Promotional Analysis and Forecasting for Demand Planning: A Practical Time Series Approach

Michael Leonard, SAS Institute Inc.
Cary, NC, USA

Abstract

Many businesses use sales promotions to increase the demand for or visibility of a product or service. These promotions often require increased expenditures (such as advertising) or loss of revenue (such as discounts), and/or additional costs (such as increased production). Business leaders need to determine the value of previous or proposed promotions. One way to evaluate promotions is to analyze the historical data using time series analysis techniques. In particular, intervention analysis can be used to model the historical data taking into account a past promotion. This type of promotional analysis may help determine how past promotions affected the historical sales and can help predict how proposed promotions may affect the future based on similar, past promotions. This paper briefly describes intervention analysis, provides practical advice for promotional analysis and forecasting using interventions, and demonstrates these practices using SAS/ETS ® Software.

Keywords
Promotional Analysis, Demand Planning, Forecasting, Intervention Analysis, Time Series Analysis, ARIMA

1. Introduction

Businesses need to plan how they will generate and satisfy demand for their products and/or services. Sales promotions are a vehicle by which businesses increase the demand for and visibility of their products and/or services. These promotions cost money and these costs must be justified by structured analysis. Additionally, proposed sales promotions affect future demand, so increased resources must be allocated to satisfy the promoted demand. With the advent of e-commerce, customer expectations are much higher than in the past. Therefore, if a business promotes a product or service, it is expected to provide these in a timely fashion.

Business leaders often ask the following questions: Was a past promotion successful? Will a proposed promotion be profitable? How will demand be affected by a planned promotion, which is similar to a past promotion? These are questions that this paper hopes to help answer.

2. Scope

The topic of promotional analysis means different things to different people. This paper only considers product or service promotions for which sufficient historical data exist. This paper presents a methodology for analyzing and forecasting promotions based on promotions that have occurred in the past in either the product or service under analysis, or a similar product or service whose underlying time series process exhibits similar properties in response to (similar) promotions.

This paper will not answer the questions of who should be targeted for a promotion, where should a promotion be advertised, nor other such categorical issues outside the scope of traditional time series analysis. These types of questions may be better answered using data mining techniques (Berry and Linoff 1997) such as cluster analysis or memory-based reasoning (although the time series analysis techniques described in this paper may also be useful in these analyses). Additionally, this paper will not address promotions related to products or services that have no historical data (new products). New product forecasting and promotional analysis may be better addressed using diffusion modeling (Parker 1994), surrogate product analysis (Parkoff and Crowler 1999), judgmental techniques (Wright and Goodwin 1998), and other new product analyses (Thomas 1993). Also, this paper will not address the questions of price and promotional elasticities. These analyses are better addressed by pricing models (Foekens et al.

1994), shrinkage estimation techniques (Blattberg and George 1989), and other econometric modeling (Foekens et al. 1994).

Although the scope of this paper is rather limited with respect to overall issues related to promotional analysis, the techniques described here may prove beneficial in practical analysis of the overall problem. It may well be likely that many different techniques are needed, possibly in combination.

3. A Practical Time Series Approach

In this paper, promotional analysis is a means to evaluate the success or failure of a promotion given the historical time series data. Established statistical techniques are one way to objectively evaluate the success or failure of a promotion. Judgmental techniques offer a subjective evaluation of the success or failure of a promotion. A combination of statistical and judgmental techniques may provide the best answers. Promotional analysis can be achieved using well-established intervention analysis techniques based on both the historical data and the judgment of the practitioner and other experts.

In time series analysis literature (Box et al. 1994), an intervention *event* is an input series that indicates the presence or absence of an event. An intervention event causes a time series process to deviate from its expected evolutionary pattern. It is assumed that the intervention event occurs at a specific time, has a known duration, and is of a particular type. The *time* of the intervention is when the event begins to cause deviation. The *duration* of the intervention is how long the event causes deviation. The *type* of the intervention is how the event's influence changes over time. The intervention *response* is how the intervention causes deviation. Generally speaking, intervention events are essentially dummy regressors or indicator variables (explicitly determined by the time, duration, and type) that are introduced through a transfer function filter (specified by the response) that results in an intervention *effect*. The term *intervention* sometimes applies to the event, response, and/or effect.

Many promotions could be considered interventions because they can cause the aforementioned deviations. In such cases, the practitioner usually knows the time, duration, and type of the promotion, which has occurred in the past or will occur in future. The response of the promotion is not generally known. However, the response may be identified by an analysis of the (similar) historical data and/or from the judgment of the practitioners and other experts based on their past experience, domain knowledge, and other statistical analyses outside the scope of this paper. If you know the intervention event and response, the intervention effect can be derived from the response parameter estimates. The end result of promotional analysis, using interventions, is that the practitioner has a better understanding of how past promotions affected the historical data. Armed with the knowledge gained by intervention analysis, the practitioner can better determine the value of past promotions and better forecast product and service demand by taking into account future promotions that are similar.

This paper provides practical advice to the practitioner for evaluating promotions using intervention analysis. The methodology presented involves decomposing the historical data into two parts: the underlying time series and the intervention effect. This methodology is in no way comprehensive because the complexity of the underlying time series process and/or the interventions may be quite complicated (for instance, several overlapping promotions). Additionally, the issues related to promotional analysis may extend beyond those related to time series analysis, as mentioned above. The methodology presented is illustrated using SAS/ETS Software.

4. Background

This section provides a brief theoretical background on both time series models and intervention models. It is intended to establish notation for the experienced practitioner and to provide the novice practitioner with motivation, orientation, and references. A more complete discussion of these topics can be found in Box et al. (1994), Hamilton (1994), and Fuller (1995). This paper follows the notation used in Box et al. (1994). Although this brief background does not include details of all the different types of time series models, SAS/ETS Software can be used to analyze all of the time series models mentioned in this section (and many others as well).

4.1 Time Series

A discrete-time series consists of a set of random variables, which are observed at equally spaced time periods (say daily or monthly) and which are ordered and indexed by time. Assuming that T observations are recorded, then a single realization of a univariate time series can be represented by y_t, where $t = 1,\ldots, T$.

4.2 Time Series Models

A typical time series model attempts to explain the behavior of a time series based on its lagged values using a disturbance filter and possibly one or more inputs. A simple example of a time series model is the first order autoregressive model or AR(1) model that explains the behavior of a time series based on previous values:

$$y_t = c + \phi y_{t-1} + a_t$$

where c is a constant, ϕ is the first-order autoregressive parameter, and a_t is the random disturbance term (also known as error or noise) of mean zero and constant variance σ^2. a_s where a_t are uncorrelated for all $s \neq t$ (sometimes called white noise). The series is stationary when ϕ is less than one in magnitude. For a stationary time series, past values influence the current value in an exponentially decreasing fashion.

More generally, the lagged relationships associated with a time series model can be quite complex. For example, an ARMA(p,q) model has the following form:

$$\phi(B)(y_t - \mu) = \theta(B)a_t \qquad \text{or alternatively} \qquad y_t = \mu + \psi(B)a_t$$

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p$$
$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q$$
$$\psi(B) = \theta(B)/\phi(B)$$

where μ is the series mean, B is the backshift operator, $By_t = y_{t-1}$, $\phi(B)$ is the p^{th} order autoregressive polynomial, $\theta(B)$ is the q^{th} order moving average polynomial, $\psi(B)$ is the (finite or infinite order) disturbance filter, and a_t is the random disturbance term that is assumed to be white noise. The time series is stationary and invertible if the roots of the characteristic equation of both the autoregressive and moving average polynomial are outside the unit circle, respectively. It is assumed that y_t is stationary or has been made stationary by appropriate simple or seasonal differencing.

There may be one or more deterministic and/or stochastic input (regressor, exogenous, or explanatory) variables that influence the time series, and their influence can be either static or dynamic. These inputs can be modeled as a time varying mean and can be better understood by a model of the following form:

$$y_t = \mu_t + \psi(B)a_t \qquad (4.2.1)$$

where μ_t varies with time and describes the influence of the inputs on the time series at each point in time. If the term μ_t is not influenced by lagged values of the inputs, the model is often called a *regression with time series errors model*; otherwise, the model is often called a *dynamic regression model*.

This paper will not address the general problem of a time series model with stochastic inputs (even though interventions can be viewed as deterministic inputs). If the inputs' influence on the time series is static, they are simply regressor variables; otherwise, they are dynamic regression variables or transfer function inputs. Additionally, the time series model may involve a transformation (for example: logarithmic, square root, logistic, or Box-Cox transformations). This paper will not address transformed models but SAS/ETS Software can analyze, model, and forecast transformed models.

4.3 Intervention Events

An intervention *event* is an input series from a deterministic source (dummy regressor or indicator variable) that is used to explain deviations from the underlying time series process. Interventions begin to influence the recorded data at a specific *time* and last for a specific *duration*. There are three commonly used types of intervention events. This paper uses ξ_t to denote an intervention event.

4.3.1 Point Intervention

A point (pulse) intervention is a dummy regressor that takes a value of one (1) at the time of the intervention, and the rest of its values are zero (0). The duration of a point intervention is one time period. In other words,

$\xi_t = 1$, if $t = time$
$\xi_t = 0$, otherwise

Point interventions are useful for evaluating promotions, which are known to occur in a single, specific time period (*time*) and whose influence on the time series will die out thereafter. Figure 1a shows an example of a point intervention.

4.3.2 Step Intervention

A step intervention is a dummy regressor whose values before the time of the intervention are zero and whose subsequent values are one. The duration of a step intervention is the number of periods from time to the end of the time series. In other words,

$\xi_t = 0$, if $t < time$
$\xi_t = 1$, otherwise

Step interventions are useful for evaluating promotions, which are known to occur during and after a specific time period (*time*) and which permanently influence the time series thereafter. Figure 1b shows an example of a step intervention. If the duration of the promotion ends prior to the end of the time series, two offsetting step interventions can be used. This type of intervention is often called a *temporary change*.

4.3.3 Ramp Intervention

A ramp intervention is a dummy regressor whose values before and during the time of the intervention are zero and whose subsequent values increase linearly thereafter. The duration of a ramp intervention is the number of periods from *time* to the end of the time series. In other words,

$\xi_t = 0$, if $t < time$
$\xi_t = (t - time)$, otherwise

Ramp interventions are useful for evaluating promotions, which are known to occur during and after a specific time period (*time*) and whose influence on the time series increases thereafter. Figure 1c shows an example of a ramp intervention. If the duration of the promotion ends prior to the end of the time series, two offsetting ramp interventions can be used.

Various combinations of point, step, ramp, and other types of interventions can be used to model more complex promotions. However, using excessive numbers of interventions can be dangerous. In the extreme case, using an intervention at every point in time can completely define a time series, in which case future modeling is useless.

Figure 1a - Point Intervention

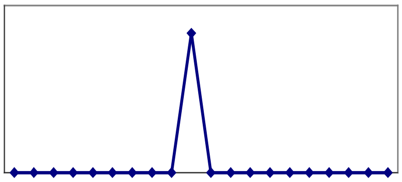


Figure 1b - Step Intervention

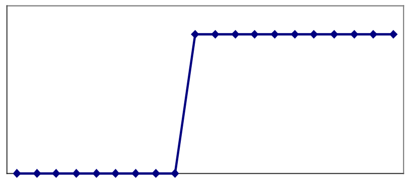
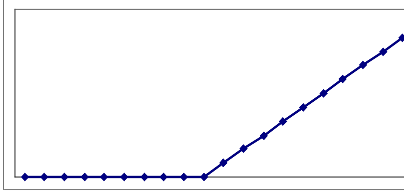


Figure 1c - Ramp Intervention



Fiigure 2a - Point Intervention with Exp(0.6) Filter

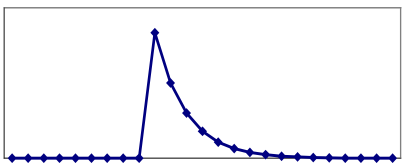


Figure 2b - Step Intervention with Exp(0.6) Filter

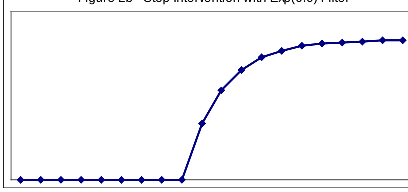


Figure 2c - Ramp Intervention with Exp(-0.95) Filter

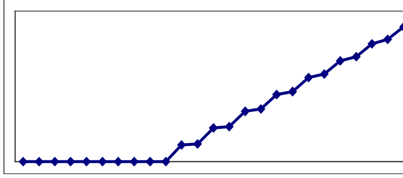


Figure 3a - Point Intervention with Wave(0.7,0.25) Filter

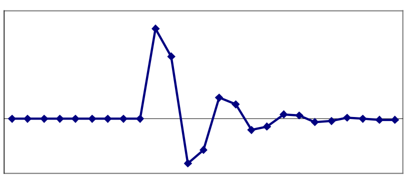


Figure 3b - Step Intervention with Wave(0.7,0.25) Filter

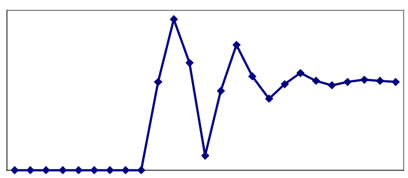
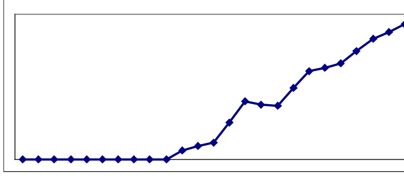


Figure 3c-Ramp Intervention with Wave(0.7,0.25) Filter



4.4 Interventions Response (Transfer Function Filters)

Interventions are introduced in a time series model through a transfer function filter. A *transfer function filter* explains how the current and previous (lagged) values of the intervention event cause deviations in the underlying time series process. A transfer function filter is similar to the disturbance filter; however, the *disturbance filter* operates on the disturbances where as the transfer function filter operates on the intervention. In general, a typical transfer function model has the form:

$$\upsilon(B) = \omega_0\,\omega(B)B^b/\delta(B)$$

$$\omega(B) = 1 - \omega_1 B - \omega_2 B^2 - \ldots - \omega_s B^s$$
$$\delta(B) = 1 - \delta_1 B - \delta_2 B^2 - \ldots - \delta_r B^r$$

where $\upsilon(B)$ is the (finite or infinite order) transfer function filter, ω_0 is the scaling factor, $\omega(B)$ is the s^{th} order numerator polynomial, $\delta(B)$ is the r^{th} order denominator polynomial, and b accounts for lagged effects. If $r = 0$, $\upsilon(B)$ is of finite order; otherwise, it is of infinite order.

If each intervention value influences both the current and past values, then the intervention is a dynamic regression variable. In the case where the input series is an intervention, the transfer function filter, $\upsilon(B)$, is also referred to as the intervention response. The overall influence of the intervention event on the underlying time series is subsequently referred to as the intervention effect ($\upsilon(B)\xi_t$) which describes the promotion's influence over time.

If each intervention value only influences the current value of the underlying time series, that is $\omega(B) = 1$ and $\delta(B) = 1$, then the intervention is a simple regression variable with a parameter, $\upsilon(B) = \omega_0$. In this case, the intervention response is constant for all time and the intervention effect ($\omega_0\xi_t$) is a scaled version of the intervention event shown in Figure 1.

Transfer function filters include *numerator* (similar to autoregressive) polynomial terms and *denominator* (similar to moving average) polynomial terms. Commonly used denominator polynomial terms include one parameter exponential decay filters (Exp, $s = 0$, $r=1$) and two parameter filters (Wave, $s = 0$, $r = 2$):

Exp $\quad \upsilon(B) = \omega_0 / (1 - \delta_1 B)$
Wave $\quad \upsilon(B) = \omega_0 / (1 - \delta_1 B - \delta_2 B^2)$

A full discussion of transfer functions is beyond the scope of this paper. However, to illustrate the above concepts, Figures 2 and 3 illustrate the intervention effect of these two common transfer function filters (response) on the basic types of interventions previously described.

4.5 Combined Model

Given an underlying time series that has deviated from its expected evolutionary path by a past promotion, this paper's goal is to aid the practitioner in developing a model that explains this deviation. One way is to combine the time series model with the intervention model. The resulting combined model may take the following form (compare with equation 4.2.1):

$$y_t = \mu_t + \psi(B)a_t + \upsilon(B)\xi_t$$

The first right-hand term models the mean term (possibly time varying), the second term models the disturbance, and the third term models the intervention (or promotion). If $\mu_t = \mu$ and $\upsilon(B) = 1$, we are dealing with a regression with time series error model; otherwise, dynamic regression or transfer function modeling is required. In the case of the regression with time series error model, the above equation has the form:

$$y_t = \mu + \psi(B)a_t + \omega_0\xi_t$$

If simple or seasonal differencing is applied to the time series, y_t, the same differencing should be applied to the intervention events, ξ_t.

$$(1-B)y_t = \mu_t + \psi(B)a_t + \upsilon(B)(1-B)\xi_t$$
$$(1-B)(1-B^s)y_t = \mu_t + \psi(B)a_t + \upsilon(B)(1-B)(1-B^s)\xi_t$$

Subsequently, we will refer to the time series model and transfer function filter for the intervention as the *combined model* (not to be confused with the technique of combining one or more forecasting models).

4.6 Multiple Promotions

A product or service may have had several overlapping promotions that occurred in the past, or past promotions that influence proposed future promotions. SAS/ETS Software can be used to analyze multiple promotions, but this paper does not specifically address this problem.

5. Promotional Analysis – A Practical Time Series Approach Defined

In order to perform promotional analysis using intervention analysis, the underlying time series model, as well as the promotion/intervention model, must be specified. Once the combined model has been specified, the parameters can be estimated by fitting the historical data. Established transfer function estimation techniques can accomplish this task such as described in Box et al. (1994). Once the parameters have been estimated, the effect of the promotion/intervention can be determined, which is the ultimate goal of promotional analysis using time series techniques. Once the effect of a past promotion has been determined, its value can be computed, as well as forecasts of similar proposed promotions.

In equation form, the above task involves decomposing historical data, y_t, into two parts: the underlying time series process, $\mu_t + \psi(B)a_t$, and the effect, $\upsilon(B)\xi_t$. Once the response parameters, $\upsilon(B)$, have been estimated, the effect, $\upsilon(B)\xi_t$, may be derived because the intervention event, ξ_t, is deterministic. The derived effect aids the practitioner in understanding how the intervention caused the underlying time series process to deviate from its expected path. This decomposition is illustrated in Figures 4 and 5.

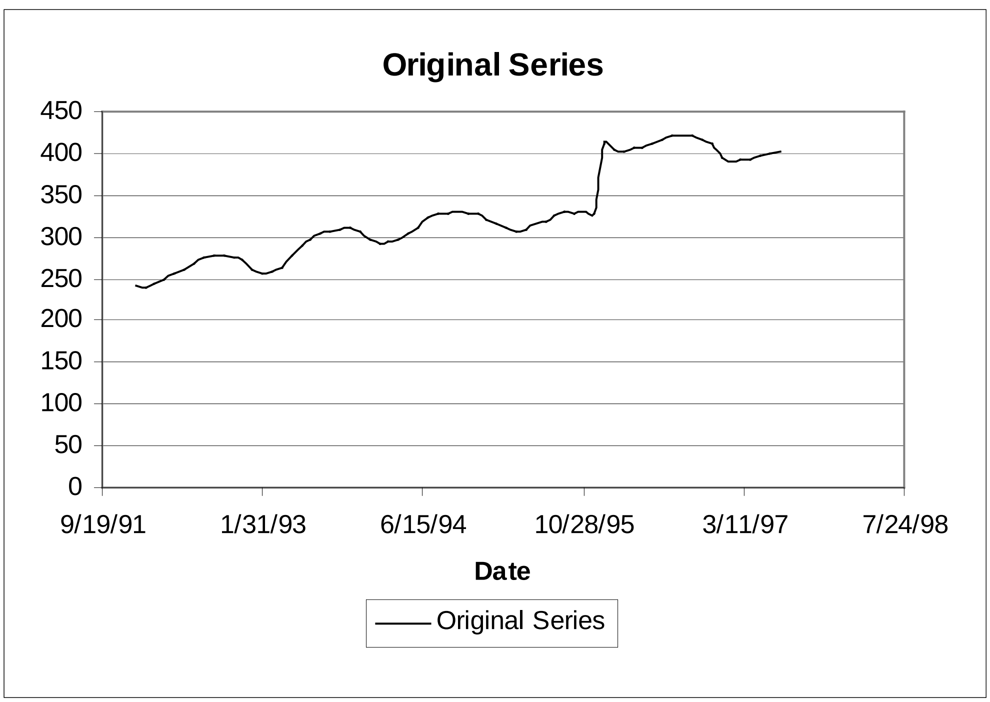


Figure 4: Original Time Series

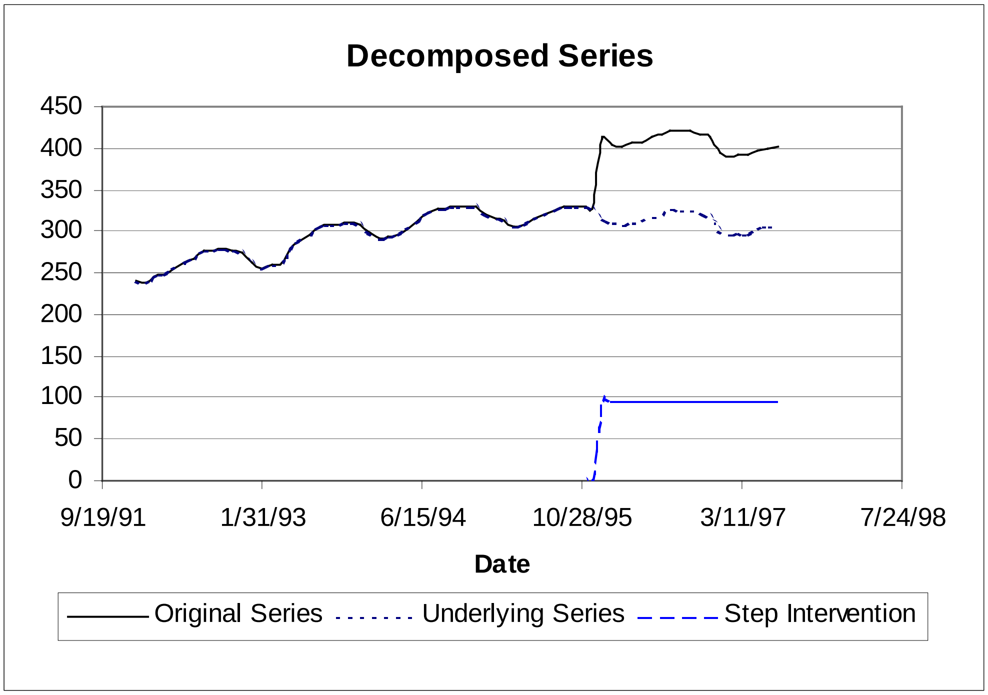


Figure 5: Decomposed Time Series

In practice, the practitioner knows the time, duration, and type of the intervention event. The underlying time series model specification and/or the intervention response specification may or may not be known. Since computer software can estimate the parameters, determining the model specification and the intervention response specification are the most important tasks for the practitioner and are crucial to effective promotional analysis. This step is often referred to as *identification*.

Once the combined model is successfully identified, fitted to the historical data, and checked for adequacy, the response is estimated. Combining the estimated response with the known intervention event (time, duration, and type), the deviation caused by the promotion is also determined because the effect can be fully derived from the response parameter estimates. Since response parameter estimates are random variables, they have an associated variance; therefore, the derived effect should have associated confidence limits. If the intervention model is dynamic, these confidence limits can be difficult to compute (Monte Carlo techniques are needed). This paper will not address this issue and leaves it for the reader to explore.

6. Combined Model Identification, Estimation, and Checking

If the underlying time series model specification is not known, it can be determined by analyzing the historical data. The statistical literature on model identification is immense and requires sophistication and experience to master. Some well-established techniques for model identification include:

1. Stationarity analysis (nonseasonal and seasonal)
2. Autocorrelation/partial autocorrelation/inverse autocorrelation analysis
3. Smallest canonical correlation analysis
4. Extended sample autocorrelation analysis
5. Minimum information criteria
6. And many others (Choi 1992)

This paper will not survey these techniques but leaves them for the reader to explore. However, since a past promotion can cause the time series to deviate from its expected evolutionary path as predicted by the underlying time series process, these model identification techniques may be less effective. If the effect of the promotion/intervention is small, these identification techniques may be effective. If the duration of the promotion/intervention is small with respect to time range, the historical data associated with the promotion can be set to missing values and the model can be more readily identified from the modified data. If there is insufficient data to eliminate the promotion/intervention, the model and intervention response must be identified jointly using the combined model.

Be cautious when using a large number of disturbance filter (moving average and/or autoregressive) parameters because they can mask the intervention effect. When selecting between candidate models, a selection criterion that penalizes for large numbers of parameters should be used, such as the Akaike Information Criterion (AIC), Schwarz Bayesian Criterion (SBC or BIC), or Amemiya's Prediction Criterion (APC).

If the intervention response specification is not known, it can also be determined by analyzing the historical data. If the underlying time series model is specified or effectively identified, an analysis of the model residuals, similar to the analysis of the historical data, may help identify the response; sometimes several iterations are needed. Once again if the model and response are unspecified, they must be jointly identified using the combined model. And likewise, a large number of transfer function (numerator and/or denominator polynomial) parameters can mask details associated with the underlying time series process.

As can be seen, there is a balancing act between the number of disturbance filter and transfer function filter parameters. The ultimate goal is to specify the disturbance filter so as to capture the underlying time series process without masking the effect of the intervention and to specify transfer function filter so as to capture the intervention effect without masking the underlying time series process. This balancing act may require several iterations of combined model identification, estimation, and checking to resolve.

Once both the model and the response have been specified, the combined model should be estimated and checked for validity. Most practitioners will rely on the computer software to do the estimation. Once again, the statistical literature on model checking is immense and requires sophistication and experience to master. In addition to the techniques for model identification (applied to the combined model residuals) described above, well-established techniques for model checking include:

1. White Noise Testing of combined model residuals
2. Significance of the parameter estimates
3. And many others

Often, more than one iteration of identification, estimation, and checking is needed (Box et al. 1994). The above discussion begs the questions: What if the specification of the model and response are unknown? What if some of the above identification techniques are not fully understood by the practitioner? This paper hopes to offer some practical advice to help answer these questions.

7. Combined Model Identification, Fitting, Checking - Practical Advice

For the practitioner, estimating the combined model to determine the effect of the promotion is the desired goal. This goal requires identification, estimation, and checking of the combined model (possibly requiring several iterations). The methodology presented uses a model selection criterion to select promising models from a larger set of candidate models. The model selection criterion should penalize for large numbers of parameters (such as AIC, BIC, APC). What is known about the combined model specification determines how to proceed. The methodology below is ordered by decreasing knowledge of the combined model specification, and the advice provided in each section applies to subsequent sections.

7.1 Specified Time Series Model, Specified Intervention Response

Given that the model and response are both specified, all that is needed is parameter estimation and checking. The computer software usually does parameter estimation and the practitioner must then check the specified combined model for adequacy using techniques described above and utilizing the judgment of the practitioner and/or domain experts. If the checking indicates that the combined model is inadequate, changes to the specifications should be considered and the new combined model re-fit. Once the combined model is deemed adequate, it is recommended that the historical data be fit to just the model with the intervention excluded. Comparing the two final candidate models (with and without the intervention) will help determine if the intervention has great influence.

7.2 Specified Time Series Model, Unspecified Intervention Response

Given that the model is specified but the response is unspecified, several candidate transfer functions should be specified utilizing the judgment of the practitioner and/or domain experts. All of the candidate transfer functions should be fit to the historical data using the specified model. The candidate transfer functions should be ordered by the model selection criterion. Starting with the first, check the adequacy of the candidate combined model similar to above. The first candidate that is determined adequate is likely a good candidate, but checking other candidates that performed well is recommended.

7.3 Unspecified Time Series Model, Specified Intervention Response

Given that the model is unspecified but the response is specified, several candidate models should be specified utilizing the judgment of the practitioner and/or domain experts. All of the candidate models should be fit to the historical data using the specified response. The candidate models should be ordered by the selection criterion. Starting with the first, check the adequacy of the candidate combined model similar to the above. The first candidate that is determined adequate is likely a good candidate but checking other candidates that performed well is recommended.

7.4 Unspecified Time Series Model, Unspecified Intervention Response

Given that the model and response are both unspecified, several candidate models should be specified, as well as several candidate transfer functions. Each possible combination of candidate model and candidate transfer functions should be fit to the historical data. The candidate combined models should be ordered by the model selection criterion. Starting with the first, check the adequacy of the candidate combined model similar to above.

8. Promotional Analysis

Once the combined model is successfully identified, fitted to the historical data, and checked for adequacy, the intervention response is fully estimated. Since the estimated response is based on the transfer function parameter estimates and the intervention event (time, duration, and type) is fully described by a dummy variable, the intervention effect (departure from the underlying time series process) caused by the promotion can be derived.

Knowing the intervention event, response, and effect of a promotion is the ultimate goal of promotional analysis (using a practical time series approach) because once these properties are known or estimated the deviation from the underlying time series process has been determined. The value of the promotion can also be determined from this deviation, usually after appropriate scaling.

9. Promotional Analysis - Practical Advice

The way in which deviations related to promotions should be evaluated can be quite complex (especially if issues outside the scope of time series analysis are considered). Once again, this paper offers simple practical advice on evaluating these deviations.

The *deviation* caused by the promotion can be determined by applying the previously estimated response to the dummy variables. To determine the overall value (benefit) of a promotion, simply integrate (or sum) the deviation from the intervention time until the present. Even though the duration of the intervention may be limited, the effect of the intervention may last longer due to the transfer function's memory. If the transfer function has no memory, then integrating throughout the duration will be all that is necessary. However, in either case, integrating from the *time* to the present will be adequate. If the time series records units, the units need to be converted to revenue.

When a transformed model is used, the parameter estimates relate to the transformed metric. For example, if a log transformation was used, the transfer function is additive in the transformed metric and multiplicative in the original metric. As stated above, this paper will not consider transformed models in detail.

10. Forecasting with Proposed Promotions

The forecasts generated by the model or combined model based solely on the historical data should be *adjusted* to account for the deviation anticipated with a proposed future promotion. After a past promotion is fully analyzed using the methodology provided above, the deviation caused by the past promotion has been derived. If a similar promotion is proposed in the future, this deviation might provide insights into the future. Specifically, the deviation, shifted in time, can be used to adjust the forecasts generated solely from the historical data.

If the proposed promotion is sufficiently distant and/or the season differs from the past, similar promotion, the deviation should possibly be re-scaled for a change in level or trend, and/or seasonally adjusted if the change of seasons is important. Additionally, statistical and judgmental analysis outside the scope of this paper may also be used in adjusting forecasts to account for the proposed promotion.

If the underlying time series process is trending upward or downward and if the level influences the deviation caused by the promotion, the forecast should be adjusted based on the scaled deviation associated with the future level. In addition, if a deviation was derived from a past promotion of a different, yet similar product or service, it should be scaled to account for the different levels between the products or services prior to adjusting the forecast. In particular, statistical analyses related to price and promotional elasticities and data mining techniques can help determine how to scale the deviation.

Likewise, if the underlying time series is seasonal and if the deviation caused by the promotion is influenced by the season, the season in which the past and proposed promotion occur needs to be considered. If the past and the proposed promotions occur in different seasons (say winter versus summer), the deviation related to the past promotion should be changed to account for different seasons prior to adjusting the forecasts.

Once again, since the transfer function parameters are random variables and the level, trend, and seasonal adjustments introduce additional uncertainty, the derived deviation should have associated confidence limits. Therefore, the forecast confidence limits should incorporate this additional uncertainty.

11. Forecasting with Proposed Promotions - Practical Advice

The way in which past deviations should be re-scaled or changed for adjusting the forecasts based on the historical data can be quite complicated and requires the judgment of the forecasting practitioner and/or domain experts. Once again, this paper offers practical advice on adjusting forecasts to account for proposed promotions.

11.1 Same Time Series, Similar Past Promotion

Based on the historical data, forecast the time series using the specified combined model without considering the proposed future promotion. Use the projected level and/or seasonal index to adjust the deviation for the level change and/or change in season. Adjust the forecast based on the historical data and on the time-shifted deviation, accordingly.

11.2 Different Time Series, Similar Past Promotion

Based on the historical data, forecast the time series using the specified model without the proposed future promotion. Use the projected level and/or seasonal index to adjust deviation for the level change and/or change in season between the two products or services. Adjust the forecast based on the historical data and on the time-shifted deviation, accordingly.

11.3 Adjusting Confidence Limits (Prediction Intervals)

Since the future deviations are not known with certainty, the adjusted forecasts' confidence limits should also be adjusted to account for this additional uncertainty. A simple way to proceed is as follows: adjust the forecasts based on the derived deviation; adjust the forecasts' lower confidence limit based on the derived deviation's lower confidence limit; and finally, adjust the forecasts' upper confidence limit based on the derived deviation's upper confidence limit. Although this approach does not fully address all possible uncertainty, it is a simple and easy approach that may provide insight to the practitioner.

11.4 Updating Forecasts

As time passes and more data about an ongoing or proposed promotion becomes available, the combined model should be re-estimated and checked. The forecasts should then be updated based on the new data.

12. Demonstration Using SAS/ETS ® Software

This paper investigated several scenarios that may be encountered in promotional analysis and forecasting by the practitioner; however, below we provide just a few of these examples using the SAS ® System. In general, the advice provided above is illustrated below.

12.1 Unspecified Time Series Model, Specified Intervention Response

Consider the problem of evaluating a promotion, which starts in January 1996 and continues to the present, say July 1997. Figure 6 provides plot of the time series under consideration.

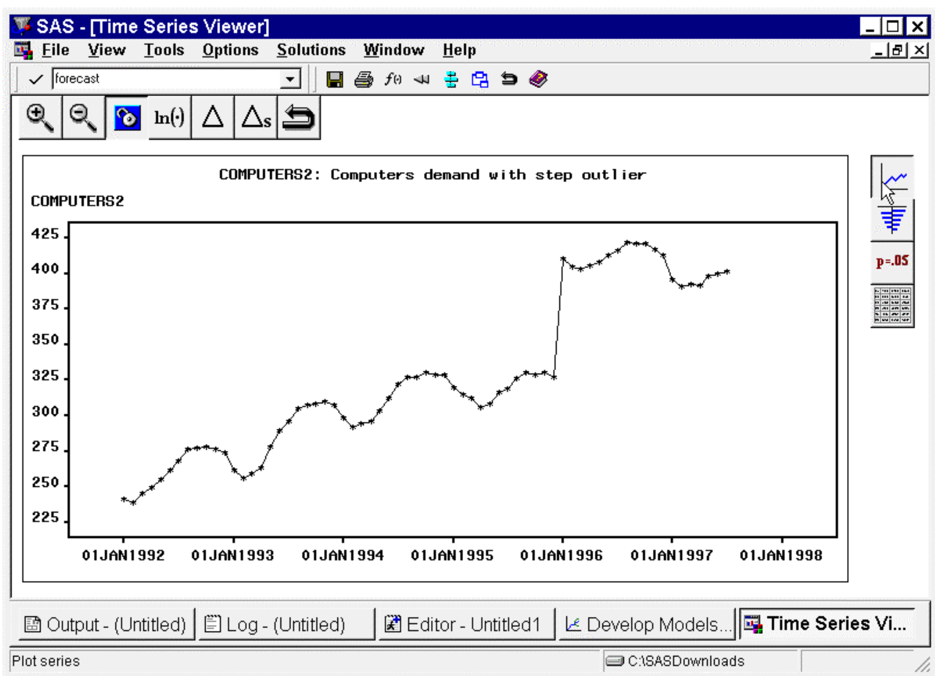


Figure 6: Time Series Plot

Assume that the underlying time series contains no dynamic inputs. Assume that the promotion deviates the underlying time series process in a consistent, step-wise fashion from the starting date until the present. The *time*, *duration*, *type*, and *response* of the intervention are known, but the underlying time series model is unknown. Based on these assumptions, the combined model is assumed to the following form (regression with time series error model):

$$y_t = \mu + \psi(B)a_t + \omega_0\xi_t$$

If the series is determined to be non-stationary, the both the series and the intervention event should be difference accordingly.

If simple differencing is required, the combined model is assumed to have the following form:

$$(1-B)y_t = \mu + \psi(B)a_t + \omega_0(1-B)\xi_t$$

If seasonal differencing is required, the combined model is assumed to have the following form:

$$(1-B^s)y_t = \mu + \psi(B)a_t + \omega_0(1-B^s)\xi_t$$

If simple and seasonal differencing is required, the combined model is assumed to have the following form:

$$(1-B)(1-B^s)y_t = \mu + \psi(B)a_t + \omega_0(1-B)(1-B^s)\xi_t$$

Exhibit 1 provides a step-by-step methodology for promotional analysis using the SAS/ETS ® Time Series Forecasting System. It is not a complete methodology for all promotional analysis problems but offers some general insight. The results of this analysis are shown in Figure 7.

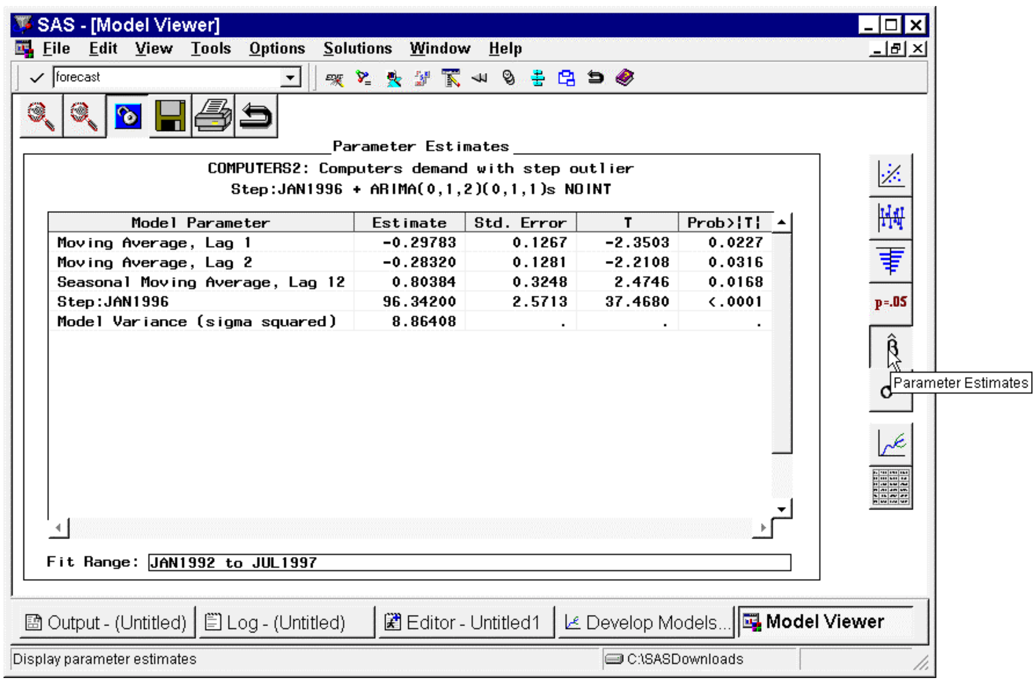


Figure 7: Combined Model Parameter Estimates

The ARIMA (0,1,2)(0,1,1)s model was selected from the candidate models based on a model selection criterion (AIC). Since the parameter estimate for the step intervention (STEP:JAN1996) was 96.342 with standard error 2.5713, then a good guess is that the value of this promotion is 96.342 plus or minus two standard errors for each period of the duration of the intervention.

12.2 Different Time Series, Similar Past Promotion

Consider the problem of forecasting for a proposed promotion that will start in December 1997 and continues beyond the forecast horizon. Figure 8 provides a plot of the time series.

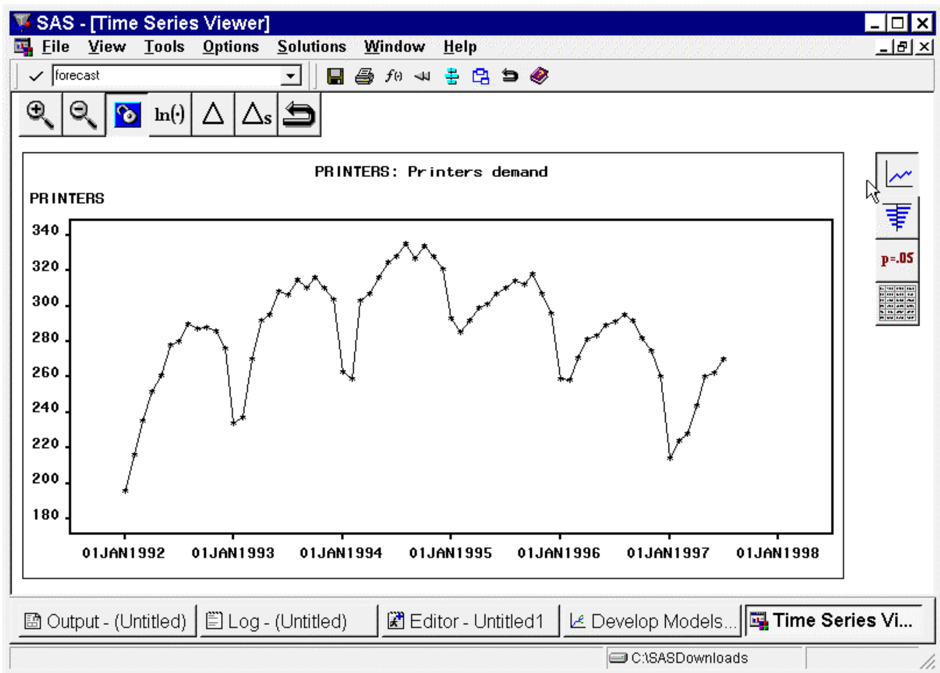


Figure 8: Time Series Plot

Assume that the underlying time series model does not contains dynamic inputs and there have been no past promotions. Based on these assumptions, the combined model is assumed to the following form (ARIMA model):

$$y_t = \mu + \psi(B)a_t$$

Assume that the promotion will deviate the underlying time series process in a step-wise fashion from the proposed starting date until the forecast horizon. Additionally, assume that the adjustment has been scaled and seasonally adjusted. In this demonstration, the adjustment is 106 units starting at December 1997. The *time*, *duration*, *type*, and *response* of the adjustment are based on a past (similar) promotion, but the underlying time series model is unknown.

Exhibit 2 provides a step-by-step methodology for forecasting proposed promotions using the SAS/ETS ® Time Series Forecasting System. It is not a complete methodology for all promotional forecasting problems but offers some general insight. The results of this analysis are shown in Figure 9.

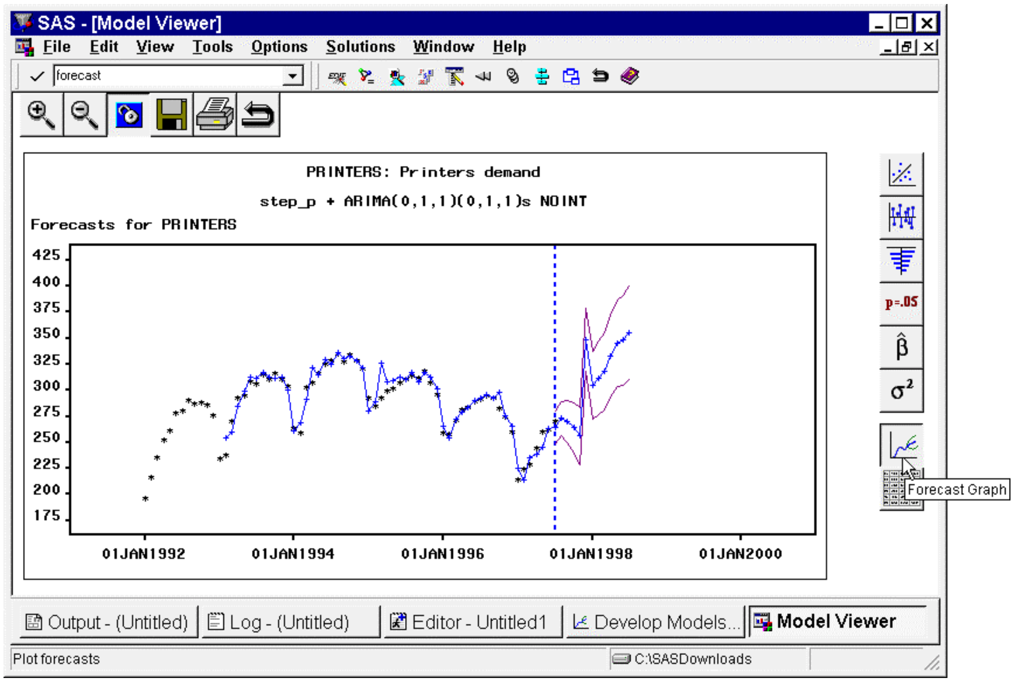


Figure 9: Forecast adjusted for Proposed Promotion

The ARIMA (0,1,2)(0,1,1)s model was selected from the candidate models based on a selection criterion (MAPE); however, further analysis suggested that an ARIMA (0,1,1)(0,1,1)s model was more desirable. The forecasts were generated using this model. The forecasts were then adjusted by the proposed deviation (106 units).

13. Conclusion

Many businesses use sales promotions to increase the demand or visibility of a product or service. Structured analysis of the historical data can aid in determining the value of a past promotion and provide insight into forecasting the value of future promotions. Though this structured analysis can have many forms, a practical time series approach, demonstrated by this paper, that utilizes human judgment might be beneficial to the practitioner for demand planning. Decomposing the historical data into two parts (the underlying time series process and the intervention/promotion effect) is one approach that may prove useful to the practitioner in analyzing past promotions. Forecasts can be adjusted based on proposed promotions using the analysis of past (similar) promotions incorporating insight provided by statistical and judgmental techniques outside the scope of this paper. This practical approach was demonstrated using SAS/ETS ® Software.

14. Acknowledgments

This paper and its associated practical demonstration relied heavily on the contributions of Evan Anderson, Virginia Clark, Gul Ege, Donald Erdman, and Rajesh Selukar of SAS Institute Inc.

15. References

Berry, M. J. A. and Linoff G. (1997), *Data Mining Techniques*, New York: John Wiley & Sons, Inc.

Blattberg, R. C. and George, E. I. (1989), "Shrinkage Estimation of Price and Promotional Elasticities," *Proceedings of the Business and Economic Statistics Section*.

Box, G. E. P, Jenkins, G. M., and Reinsel, G.C. (1994), *Time Series Analysis: Forecasting and Control*, Englewood Cliffs, NJ: Prentice Hall, Inc.

Choi, B. (1992), *ARMA Model Identification*, New York: Springer-Verlag, Inc.

Foeckens, E. W. , Leeflang, P. S. H., and Wittink, D. R. (1994), "A Comparison and an Exploration of the Forecast Accuracy of a Loglinear Model at Different Levels of Aggregation," *International Journal of Forecasting*.

Foeckens, E. W., Leeflang, P. S. H., and Wittink, D. R. (1999), "Varying Parameter Models to Accommodate Dynamic Promotion Effects," *International Journal of Forecasting.*

Fuller, W. A. (1995), *Introduction to Statistical Time Series*, New York: John Wiley & Sons, Inc.

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

Parker, P. M. (1994), "Aggregate Diffusion Forecasting Models in Marketing," *International Journal of Forecasting*.

Parkoff, S. B. and Crowther, J. E. (1999), "New Product Forecasting in Fast-Moving Consumer Products Companies," Presentation at the International Symposium on Forecasting.

Thomas, R. J. (1993), *New Product Development: Managing and Forecasting for Strategic Success*, New York: John Wiley & Sons, Inc.

SAS Institute Inc. (1993), *SAS/ETS User's Guide, Version 6, Second Edition*, Cary, NC: SAS Institute Inc.

Wright, G. and Goodwin, P. (1998), *Forecasting with Judgment*, New York: John Wiley & Sons, Inc.

16. Exhibits

Fullscreen details of the demonstrations are too large to include in this paper. This paper's exhibits can be found at http://www.sas.com/statistics. Once at this Web site, select *Papers*.