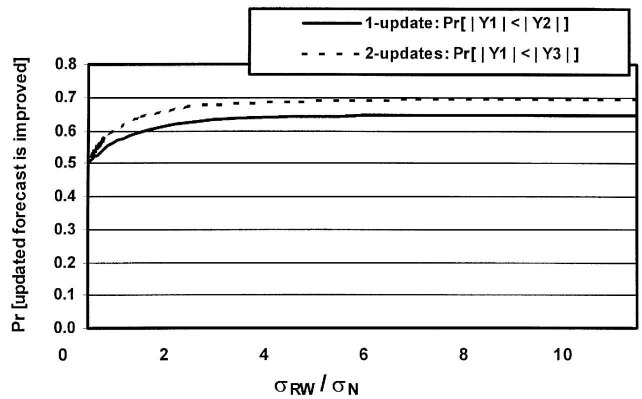
Changes in forecast accuracy can have a significant impact on a firm's activities, causing sometimes costly and possibly unnecessary reactions in numerous functional areas. In examining such changes over two consecutive updates, there are four possible outcomes:[1]

(1) $t - 3$ to $t - 2$ forecast accuracy better, $t - 2$ to $t - 1$ forecast accuracy better;
(2) $t - 3$ to $t - 2$ forecast accuracy better, $t - 2$ to $t - 1$ forecast accuracy worse;
(3) $t - 3$ to $t - 2$ forecast accuracy worse, $t - 2$ to $t - 1$ forecast accuracy better; and
(4) $t - 3$ to $t - 2$ forecast accuracy worse, $t - 2$ to $t - 1$ forecast accuracy worse.

In case 1, of course, any reactions to the updated

[1]With exponential smoothing, stationary demand, and low values of α, each outcome occurs with probability around 1/4. For higher values of α, the probabilities approach 1/6, 1/3, 1/3, and 1/6 for the four outcomes, respectively. (See the proof of Lemma 1 in Appendix A.)

Figure 2 Forecast-Update Performance with Random Walk



forecast would be appropriate. Note that in cases 2, 3, and 4 there will be "churning" (i.e., a response to a forecast update where the new forecast is worse). In case 2, a reaction to the $t - 2$ forecast turns out to be appropriate, but a reaction to the $t - 1$ forecast would create churning. In cases 3 and 4, a reaction to the new forecast in period $t - 2$ can cause inefficiencies in material purchases, labor planning, capacity calculations, subcontracting, etc., since the changes turn out to be in the wrong direction. Such changes usually become more costly the closer to the demand event. If forecast updates are likely to fail to improve the forecasts, and such updates cause increased efforts in purchasing and other functional areas, this may be a compelling cause to increase the forecaster's bias toward unchanged forecasts at some point prior to the demand event.

5.2. Forecast Error Versus Magnitude of Forecast Update

Better forecast-update performance may be associated with larger, more significant changes to the forecast. We tested this hypothesis using the data in the second example of §3. We partitioned the subset of 146 forecast updates which resulted in an altered forecast (of which 97 or 66.4% improved the forecast) into two groups of equal size based on whether the size of the update (measured in percentage terms) was greater or less than the median. The higher-percentage-update group had reduced forecast error 78.1% of the time, whereas the lower-percentage-update group had reduced forecast error only 54.8% of the time. This difference is significant at the 2% level using a normal approximation to the binomial distribution. We conclude that forecast-update performance is related statistically to the size of the forecast update in this sample. However, although those forecast updates that were larger than the median in percentage terms performed significantly better than those that were smaller than the median, even in the former case 22% of these updates degraded the forecast.

5.3. Forecast-Update Performance Versus Forecast Accuracy

Managerial concerns about poor forecast-update performance may be misplaced. In fact, prescriptions that increase the forecast accuracy in earlier periods likely will reduce update performance, since it is easier to improve upon a bad forecast than upon a good forecast. On the other hand, if update performance is very good, it may be a signal that early forecasts are highly inaccurate. Since high levels of forecast-update performance may be a symptom of inaccurate forecasts, forecast-update performance should be considered only in conjunction with forecast accuracy.

5.4. Other Causes of Poor Forecast-Update Performance

There are several other possible causes of poor forecast-update performance that remain as topics for future research. Badinelli (1990) examines the costs of using the incorrect forecasting model. While we have shown here that even using the correct forecasting model can generate poor forecast-update performance, it is conceivable that using an incorrect model can generate worse forecast-update performance.

Forecast-update performance may suffer if demand data are not from the true, underlying demand. This may occur if shipment data are used as a proxy for demand data, or if demand is distorted in the supply chain in a bullwhip effect (see Lee et al. 1997), or if quarterly or annual sales incentives create end-of-period anomalies.

Finally, updated forecasts may worsen in accuracy if forecasters overreact to the most current data. In this case, forecast updates might benefit from being dampened, similar to the effect of the smoothing parameter in exponential smoothing.

6. Conclusions

The phenomenon of forecast updates that worsen forecast error contradicts the commonly accepted notion that the shorter the forecast horizon, the more accurate the forecast will be. We have shown that demand-forecast updates in the last several periods before the demand event decrease forecast accuracy 30% to 50% of the time. We have also presented several theoretical results demonstrating that this phenomenon can occur with random demand and common exponential-smoothing forecast models.

We conclude that many manufacturers and business managers may have unrealistic expectations concerning forecast accuracy improvements as forecasts are updated in the final periods before the demand event.

Acknowledgments The authors wish to thank Corey Billington and Tom Davis of HP and Chris Givens of Adexa for sharing industrial instances of this problem. Helpful suggestions by Margaret Brandeau, Robert Carlson, Hau Lee, Seungjin Whang, and several referees and editors of *M&SOM* are also appreciated.

Appendix A

Proof of Theorem 1. Let α be the smoothing coefficient for the exponential-smoothing forecast. Brown (1963) showed that as $t \to \infty$, $E[F_t] = \mu$ and $\mathrm{Var}[F_t] = \alpha\sigma^2/(2 - \alpha)$.
Then

$$Y_i \sim N\left(\mu - \mu, \sigma^2 + \frac{\alpha}{2 - \alpha}\sigma^2\right) = N\left(0, \frac{2}{2 - \alpha}\sigma^2\right),$$

since x_t is independent of $F_{t-i,t}$.

Also

$$Y_j \sim N\left(\mu - \mu, \sigma^2 + \frac{\alpha}{2 - \alpha}\sigma^2\right) = N\left(0, \frac{2}{2 - \alpha}\sigma^2\right),$$

since x_t is independent of $F_{t-j,t}$.

We need to show that $\Pr[|Y_i| \leq |Y_j|] = 0.5$. We note that Y_i and Y_j are *not* independent; let ρ be the correlation between Y_i and Y_j. Y_i and Y_j are distributed normally with identical means and variances ($\mu' = 0, \sigma'^2 = 2\sigma^2/(2 - \alpha)$) and their joint density is multinormal:

$$f(y_i, y_j) = \left(\frac{1}{2\pi\sigma'^2\sqrt{1 - \rho}}\right)e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{y_t^2}{\sigma'^2}\right) - 2\rho\left(\frac{y_i y_j}{\sigma'^2}\right) + \left(\frac{y_j^2}{\sigma'^2}\right)\right]}.$$

Because the variances of Y_i and Y_j are equal, the joint density is symmetric about the lines $y_i = y_j$ and $y_i = -y_j$. This symmetry directly leads to the result: $\Pr[|Y_i| \leq |Y_j|] = 0.5$ □

Proof of Lemma 1. The covariances of Y_1, Y_2, and Y_3 are: $\mathrm{Cov}[Y_1, Y_2] = \mathrm{Cov}[Y_2, Y_3] = \sigma^2[(2 - \alpha + \alpha(1 - \alpha))/(2 - \alpha)]$ and $\mathrm{Cov}[Y_1, Y_3] = \sigma^2[(2 - \alpha + \alpha(1 - \alpha)^2)/(2 - \alpha)]$. The variances were determined in Theorem 1. Thus, the joint density $f(y_1, y_2, y_3)$ of Y_1, Y_2, and Y_3 is $(1/(2\pi|\Sigma|^{1/2}))e^{-Q/2}$ where Σ is the covariance matrix and $Q = [y_1\ y_2\ y_3]\Sigma^{-1}[y_1\ y_2\ y_3]$. The probability is calculated through integration of $f(y_1, y_2, y_3)$ over the appropriate space: $\iiint_{|y_1|\le|y_2|\le|y_3|} f(y_1, y_2, y_3)dy_3dy_2dy_1$. We calculated the integral numerically in Mathematica for values of α between 0 and 1 using a left-point approximation of the triple integral with each variable divided into 200 increments. The results are shown in Figure 3. □

The bounds occur at $\alpha = 0$ (1/4) and at $\alpha = 1$ (1/6). These bounds can be understood intuitively as follows. For very small α, the forecasts are essentially unchanged, although they are nudged very slightly toward the latest data point. Forecasts are unbiased and remain so with each very small update. The lemma holds for very small α if the last three demands are either all above the forecast (which occurs with probability 1/8) or all below the forecast (also with probability 1/8) for a total probability of 1/4. The updated forecasts will have been nudged twice toward the actual value, ensuring that $[|Y_3|\le|Y_2|\le|Y_1|]$. For $\alpha = 1$, each forecast equals the prior data point. This means that the errors Y_i are merely a comparison of X_t with X_{t-1}, X_{t-2}, and X_{t-3}. Since each X is a random draw from a stationary distribution, all orderings of the resulting errors are equally probable. The six possible orderings are $[|Y_3|\le|Y_2|\le|Y_1|]$, $[|Y_3|\le|Y_1|\le|Y_2|]$, $[|Y_2|\le|Y_3|\le|Y_1|]$, $[|Y_2|\le|Y_1|\le|Y_3|]$, $[|Y_1|\le|Y_2|\le|Y_3|]$, $[|Y_1|\le|Y_3|\le|Y_2|]$, each with probability 1/6.

Appendix B

PROOF OF LEMMA 2. Muth (1960) shows the optimal smoothing constant, α^*, for this forecasting process to be

$$\alpha^* = \frac{\sigma_{RW}}{\sigma_N}\sqrt{1 + \frac{\sigma_{RW}^2}{4\sigma_N^2}} - \frac{\sigma_{RW}^2}{2\sigma_N^2}. \tag{B1}$$

We show the calculations for $\Pr[|Y_2|\le|Y_1|]$. The calculation for forecast updates over two periods, i.e., $\Pr[|Y_3|\le|Y_1|]$, is similar, and we show the results only. Use Equations (4) and (1) to obtain variances of Y_2 and Y_1:

$$\sigma_{Y_2}^2 = \frac{2\sigma_N^2}{2-\alpha} + \frac{(1 + 2\alpha - \alpha^2)\sigma_{RW}^2}{\alpha(2-\alpha)} \text{ and } \sigma_{Y_1}^2 = \frac{2\sigma_N^2}{2-\alpha} + \frac{\sigma_{RW}^2}{\alpha(2-\alpha)}.$$

The covariance between Y_2 and Y_1 can be shown to be:

$$\mathrm{Cov}[Y_2, Y_1] = \sigma_N^2\left[\frac{2-\alpha^2}{2-\alpha}\right] + \sigma_{RW}^2\left[\frac{1+\alpha-\alpha^2}{\alpha(2-\alpha)}\right].$$

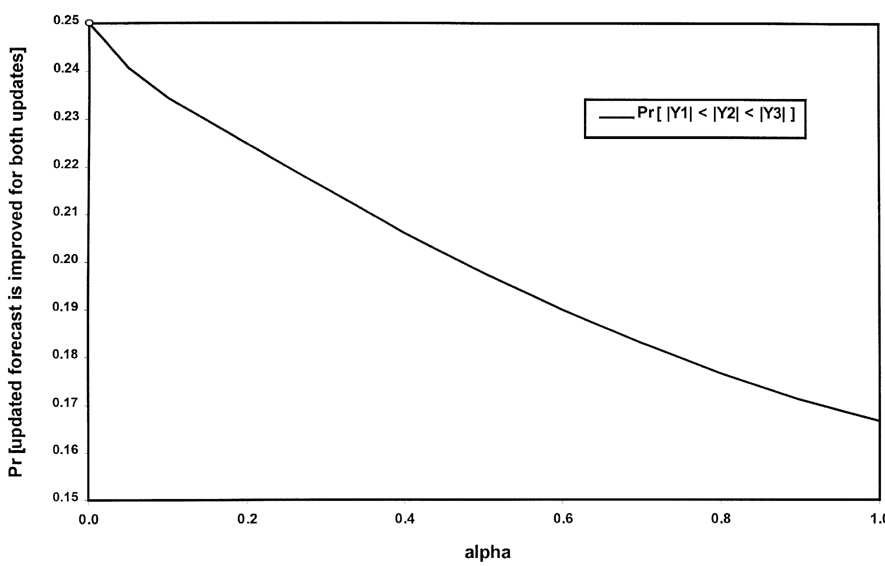
The correlation between Y_2 and Y_1 is $\rho = \mathrm{Cov}[Y_2, Y_1]/\sigma_{Y_2}\sigma_{Y_1}$. Using the optimal values for α (Equation (B1)), we can determine $\Pr[|x_t - F_{t-1}|\le|x_t - F_t|]$ (i.e., $\Pr[|Y_2|\le|Y_1|]$) by integrating the bivariate density function $f(y_1, y_2)$ over the Y_2Y_1 plane where $|Y_2| \le |Y_1|$: $\iint_{|y_2|\le|y_1|} f(y_1, y_2)dy_2dy_1$. The only independent variables are σ_N and σ_{RW}, and only their ratio is significant. We used Mathematica and numerically integrated to create Figure 2. □

Appendix C

THEOREM 3. *Assume demand is from the time series $x_t = \mu + \epsilon_t$, where μ is an unknown constant corresponding to the mean of the series and ϵ_t is i.i.d. random error with a distribution $\sim N(0, \sigma^2)$. Let $F_{t+1} = 1/t\ \Sigma_{i=1}^t x_i$ be the forecast at time t and let $Y_k = x_t - F_{t-k,t}$, the k-step-ahead forecast error. Then, as $t \to \infty$, $\Pr[|Y_i| \le |Y_j|] = 0.5$, $\forall \ne j$ where i and j are small relative to t.*

PROOF. One can easily show that $E[F_{t+1}] = \mu$ and $\mathrm{Var}[F_{t+1}] = \sigma^2/t$, which approaches 0 for large t.
Then

Figure 3 Forecast-Update Performance under Consecutive Updates



$$Y_i \sim N\left[\mu - \mu, \sigma^2 + \frac{1}{t - i - 1}\sigma^2\right) \xrightarrow[t\to\infty]{} N(0, \sigma^2),$$

since x_t is independent of $F_{t-i,t}$.

Also

$$Y_j \sim N\left(\mu - \mu, \sigma^2 + \frac{1}{t - j - 1}\sigma^2\right) \xrightarrow[t\to\infty]{} N(0, \sigma^2),$$

since x_t is independent of $F_{t-j,t}$.

We need to show that $\Pr[|Y_i| \le |Y_j|] = 0.5$. We note that Y_i and Y_j are *not* independent; let ρ be the correlation between Y_i and Y_j. Y_i and Y_j are distributed normally with identical means and variances that are asymptotically identical. Their joint density is multinormal:

$$f(y_i, y_j) = \left(\frac{1}{2\pi\sigma'^2\sqrt{1-\rho}}\right)e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{y_i^2}{\sigma'^2}\right) - 2\rho\left(\frac{y_i y_j}{\sigma'^2}\right) + \left(\frac{y_j^2}{\sigma'^2}\right)\right]}.$$

Because the variance of Y_i and Y_j are equal, the joint density is symmetric about the lines $y_i = y_j$ and $y_i = -y_j$. This symmetry directly leads to the result: $\Pr[|Y_i| \le |Y_j|] = 0.5$ □

References

Azoury, K. S. 1985. Bayes solution to dynamic inventory models under unknown demand distribution. *Management Sci.* **31**(9) 1150–1160.

Badinelli, R. D. 1990. The inventory costs of common misspecification of demand-forecasting models. *Internat. J. Production Res.* **28**(12) 2321–2340.

Brown, R. G. 1963. *Smoothing, Forecasting and Prediction of Discrete Time Series.* Prentice Hall, Englewood Cliffs, NJ.

Gardner, E. S. 1990. Evaluating forecast performance in an inventory control system. *Management Sci.* **36**(4) 490–499.

Hausman, W. H. 1969. Sequential decision problems: A model to exploit existing forecasts. *Management Sci.* **16** B93–B111.

Heath, D. C., P. L. Jackson. 1994. Modeling the evolution of demand forecasts with application to safety stock analysis in production/distribution systems. *IIE Trans.* **26**(3) 17–30.

Lee, H. L., V. Padmanabhan, S. Whang. 1997. Information distortion in a supply chain: The bullwhip effect. *Management Sci.* **43**(4) 546–558.

Muth, J. F. 1960. Optimal properties of exponentially weighted forecasts. *J. Amer. Statist. Assoc.* **55** 299–306.

Nahmias, S. 1997. *Production and Operations Analysis,* 3rd ed. Irwin, Chicago, IL.

Sanders, N. R., K. B. Mandrodt. 1994. Forecasting practices in US corporations: Survey results. *Interfaces* **24**(2) 92–100.

Sterman, J. D. 1989. Modeling managerial behavior: Misperceptions of feedback in a dynamic decision making experiment. *Management Sci.* **35**(3) 321–329.

Winters, P. R. 1960. Forecasting sales by exponentially weighted moving averages. *Management Sci.* **6** 324–342.

The consulting Senior Editor for this manuscript was Stephen Graves. This manuscript was received on July 22, 1997, and was with the authors for 5 revisions. The average review cycle time was 62.6 days.