

# Should Music Labels Pay for Radio Airplay? Investigating the Relationship Between Album Sales and Radio Airplay

by

Alan L. Montgomery and Wendy W. Moe

*August 2002*

The authors wish to thank Pete Fader and Capitol Records for the data used in this research. Alan L. Montgomery is an Associate Professor of Industrial Administration at Carnegie Mellon University, GSIA, 5000 Forbes Ave., Pittsburgh, PA 15213 (e-mail: alan.montgomery@cmu.edu) and Wendy W. Moe is an Assistant Professor of Marketing at the University of Texas at Austin, McCombs Schools of Business, Austin, TX 78712-1178 (e-mail: wendy.moe@bus.utexas.edu).

# Should Music Labels Pay for Radio Airplay? Investigating the Relationship Between Album Sales and Radio Airplay

## **Abstract:**

Managers in the music industry closely monitor both radio airplay of an album as well as the album's sales. Their interest in radio airplay is due to the belief that airplay can increase an album's sales. Therefore it is natural for managers to attempt to influence radio airplay so as to subsequently impact album sales and ultimately profits. Over the past several years the concept of "pay-for-play" has resurfaced. If direct payments for radio airplay are to be made, then a precise understanding of the dynamic relationship between sales and airplay is needed. Typically radio airplay and album sales both show an exponential declining pattern. It is natural to ask whether both series are evolving concurrently—but independently—or is there some type of dependence? If there is a causal relationship, what is the direction of causality, or is there be a feedback relationship where both series influence each other? The purpose of this paper is to address these modeling questions using vector autoregressive models (VARMA), and show how these models can be used to answer the substantive question of whether the music industry should pay for airplay.

**Keywords:** Music Sales; Vector Autoregressive-Moving Average Models; Estimation; Model Specification; Forecasting; Time Series Outliers and Innovations

## 1 Introduction

Managers in the music industry closely monitor both radio airplay of an album as well as the album's sales. Their interest in radio airplay is due to the belief that airplay can increase an album's sales. Therefore it is natural for managers to attempt to influence radio airplay so as to subsequently impact album sales and ultimately profits. Over the past several years the concept of "pay-for-play" has resurfaced (Billboard 1999, WSJ 1998). During the 1950 and 1960's it was common for music labels (or record companies) to offer side payments to radio station employees to play songs on the air, which was often referred to as "payola". However, this practice died down due to federal legislation.

The new "pay-for-play" proposals are not covert payments but the direct purchase of radio airplay by music labels similar to television infomercials. For example, one proposal is that a music label, say Capitol Records, pay \$1 million to buy an hour of prime time programming to schedule as they wish. Federal regulators are willing to let these types of arrangements proceed as long as the listener hears a disclaimer such as "presented by Capitol Records". Additionally, there is a growing practice of music labels paying for access to radio programmers or independent promoters that may indirectly influence radio airplay (WSJ 2002). Although some music industry representatives object that these new "pay-for-play" schemes amount to "payola" (NYT 2002). Many music labels are now worried that it is becoming too expensive to promote albums (WSJ 2002) and the issue of "payola" is being hotly debated by industry participants (Billboard 2002).

At the heart of an informed debate over payment of radio airplay is the need for a precise understanding of the dynamic relationship between music album sales and radio airplay. This is the focus of our research. Music album sales and radio airplay can be characterized as a bivariate time series which we propose to model using a vector autoregressive moving average (VARMA) model. This model allows us to answer the question of the value of radio airplay in terms of increased album sales, as well as to understand the effect of album sales on radio airplay. In our research we do not advocate "pay-for-play" for the radio industry, but instead answer the more pragmatic question of if it is permitted what is its value? Our findings also should enlighten the debate within the music industry about the efficacy of radio airplay as a promotional vehicle.

Typically radio airplay and album sales both show an exponential declining pattern. It is natural to ask whether both series are evolving concurrently—but independently—or is there some type of dependence? If there is a causal relationship, what is the direction of causality, or is there be a feedback relationship where both series influence each other?

There are many reasons to expect a dynamic relationship exists between album sales and radio airplay and that direct payment of radio airplay might be profitable for music labels. The most obvious reason to believe that airplay positively affects sales is that airplay may be perceived as a form of advertising. Alternatively, the role of radio airplay in the purchase process may be more akin to a news report from a mass media outlet than a simple advertisement. Repeated playing of an album can positively impact the public's perception of what type of music is enjoyable, just as a news report may lead listeners to believe an announcement is important. Certainly radio stations have a big impact on what is heard, since stations filter through more than 500 new releases each week and choose just three or four to introduce to the listening public. These types of effects are known in the communications literature as an agenda-setting effect (Funkhauser 1973, Iyengar and Kinder 1987, McCombs and Shaw 1972) and would predict more radio airplay results in increased album sales.

Not only is a radio station able to influence the public, but the public can also affect what is aired on the radio. Increased album sales may lead radio stations to play an album more. In the music business, this feedback relationship of album sales on radio airplay has been institutionalized in the industry via the Billboard Charts which ranks songs and albums on a weekly basis according to sales. Many stations play only songs that rank in the top 40 of these charts, while other stations use them as a guide to select songs to air that reflect their audience's preferences. Since the Billboard Charts are published on a weekly basis, these rankings can take time to alter a radio's programming, indicating lagged relationships between sales and airplay. For a study of the movement of songs on the Billboard Charts see Bradlow and Fader (2001).

A possible negative impact of increased radio airplay is that it may be a substitute for owning an album, hence a consumer may not purchase an album as long as it is being played frequently on the radio. Additionally, if an album is aired too much, then the public may tire of it quickly. Therefore, the positive aspects of radio

airplay will not go unchecked.

This discussion suggests that we cannot a priori determine the potential relationship between radio airplay and music album sales and need a general and flexible empirical approach to studying this multivariate time series. We choose the class of vector autoregressive moving average models (VARMA) after appropriately transforming the underlying series logarithmically and introducing holiday intervention effects. VARMA models can account for independence, contemporaneous dependence, and lagged dependence in a parsimonious manner. A primary advantage of VARMA models is that they do not require the analyst to know the dynamic structure beforehand, but can identify it from the data itself. Additionally, distributed lag, transfer function, univariate time series, and multivariate regression models are all nested within the class of VARMA models. Furthermore we derive conditions under which the forecasts of a VARMA model will be similar to those of the Moe and Fader's (2002) adaptation of a multivariate version of the Fourt and Woodlock (1960) diffusion model. The Fourt and Woodlock (1960) diffusion model is a special case of the popular Bass model (1969).

While there are a plethora of techniques to forecast sales either using past observations or marketing mix inputs, VARMA models have received limited attention in the marketing literature. (No applications of VARMA models have appeared within the *Journal of Marketing Research* or *Marketing Science*.) Although a subclass of VARMA models known as vector autoregressive (VAR) models have been used quite successfully to study the persistence of marketing variables, see work by Dekimpe and Hanssens (1995a, 1995b, 1999), Dekimpe, Hanssens, and Silva-Risso (1999), and Dekimpe et al (1997). In the statistical literature VARMA models have been well studied (Tiao and Box 1981, Reinsel 1995). (For an example of VARMA models applied to forecasting the unemployment rate see Montgomery et al. 1998.) An advantage of VARMA models over VAR models is that they can more parsimoniously capture longer memory effects.

Another problem that requires the analyst to exercise some caution is that it is possible that album sales and radio airplay are derived from independent diffusion processes. If both series diffuse independently of one another, but at the same time, a regression of one on the other may identify a spurious relationship between the two (Granger and Newbold (1986), §6.4). Bass and Srinivasan (2002) have also noted the that some forms of

the Bass model are quite susceptible to spurious regression when the covariates are smoothly trending. In this paper we illustrate the problems that spurious correlations can cause using the multivariate diffusion model employed by Moe and Fader (2001). We demonstrate that estimates of this model are biased when album sales and radio airplay follow independent, but similar, diffusion processes. This bias may lead a researcher to incorrectly conclude that additional airplay would increase album sales, when in fact there is no gain from additional airplay. We note that the objective of their study was to segment albums based on sales patterns and not to understand the relationship between airplay and sales.

§2 discusses how VARMA models can be employed to measure the dynamic relationship between album sales and radio airplay. In §3 we relate VARMA models to diffusion models, specifically we show under certain circumstances that diffusion models are a special case of ARMA models. We discuss the advantages of VARMA modeling when the structure of the interdependency of the series is not known beforehand. We then continue the analysis of §2 and derive the implications of the VARMA model on the valuation of radio airplay in terms of sales for a single album in §4. §5 discusses these results for thirteen different albums released by Capitol Records between 1993 and 1995. We find that the relationships between airplay and sales can vary widely. Some albums have sales that show a distinct feedback relationship with radio airplay, while other show only weak contemporaneous relationships. We show that the diversity of relationships that exist, result in much different valuations for radio airplay, some show no additional profits from radio airplay, while for others the profitability depends upon the lagged relationships. §6 concludes with a summary and discussion of these results. In summary, we find that while radio airplay can be profitable, the profitability depends heavily upon dynamic and lagged relationships.

## **2 Modeling the Relationship between Album Sales and Radio Airplay**

To answer the question that has motivated this research, “Should music labels pay for radio airplay?”, we need to determine the value of radio airplay in terms of profitability of increased album sales. There are a couple of industry practices that help simplify this problem. First, the costs of producing an album are large and

fixed, while the incremental cost of physically creating an additional album are small and relatively constant. Second, the price of music albums is typically held constant<sup>1</sup>. These practices mean that incremental profits are proportional to incremental sales. However, we need to be concerned with not only the effect of additional radio airplay on the current period of album sales, but also the discounted stream of incremental album sales associated with any additional radio airplay in the current period.

Formally, we calculate the value of radio airplay on music album sales by computing the discounted stream of incremental expected album sales due to increased airplay at time  $T$ . The stream of expected sales is used, since changes in airplay in one time period may affect future sales. Formally, this measure is:

$$\sum_{t=T}^{\infty} e^{-r(t-T)} (\mathbb{E}[s_t | \mathbf{a}_t + a_T^* \mathbf{1}_t, \mathbf{s}_{t-1}] - \mathbb{E}[s_t | \mathbf{a}_t, \mathbf{s}_{t-1}]) \quad (1)$$

where  $s_t$  and  $a_t$  are album sales and radio airplay at time  $t$ ,  $\mathbf{s}'_t = (s_1, s_2, \dots, s_t)$ ,  $\mathbf{a}'_t = (a_1, a_2, \dots, a_t)$ ,  $a_T^*$  the incremental radio airplay purchased at time  $T$ ,  $\mathbf{1}_t$  is a  $t \times 1$  vector of zeros with the  $T$ th element equal to one, and  $r$  is the discount rate of money. The key element of this calculation is the computation of the conditional expectation of album sales given past sales and radio airplay. This approach can be easily generalized to changed radio airplay schedules in multiple periods..

In the remainder of this section, we consider the calculation of this conditional expectation using a time series framework and illustrate this methodology with an empirical application to Bonnie Raitt's *Longing in Their Hearts*. The sales data is collected weekly by Soundscan, Inc. Soundscan pays for the exclusive rights to the point of purchase data obtained from over 14,000 retail stores representing more than 85% of all over the counter album sales in the United States. One weakness of the data is that small music stores are under-represented in the sample, which is of some concern since this is where some newly introduced artists first begin to sell. Information about weekly radio airplay for each album is collected by Broadcast Data System (BDS). BDS identifies and tracks the airplay of songs by electronically monitoring over 560 radio station signals across the United States. The computer system can recognize over 110,000 different songs by each song's unique patterns

---

1. In fact prices have exhibited such stability that price fixing law has been alleged (Holland 1997).

that are present in its digitized broadcast signal and records. BDS records the song played, the date and time of the broadcast, the radio station that played the song, and the estimated size of the audience reached in a measure similar to GRPs.

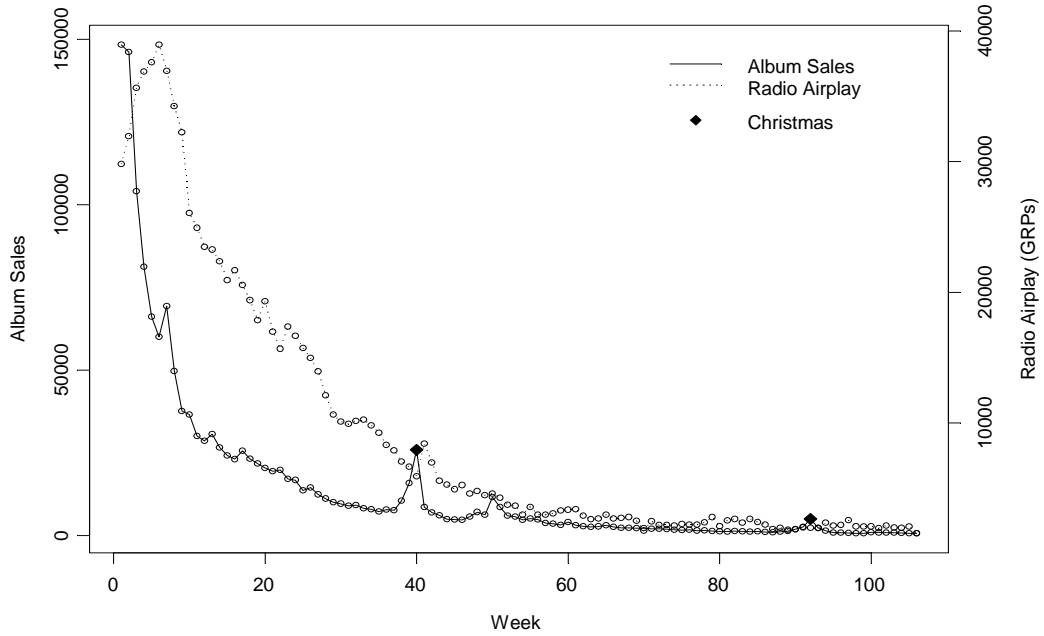
### 2.1 Modeling Album Sales with Univariate ARIMA Models

Since VARMA models are not common in the marketing literature, we begin by considering univariate ARIMA models of sales which is a special case of VARMA models. A commonly used statistical model in linear time series analysis is the univariate ARIMA model (Box, Jenkins and Reinsel 1994). We say  $y_t$  follows an ARIMA( $p, d, q$ ) model if,

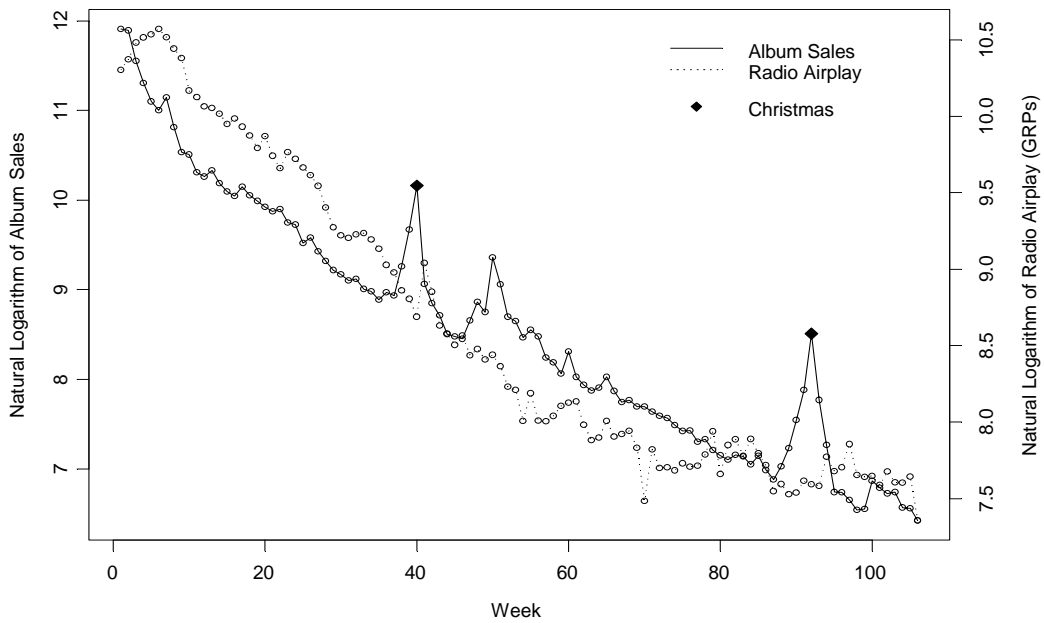
$$(1 - \phi_1 B - \dots - \phi_p B^p)(1 - B)^d y_t = c + \beta \mathbf{x}_t + (1 - \theta_1 B - \dots - \theta_q B^q) \epsilon_t \quad (2)$$

where  $p$ ,  $d$ , and  $q$  are non-negative integers,  $c$ ,  $\phi_p$ ,  $\theta_j$ , and  $\beta$  are parameters,  $B$  is the backshift operator such that  $B y_t = y_{t-1}$ ,  $\{\mathbf{x}_t\}$  is an  $M \times 1$  vector of explanatory variables, and  $\{\epsilon_t\}$  is a sequence of independently and identically distributed normal variables with mean zero and variance  $\sigma^2$ . We consider  $\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$  and  $\theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q$  as polynomials, and assume that the two polynomials have no common factors and have all of their zeros outside the unit circle. The addition of the moving average error structure to the autoregressive process permits parsimonious representations of time series with long-term memory effects.

There are several approaches available in the literature for building ARIMA models. For an excellent introduction to ARIMA modeling for marketing problems, we refer the reader to Hanssens et al (1990). In this paper, we adopt a three-stage iterative approach (see Box, Jenkins and Reinsel 1994) consisting of model specification, estimation, and diagnostic checking. In model specification, we use the sample autocorrelation functions, the sample partial autocorrelation functions, and the extended sample autocorrelation functions of Tsay and Tiao (1984) to identify the values of  $p$  and  $q$ . Once the order ( $p, q$ ) is given, we employ the exact maximum likelihood method to estimate the unknown parameters under the normality assumption (see Ansley 1979, Hillmer and Tiao 1979). For diagnostic checking, we perform various residual analyses including examining



**Figure 1.** Weekly album sales and radio airplay for Bonnie Raitt's *Longing in Their Hearts*.



**Figure 2.** Natural Logarithm of weekly album sales and radio airplay for Bonnie Raitt's *Longing in Their Hearts*.

residual serial correlations by using the Ljung and Box (1978) statistics.

*Empirical result:* To illustrate the ARIMA model consider the album sales of Bonnie Raitt's *Longing in Their Hearts*. Figure 1 plots the weekly sales and airplay support for the album. There is a clear, declining exponential trend in the sales pattern of this album—which is typical of other albums in our dataset. To represent this series in a linear model, we employ a natural log transformation and plot the transformed series in Figure 2.

Additionally, there is a Christmas effect on album sales. Specifically, sales gradually build up before the actual week of Christmas, reach a peak Christmas week, and then return to its usual pattern. To capture the Christmas sales effect we create three intervention variables: a dummy indicator variable for the week of Christmas ( $x_{1t}$ ), a buildup effect that grows exponentially in the weeks prior to Christmas ( $x_{2t}$ ), and a decay variable that declines exponentially in the weeks following Christmas ( $x_{3t}$ ). The variables are defined as:

$$x_{1t} = \begin{cases} 1 & \text{if } t \text{ is Christmas} \\ 0 & \text{otherwise} \end{cases}, \quad x_{2t} = \begin{cases} \delta & \text{if } x_{1,t+1} = 1 \\ \delta x_{2,t+1} & \text{otherwise} \end{cases}, \quad x_{3t} = \begin{cases} 1 & \text{if } x_{1t} = 1 \\ \delta x_{3,t-1} & \text{otherwise} \end{cases} \quad (3)$$

These variables are variants of outlier specifications, such as the temporary change (Tsay 1988), and follow a customary practice in intervention analysis by fixing  $\delta$  to .5.

After examination of the PACF, ACF, and EACF we estimate<sup>2</sup> an ARIMA(1,0,1) model with additional intervention terms to account for Christmas effects:

$$\begin{aligned} (1 - 0.98B)\ln(s_t) = & 0.15 + 2.22x_{1t} + 1.02x_{2t} - 1.62x_{3t} + (1 - 0.22B)\varepsilon_t \\ & (0.01) \quad (0.07) \quad (0.19) \quad (0.15) \quad (0.15) \quad (0.10) \end{aligned} \quad (4)$$

$$\varepsilon \sim N(0, \sigma^2), \quad \sigma = 0.13$$

The standard errors of the parameter estimates are given in parentheses below the estimate. The Ljung-Box statistics for this fitted model yields  $Q(12) = 7.5$ . This statistic is asymptotically distributed as chi-square with six degrees of freedom and indicates that the residual serial correlations are small and that this model specification

---

2. Estimation of this model is performed using the SCA System.

is adequate. All parameter estimates in this model are statistically significant at the 5% level.

Our objective as stated in equation (1) is to derive the conditional expectations for this model. Forecasting with ARIMA models has been well studied. The general ARMAX(1,1) model as used in (4) is:  $\ln(s_t) = \phi_0 + \phi_1 \ln(s_{t-1}) + \beta' \mathbf{x}_t + \epsilon_t - \theta \epsilon_{t-1}$ , and its forecasts are:

$$\ln(\hat{s}_{t+l}) = \begin{cases} \phi_0 + \phi_1 \ln(s_t) + \beta' \mathbf{x}_t + \theta \epsilon_t & \text{if } l=1 \\ \phi_0 + \phi_1 \ln(\hat{s}_{t+l-1}) + \beta' \mathbf{x}_{t+l} & \text{if } l \geq 2 \end{cases} \quad (5)$$

These forecasts are unbiased, normally distributed, and  $\text{var}(\ln(\hat{s}_{t+1})) = \sigma^2$  and  $\text{var}(\ln(\hat{s}_{t+l})) = (1+\theta^2)\sigma^2$  for  $l \geq 2$ . The conditional expectation of sales at time  $l$  in its natural units can be found by taking the exponential of the forecasts and adjusting for the mean of the log-normal distribution (i.e., if  $\ln(y) \sim N(\mu, \sigma^2)$  then  $E[y] = \exp\{\mu + \frac{1}{2}\sigma^2\}$ ; also see Arino and Franses 2000). The solution is:

$$E_{[s_{t+l} | s_t, \mathbf{x}, \phi, \theta, \sigma]} = \begin{cases} \exp\{\phi_0 + \phi_1 \ln(s_t) + \beta' \mathbf{x}_t + \theta \epsilon_t + \frac{1}{2}\sigma^2\} & \text{if } l=1 \\ \exp\{\phi_0 + \phi_1 \ln(\hat{s}_{t+l-1}) + \beta' \mathbf{x}_{t+l} + \frac{1}{2}(1+\theta^2)\sigma^2\} & \text{if } l \geq 2 \end{cases} \quad (6)$$

To better understand the behavior of this expectation suppose that  $l$  is large and that  $|\phi_1| < 1$  (i.e.,  $\phi_1 \neq 0$ ) and employing a geometric series as an approximation to  $\phi_0(1+\phi_1+\phi_1^2+\dots+\phi_1^l)$ , we can show that (6) is approximately:

$$\lim_{l \rightarrow \infty} E_{[s_{t+l}]} \approx \exp\left\{\frac{\phi_0}{1-\phi_1} + \sum_{i=0}^l \phi_1^i \beta' \mathbf{x}_{t+l-i} + \frac{1}{2}(1+\theta^2)\sigma^2\right\} \quad (7)$$

Equation (7) implies that the series will eventually settle down to the mean of the process ( $\phi_0/(1-\phi_1)$ ).

The assumption that  $|\phi_1| < 1$  may be in doubt since the estimate of this autoregressive coefficient in (4) suggests that the process may be integrated, i.e.,  $\phi_1 = 1$ . In other words the process has a long-term memory such that shocks which occur in earlier periods persist permanently and shift the expectation of sales in subsequent periods. To test this hypothesis we perform an augmented Dickey-Fuller test (1979), for a marketing application of this test see Dekimpe and Hanssens (1995b). The asymptotic p-value for the test that the logarithm of sales

follows a random-walk with drift is .03, i.e.,  $H_0: \phi_1=1$  versus  $H_A: \phi_1 \neq 1$ , which tends to support the random-walk hypothesis. Under the assumption that  $\phi_1=1$  then the conditional expectation from (6) can be simplified as:

$$E_{[s_{t+l} | s_t, \mathbf{x}_t]} = \exp \left\{ (l+1) \left( \phi_0 + \frac{1}{2} (1-\theta^2) \sigma^2 \right) + \sigma^2 \theta^2 + \sum_{i=0}^l \beta' \mathbf{x}_{t+l-i} + \theta \epsilon_t \right\} s_t \quad (8)$$

Again if we ignore the contribution of the seasonal factors, we can show that this expectation will be decreasing as long as  $\phi_0$  is sufficiently negative, i.e.,  $\phi_0 < -\frac{1}{2}(1-\theta)^2 \sigma^2$ . This relation shows that both the first and second moments of the log transformed process are important when forecasting sales, and the contribution of the variance of the process should not be ignored when forecasting sales.

The implications of unit-roots have been thoroughly researched in the time series literature (Dickey and Fuller 1976) and their impact on marketing research has also been recognized (Dekimpe and Hanssens 1995a). The most critical impact of the assumption that  $|\phi_1| < 1$  is that the long term forecasts of the process will be positive. Using the approximation given in (8) we would expect sales to average around 2,480 units in the long term, while under model (9) sales would be driven to zero. For near-term forecasts there is little difference between an autoregressive model with and without the assumption that  $\phi_1=1$ . The existence of a unit-root in a bivariate VARMA model will have special significance and we will return to this later in the next section.

## 2.2 Modeling Album Sales and Radio Airplay Jointly with VARMA Models

A multivariate generalization of the univariate ARIMA model given in (2) is the vector autoregressive moving-average (VARMA) model. A vector of time series  $\mathbf{y}_t = [y_{1t} \dots y_{kt}]'$  is said to follow a VARMA model if,

$$(I - \Phi_1 B - \dots - \Phi_p B^p) \mathbf{y}_t = \mathbf{c} + \mathbf{B} \mathbf{x}_t + (I - \Theta_1 B - \dots - \Theta_q B^q) \boldsymbol{\epsilon}_t \quad (9)$$

where  $\mathbf{c}$  is a constant vector,  $\Phi_j$  and  $\Theta_j$  are  $k \times k$  matrices,  $I$  is the  $k \times k$  identity matrix,  $\mathbf{x}_t$  is a matrix of explanatory variables—which in our case are seasonal effects, and  $\{\boldsymbol{\epsilon}_t\} = \{[\epsilon_{1t} \dots \epsilon_{kt}]'\}$  is a sequence of independently and identically distributed random vectors with mean zero and positive-definite covariance matrix  $\Sigma$ .

Notice that in the VARMA representation, vectors take the place of scalars in the ARIMA model, and parameters become matrices. We assume that the two matrix polynomials  $\Phi(B)$  and  $\Theta(B)$ , defined in a similar way as those of (2), are left-coprime<sup>3</sup> and that all of the zeros of the determinants  $^*\Phi(B)^*$  and  $^*\Theta(B)^*$  are on or outside the unit circle. For further identifiability properties of this model, see Hannan and Deistler (1988) and Tiao and Tsay (1989).

The VARMA models in (10) have not been widely employed in practice. However, VARMA models are very flexible and encompass many familiar models in the literature. The vector autoregressive (VAR) model which is commonly used in the econometric literature is a special case. For a marketing study using VAR models see Dekimpe and Hanssens (1995b). Another common example is the transfer function or dynamic regression model. For example, consider the case of  $k=2$ . If all the lower off-diagonal elements of  $\Phi_j$  and  $\Theta_j$  are zero, then the model reduces to a transfer function model, e.g., regression model with lagged inputs, with  $y_{2t}$  and  $y_{1t}$  as its input and output variables, respectively. For a further discussion of VARMA models see Tiao and Box (1981).

The flexibility of VARMA methods allows us to identify and model very different relationships without having to first specify a causal structure between variables. Independent, contemporaneous, or lagged relationships are all special cases of this model. For our bivariate time series, album sales and radio airplay, if these series are unrelated, then  $\Phi$ ,  $\Theta$ , and  $\Sigma$  are diagonal. Hence, each series can be represented by an independent ARIMA model. A contemporaneous relationship exists when  $\Sigma$  is non-diagonal. If  $\Phi$  and  $\Theta$  are upper triangular, then previous radio airplay impacts album sales but there is no feedback relationship of previous sales on airplay. Alternatively, if  $\Phi$  and  $\Theta$  are lower triangular, then album sales affects airplay. Finally, if  $\Phi$ ,  $\Theta$ , and  $\Sigma$  are fully populated, we have a contemporaneous and dynamic relationship between the two series. In summary, VARMA presents a flexible and complete model for studying the inter-relationships between these two series and can easily be extended beyond the bivariate case to capture higher-order multivariate relationships.

*Empirical result:* We build the following VARMA model for the album sales and radio airplay of Bonnie

---

3. Two matrix polynomials are left-coprime if they do not have common left factors. This is a matrix generalization of the univariate ARIMA models in (1) where the AR and MA polynomials are assumed to have no common factors.

Raitt's *Longing in their Hearts* used in the previous section. Here we let  $y_t = [\ln(s_t) \ln(a_t)]'$  where  $s_t$  and  $a_t$  denote sales and radio airplay in GRPs respectively. The model we estimate<sup>4</sup> is:

$$(I - \Phi B)y_t = c + Bx_t + (I - \Theta B)\varepsilon_t, \quad \text{cov}(\varepsilon_t) = \Sigma \quad (10)$$

where

$$\Phi = \begin{bmatrix} .91 & .08 \\ (.005) & (.01) \\ .05 & .92 \\ (.02) & (.03) \end{bmatrix}, \quad \Theta = \begin{bmatrix} .04^* & .14^* \\ (.07) & (.14) \\ .07^* & .27 \\ (.08) & (.11) \end{bmatrix}, \quad c = \begin{bmatrix} -.08 \\ (.02) \\ .21 \\ (.09) \end{bmatrix}, \quad \beta = \begin{bmatrix} 2.19 \\ (.18) \\ 1.00 \\ (.16) \\ -1.51 \\ (.18) \end{bmatrix}, \quad \Sigma = \begin{bmatrix} .02 & .00 \\ .00 & .010 \end{bmatrix}, \quad B = \begin{bmatrix} \beta' \\ 0 \end{bmatrix}$$

The standard errors of the parameter estimates are given in parentheses below the estimate. The “\*” superscript denotes the associated parameter as statistically insignificant at the 5% level based on its asymptotic distribution. As expected, the seasonal variables have a significant effect on sales. Additional analyses of the radio airplay reveals that it is not effected by the Christmas season. Therefore, seasonal effects on airplay were not included in the model.

Upon further inspection of the parameters we can assess the dynamic relationship between sales and airplay. The main diagonal along the  $\Phi$  matrix indicates that both the sales and the airplay series have significant autoregressive components. This means that observations for each of the two series do indeed depend on their past observations. The off-diagonal also reflects a feedback relationship between sales and airplay. That is, past sales increases airplay, and in turn, past airplay increases sales. Looking at the off-diagonals of the  $\Phi$  matrix, airplay seems to have a stronger effect on sales in terms of magnitude than sales on airplay for this particular album. Compare these results with those from the univariate ARIMA model. The moving-average (MA) term for sales was significant when sales was analyzed as a univariate series but failed to be significant when a VARMA

---

4. The estimation of this model using the exact likelihood is performed using MATLAB scripts that are available from the authors upon request. Standard statistical time series software packages do not allow the estimation of VARMA models with exogenous terms.

model was applied. The simplification of the structure of  $s_t$  is due to the inclusion of airplay, which illustrates that when the multivariate nature of the data is exploited, relationships can be simplified. However, the marginal univariate representation implied by the VARMA model for the log of sales is still an ARMA(1,1) model.

The error structure also provides information about the contemporaneous relationship between sales and airplay. The correlation coefficient computed from the error covariance matrix is .022. Statistical tests indicate that the errors from the sales and airplay models are uncorrelated. This implies– that although–there may be a lagged relationship between sales and airplay, the contemporaneous relationship is negligible.

A final point that the VARMA model clarifies about the dependence between sales and airplay is the existence of a joint trend. Notice that one of the eigenvalues of  $\Phi$  is close to 1. In the econometric literature this is known as cointegration (Engle and Granger 1987), for an marketing application refer to Frances (1994). Cointegration means that album sales and radio airplay share a common trend that influences both series, in other words they are evolving together and not independently of each other. The common trend can be constructed by rewriting (11) in terms of the following transformations of  $\mathbf{z}_t = \mathbf{P}^{-1}\mathbf{y}_t$ ,  $\mathbf{c}^* = \mathbf{P}^{-1}\mathbf{c}$ ,  $\Theta^* = \mathbf{P}^{-1}\Theta\mathbf{P}$ ,  $\epsilon_t^* = \mathbf{P}^{-1}\epsilon_t$ ,  $\Sigma^* = \mathbf{P}^{-1}\Sigma\mathbf{P}^{-1}$ ,  $\mathbf{B}^* = \mathbf{P}^{-1}\mathbf{B}$ , and  $\mathbf{P}$  is defined using the Jordan decomposition of  $\Phi = \mathbf{P}\Lambda\mathbf{P}^{-1}$ :

$$\mathbf{z}_t = \mathbf{c}^* + \Lambda\mathbf{z}_{t-1} + \mathbf{B}^* \mathbf{x}_t + \epsilon_t^* + \Theta^* \epsilon_{t-1}^* \quad (11)$$

where

$$\mathbf{P}^{-1} = \begin{bmatrix} .66 & -.79 \\ .62 & .82 \end{bmatrix}, \Lambda = \begin{bmatrix} .85 & 0 \\ 0 & .98 \end{bmatrix}, \Theta^* = \begin{bmatrix} .05 & -.10 \\ -.11 & .26 \end{bmatrix}$$

$$\mathbf{c}^* = \begin{bmatrix} -.22 \\ .12 \end{bmatrix}, \mathbf{B}^* = \begin{bmatrix} 1.44 & .65 & -.99 \\ 1.35 & .62 & -.93 \end{bmatrix}, \Sigma^* = \begin{bmatrix} .015 & .002 \\ .002 & .014 \end{bmatrix}$$

Notice that the first element ( $z_{1t}$ ) is stationary, while the second ( $z_{2t}$ ) follows a random walk. Alternately we can think of the original series as a composition of these two unobserved components. The model for album sales is a linear combination of a random walk ( $z_{2t}$ ) and a stationary ARMA process ( $z_{1t}$ ). Similarly radio airplay can also be expressed as a different linear combination of the same processes. In summary, although both album

sales and radio airplay follow trends, an appropriate bivariate model does not yield two separate trends as one might attempt if separate univariate models were constructed separately, but a single, common trend.

Forecasting the VARMA is similar to the ARIMA model. The forecasts for the VARMAX(1,1) model as used in (10) are:

$$\hat{y}_{t+l} = \begin{cases} \mathbf{c} + \Phi \mathbf{y}_t + \mathbf{B} \mathbf{x}_t - \Theta \hat{\mathbf{e}}_t & \text{if } l=1 \\ \mathbf{c} + \Phi \hat{y}_{t+l-1} + \mathbf{B} \mathbf{x}_t & \text{if } l \geq 2 \end{cases} \quad (12)$$

These forecasts are unbiased, normally distributed, and  $\text{var}(\hat{y}_{t+1}) = \Sigma$  and  $\text{var}(\hat{y}_{t+l}) = \Sigma + \Theta \Sigma \Theta'$  for  $l \geq 2$ . The conditional expectation of sales at time  $l$  in its natural units can be found by taking the exponential of the forecasts and adjusting for the mean of the log-normal distribution as before. The solution is:

$$E_{[s_{t+l} | s_t, a_t, \mathbf{x}, \Phi, \Theta, \Sigma]} = \begin{cases} \exp(\mathbf{v}' [\mathbf{c} + \Phi \mathbf{y}_t + \mathbf{B} \mathbf{x}_t - \Theta \hat{\mathbf{e}}_t] + \frac{1}{2} \sigma_{11}) & \text{if } l=1 \\ \exp(\mathbf{v}' [\mathbf{c} + \Phi \hat{y}_{t+l-1} + \mathbf{B} \mathbf{x}_t] + \frac{1}{2} \gamma_{11}) & \text{if } l \geq 2 \end{cases}, \quad (13)$$

where  $\mathbf{v}' = [1 \ 0]$ ,  $\sigma_{11} = \mathbf{v}' \Sigma \mathbf{v}$ , and  $\gamma_{11} = \mathbf{v}' (\Sigma + \Theta \Sigma \Theta') \mathbf{v}$ . In equation (13) we assume that we must forecast radio airplay along with album sales, if radio airplay is known then we can treat it as an exogenous variable and forecast the problem as an ARIMAX model as done in (6).

### 3 Comparison of Time Series and Diffusion Modeling

Diffusion models have proved to be very popular in the marketing literature for forecasting sales (Bass 1969). Given the prevalence of diffusion models (Mahajan, Muller, and Bass 1990) and the lack of experience with VARMA models in the marketing literature, we compare these two types of models. We show that under certain circumstances VARMA models can well approximate the predictions of diffusion models, but do not suffer from some of the shortcomings as multivariate diffusion models with trending covariates (Bass and Srinivasan 2002).

The diffusion pattern observed for music album sales and radio airplay is different, although not unrelated, from those observed for new durables (Bass 1969). Usually, the sales peak occurs at introduction and

then slowly declines through time, instead of a bell-shaped sales curve. This pattern is similar to those observed with other experiential products such as books and movies (Sawhney and Eliashberg 1996). The exponential decline over time for sales implies that a good one is the Fournier and Woodlock (1960) model:

$$F(t) = 1 - e^{-\lambda t} \quad (14)$$

where  $F(t)$  denotes the cumulative probability of trial for an individual at time  $t$ ,  $\lambda$  is the rate of trial, and  $\lambda > 0$ . Notice that this is a special case of the Bass (1969) model when his innovation of diffusion parameter equals zero. Empirical estimates of the innovation of diffusion model using the Bass model for the data set employed in this paper indicate that  $q$  is indeed close to 0. The predicted sales from the Fournier and Woodlock model during any period can be computed as follows:

$$E[s_t | \lambda, m] = m(F(t) - F(t-1)) = m(e^{-\lambda(t-1)} - e^{-\lambda t}) \quad (15)$$

where  $m$  is the market size.

The  $t$ -step ahead forecast of sales from the logarithm of sales that follows a random walk with a negative trend, such as the ARIMA(0,1,0) model given in (2), is:

$$E[s_t | s_0] = \exp\left\{t\left(c + \frac{1}{2}\sigma^2\right)\right\} s_0 \quad (16)$$

This derivation makes use of the mean of the log-normal distribution,  $E[\exp\{\epsilon\}] = \exp\{\frac{1}{2}\sigma^2\}$ .

Notice that upon making the following substitutions  $s_0 = m(\exp\{t-\lambda\})$  and  $\lambda = c + \frac{1}{2}\sigma^2$  that (15) and (16) are identical. This means that the expectation of the ARIMA(0,1,0) model under these circumstances (e.g., log transformation and random walk model with negative trend) is the same as the Fournier and Woodlock (1960) model. To the best of our knowledge this is the first time this relationship between ARIMA and diffusion models has been pointed out.

A generalization of the Fournier and Woodlock model that incorporates covariates can be written as follows (see Kalbfleisch and Prentice 1980, Lawless 1982, Radas and Shugan 1998, Moe and Fader 2001).

$$F(t) = 1 - \exp\left\{-\lambda \sum_{i=1}^t \exp\{\beta' \mathbf{x}_i\}\right\} \quad (17)$$

If we specify  $\beta$  and  $\mathbf{x}_t$  as follows,

$$\beta = \begin{bmatrix} \delta \\ 1 \end{bmatrix}, \quad \mathbf{x}_t = \begin{bmatrix} 1 \\ \ln(a_t) \end{bmatrix} \quad (18)$$

then the sales predicted by (17) is

$$s_t = m \exp\left\{-\lambda \sum_{i=1}^{t-1} a_i\right\} [1 - \exp\{-\lambda a_t\}] \quad (19)$$

The multivariate generalization of the ARIMA model, the VARMA model, can also be used to approximate the multivariate version of the Fourt and Woodlock model in (19). Consider the following VARMA model:

$$\left( \mathbf{I} - \begin{bmatrix} 1 & 0 \\ \Phi_{21} & \Phi_{22} \end{bmatrix} \mathbf{B} \right) \begin{bmatrix} \ln(s_t) \\ a_t \end{bmatrix} = \begin{pmatrix} c \\ c_a \end{pmatrix} + \boldsymbol{\varepsilon}_t, \quad \text{cov}(\boldsymbol{\varepsilon}_t) = \boldsymbol{\Sigma} \quad (20)$$

and the forecasts generated by such a model:

$$E_{[s_t | s_0, a_1, \dots, a_t]} = \exp\left\{t\left(c + \frac{1}{2}\sigma^2\right) + \sum_{i=1}^t \delta a_i\right\} s_0 \quad (21)$$

If we assume that  $c + \frac{1}{2}\sigma^2 = 0$  and  $m = s_0$ , then

$$E_{[s_t | s_0]} = m \exp\left\{\delta \sum_{i=1}^t a_i\right\} \quad (22)$$

This shows that under certain conditions the Fourt and Woodlock and the VARMA models can provide similar forecasts.

The flexibility of VARMA models help avoid  $\lambda$  problems that may develop when using diffusion models of multivariate processes. One problem that can arise when using Fourt and Woodlock models on multivariate processes is that they can identify spurious relationships between the two independent processes. For example,

suppose we have two diffusion processes that evolve independently of one another, say album sales and radio airplay. Let the model for album sales be as in (13) with  $m=1,500,000$  and  $\lambda=.06$ , and let  $m=1,000,000$  and  $\lambda=.04$  for radio airplay. If we simulate each of these models with 104 observations, the MLE estimates of the parameters for the sales series are  $\lambda=.0618$  (.0005) and  $M=1,502,000$  (1,200). Now suppose we did not know a priori that sales and airplay were independent, but instead estimated a multivariate Fourt and Woodlock model of album sales with radio airplay as a covariate as in (16). The MLE estimates from this sales model are  $\lambda=.0567$  (.0012),  $M=1,503,000$  (1,100), and  $\beta=.0321$  (.0007). Notice that in this multivariate Fourt and Woodlock model we would find that airplay has a significant effect on sales (i.e.,  $\beta \neq 0$ )—a spurious relationship. This spurious relationship exists because we have modeled the levels of two series with similar trends. This could lead to serious managerial problems, since the multivariate Fourt and Woodlock model may lead to music publishers investing substantial amount of money in radio airplay to increase album sales, when in truth no such relationship exists. This could present a potentially costly error.

Problems of spurious relationships are well documented in the regression literature, for a good discussion see Granger and Newbold (1986, pp. 205-215). The motivation for these spurious relationships is that simple regressions of a trended series upon another trended series will lead to seemingly significant regression coefficients. The problem is that these types of regressions violate the underlying assumptions of our model. A possible way of correcting this problem is to relate the changes in the series (i.e., the first order differences). Since models on the differenced data are a special case of a VARMA model created from the raw data, the VARMA models are much less susceptible to these problems of spurious relationships than multivariate forms of diffusion models without requiring the analyst to difference the data before modeling the data. We point out this shortcoming of multivariate diffusion models since past researchers have suggested these types of multivariate models without verifying whether their results are due to spurious correlations, and they demonstrate an important advantage of VARMA models with diffusion type data when we are relating variables with trends without a priori knowledge of the feedback relationship.

## 4 Valuation of Radio Airplay in terms of Album Sales

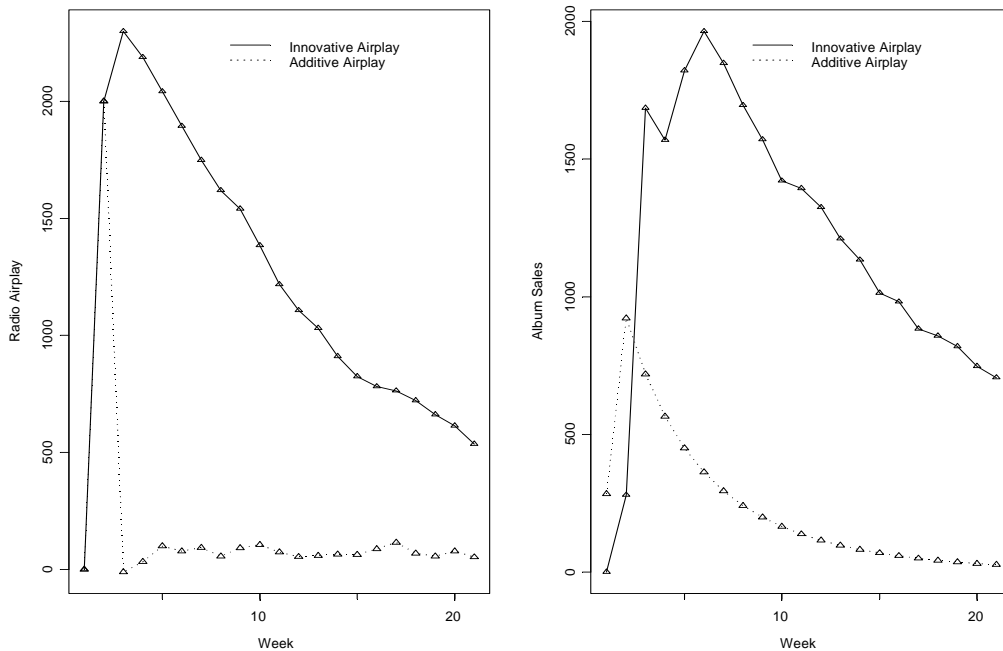
In this section we use the VARMA model and forecasting techniques presented in §2 to answer the question: How valuable is increased radio airplay in terms of increased album sales of Bonnie Raitt's *Longing in their Hearts*? We refer the reader to the VARMA estimates in equation (10) for the estimates associated with the model for this series. These estimates do not indicate a contemporaneous relationship between radio airplay and album sales since the error covariance matrix is diagonal<sup>5</sup> (i.e., an increase in radio airplay at time  $t$  does not increase album sales at time  $t$ ). However, the estimates of  $\Phi$  do show a feedback relationship between the two series (i.e., an increase in radio airplay at time  $t$  will increase album sales in subsequent periods). The logarithmic scale means that we can approximately interpret the coefficients as percentage changes. The .08 estimate of the (1,2)<sup>th</sup> position in  $\Phi$  can be directly interpreted as a 1% increase in radio airplay will result in a 8% increase in expected album sales in the following period, the .05 estimate of the (2,1)<sup>th</sup> position implies a 1% increase in album sales will result in a 5% increase in expected radio airplay in the following period. Both of these effects are statistically significant at the 5% level and illustrate the dynamic relationship between the two series. These effects give only the effects for the first period following the change, the autoregressive nature of the model means that these shocks will persist through the system, although they will diminish through time. Although the moving average parameter of radio airplay, which corresponds with the (2,2)<sup>th</sup> position of  $\Theta$ , illustrates that random shocks to radio airplay are more persistent than album sales.

To better illustrate the results on album sales of an increase in radio airplay for a single week, we plot the expected contribution radio airplay on incremental album sales through each of the subsequent weeks in the second panel of Figure 3 for a increase of 2 million GRPs of radio airplay, see the line denoted as *innovative airplay* in the first panel. Innovative forecasts are generated by adding 2 million GRPs to the observed radio airplay

---

5. Even though the contemporaneous terms do not occur in a VARMA model, contemporaneous effects are captured through correlated error structures. It needs to be remembered that VARMA models are in reduced form. Consider the VARMA model:  $\mathbf{y}_t = \Phi \mathbf{y}_{t-1} + \varepsilon_t$ ,  $\varepsilon_t \sim N(0, \Sigma)$ , where  $\mathbf{y}_t = \{y_{1t}, y_{2t}\}$ , and the transformation  $P$  where  $\Sigma = P\Lambda P'$ . After multiplication of this model by  $P$ ,  $P\mathbf{y}_t = \Phi^* \mathbf{y}_{t-1} + \mathbf{u}_t$ ,  $\mathbf{u}_t \sim N(0, \Lambda)$ , we can directly observe that  $p_{11}y_{1t} + p_{12}y_{2t} = \phi_{11}^*y_{1,t-1} + \phi_{12}^*y_{2,t-1} + u_{1t}$ , or in structural form there is a direct relationship between  $y_{1t}$  and  $y_{2t}$  induced through the correlation in the errors.

series in the first week and treating this as the realized value when forecasting. (We should not that these are the projections from the model, and this experiment was not conducted in practice.) Notice there would be no expected increase in sales during the week with the additional airplay, but in the week immediately following sales would increase by 280 units and in the following week by 1,685 additional units. The autoregressive and feedback relationship between the two series will result in album sales gradually diminishing. In total this additional 2 million GRPs in the first week results in additional album sales of more than 24,900 units during the following twenty weeks. During low season these additional GRPs would cost \$4,800, while during high season this may cost \$8,800<sup>6</sup>. The retail price of a CD is \$18, so these additional sales would result in additional retail revenues of almost \$449,000. The wholesale price of a CD is around \$9 although the incremental cost of producing an album is less than \$1, hence the incremental income from an additional album sale is \$8. So the increased radio airplay from these additional 2 million GRPs would generate \$199,400 in additional profit less the cost of the



**Figure 3.** Increase in expected radio airplay (first panel) and album sales (second panel) from an increase of 2 million GRPs of radio airplay in week 1 for Bonnie Raitt's *Longing in their Hearts*.

6. These figures are based on a survey of radio stations in the Philadelphia market.

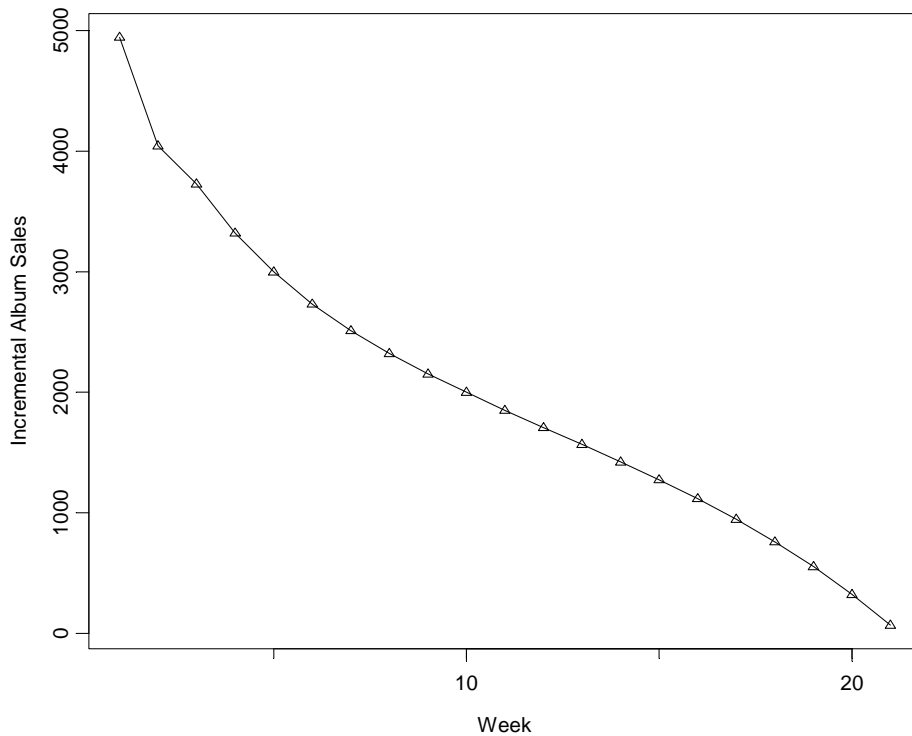
airplay. Clearly this increased radio airplay could be very profitable.

The reason that it appears that increased radio airplay would be so profitable is that the autoregressive relationships of model (10) magnify the initial increase in airplay. In other words the radio stations continue to increase radio airplay of the album even after the purchase radio airplay is completed. The first panel of Figure 3 shows the expected radio airplay. Notice the expected increase in radio airplay and the slowly decaying pattern. Our model assumes that the additional radio airplay that is purchased will be incorporated into the series just as traditional programmed radio airplay. We can think of this incremental radio airplay as an intervention or outlier (Tsay 1988). In the time series literature these effects would be called innovative outlier (IO) since it becomes interleaved into the usual dynamic effects. However, we would argue that radio stations will make a distinction between the radio airplay that they program versus the radio airplay that is sponsored. It is more likely that the effects of the purchased airplay will not become dynamically interleaved in future radio airplay periods, but simply result in a one-period increase such as with an additive outlier (AO). An additive outlier assumes that the incremental airplay will not be directly incorporated into the time series process of airplay, although there could still be some feedback relationship from higher album sales.

We recalculate the effect of an additional 2 million GRPs of radio airplay assuming that this purchased airplay will be treated as an AO and forecast the results on expected airplay and album sales in Figure 3, see the lines denoted as *additive airplay*. Forecasts with the AO are generated by adding 2 million GRPs to the realized radio airplay series for the first week, but when generating forecasts from this point ignoring these additional GRPs. Notice that under this scenario the incremental effects of the additional radio airplay would result in 4,950 additional album sold, for a total increase in expected profits of \$39,600 less the cost of the additional radio airplay. We still find that additional radio airplay would be quite profitable to the music label. Notice that the incremental effect on radio airplay is quite dramatic from this assumption (see the first panel of Figure 3), effectively limiting the effect of the increased airplay to the first week.

Regardless of whether additional radio airplay is treated as either an AO or IO, we find that additional airplay is quite profitable. However, the only way to determine how radio stations will actually respond to

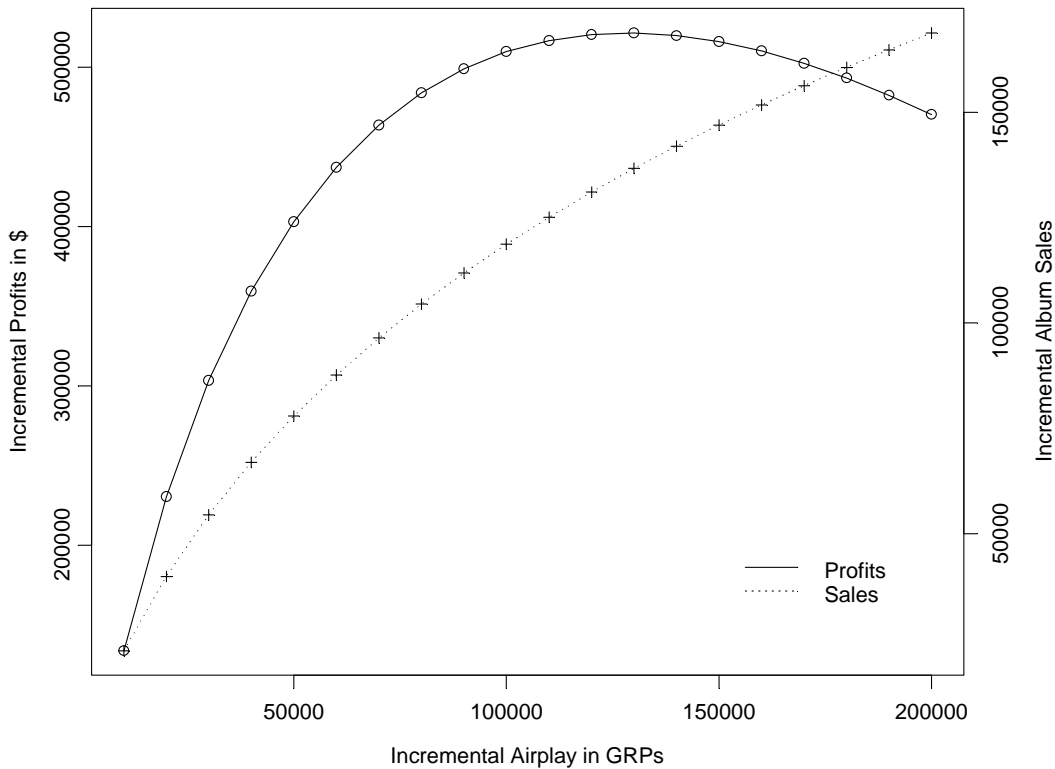
sponsored radio airplay is through experimentation, which may occur naturally if sponsorship is adopted or through artificial experiments. To the best of our knowledge these experiments have not been carried out. In the remainder of the paper we assume that any purchased radio airplay will behave as an AO since we would expect that radio stations will make a distinction between sponsored airplay and airplay that is chosen by their program directors. Our treatment of sponsored radio airplay as an AO will also yield a more conservative estimate to the profitability of sponsored radio airplay.



**Figure 4.** Expected incremental sales from 2 million GRPs of additional radio airplay purchased at the corresponding week.

Another point that we wish to make is that since sales follow an exponential model where airplay and sales are expected to diminish through time, we can expect that the contribution of additional radio airplay will also decrease through time. Effectively this means that increased radio airplay in week 2 will be less effective than in week 1. To illustrate this effect in Figure 4 we plot the incremental sales that would accrue for 20 weeks if airplay were increased by 2 million GRPs in the corresponding week, but only in that week. In the previous

discussion we found that increasing radio airplay in week one would result in 4,950 additional album sales. If this increased radio airplay was purchased in week 2, then there would expect only 4,050 additional album sales. By the 21<sup>st</sup> week the same 2 million GRPs would only generate an additional 70 album sales. Although increased radio airplay can retard the natural sales rate, it would be prohibitively expensive to attempt to totally overcome it.



**Figure 5.** Incremental album sales and profits from sponsored radio airplay at week 1.

A final question that we consider is what is the optimal amount of radio airplay? To answer this question we compute the expected profits sales associated with 10 million through 200 million additional GRPs and plot the results in Figure 5. We translate increased album sales into profits by assuming that one million GRPs cost \$4,400 and the marginal profit from each album is \$8 as before. We find that greatest increase in profits is associated with an additional 130,000 GRPs which we would expect to increase album sales by 136,700 units for an increase in profits of \$521,400. We note that this represents more than a four-fold increase over the 29,810

units that Bonnie Raitt's *Longing in Their Hearts* album actually received. Again we remind the reader that these projections occur out of range of our dataset, hence some caution needs to be exercised in determining whether this dramatic an increase will change the behavior of the observed series.

## 5 Application to a Portfolio of Albums

To further illustrate these techniques we apply them to a set of thirteen albums released by Capitol Records between 1993 and 1995. These albums are listed in Table 1 along with some descriptive information. The albums fall predominantly in the popular genre, but represent a fairly wide good sample of main-stream and groups that appeal to narrow audiences. Additionally, we have series for both new and existing artists. Notice the wide disparity in album sales in the launch week, which falls anywhere between a few hundred albums to a couple hundred thousand, and radio airplay.

	Group	Album	Genre	Number of previous releases	Album Release Date	Sales in Launch Week	Airplay in Launch Week
1	Beastie Boys	Ill Communication	Rap	5	6/5/1994	220,052	5,795
2	Richard Marx	Paid Vacation	Pop	4	2/13/1994	27,221	3,564
3	Everclear	Sparkle & Fade	Pop	1	5/28/1995	2,061	4,349
4	Bob Seger	It's a Mystery	Pop	10	10/29/1995	33,302	3,746
5	Adam Ant	Wonderful	Pop	7	3/12/1995	6,640	4,188
6	Blind Melon	Soup	Pop	1	8/20/1995	33,409	15,068
7	Charles & Eddie	Chocolate Milk	R&B	1	9/03/1995	261	376
8	Dink	Dink	Pop	0	11/20/1994	1,375	1,550
9	John Hiatt	Walk On	Pop	14	10/29/1995	21,493	3,109
10	Radiohead	The Bends	Pop	1	4/9/1995	5,250	1,710
11	Bonnie Raitt	Longing In Their Hearts	Pop	12	3/27/1994	148,337	29,810
12	Bonnie Raitt	Road Tested	Pop	13	11/12/1995	25,841	9,185
13	Smoking Popes	Born to Quit	Pop	1	7/9/1995	3,107	4,003

**Table 1** Description of albums from Capitol Record in dataset.

We apply the VARMA model with explanatory variables (VARMAX) specified in equation (5) to these other twelve albums. Usual time series identification techniques led to the specification of the VARMA(1,1) model for all albums. It is generally recognized that the ARMA(1,1) model is a good approximation to many observed time series. In other applications higher order models may be necessary, but residual diagnostics show that the VARMA(1,1) in our dataset is satisfactory. Table 2 displays the parameter estimates for these models.

Album	$c$	$\Phi$			$\Theta$		$\beta_1$	$\beta_2$	$\beta_3$	$\Sigma$	
<b>Adam Ant:</b>	-.62	.90	.15	<i>-.00</i>	<i>-.00</i>	.95	.90	-.39	.018	.005	
<i>Wonderful</i>	.96	.23	.68	<i>-.00</i>	<i>-.00</i>				.005	.018	
<b>Beastie Boys:</b>	.41	.93	.02	<i>.00</i>	<i>.00</i>	1.21	.73	-.63	.014	.000	
<i>Ill Communication</i>	.69	.09	.80	<i>.00</i>	<i>.01</i>				.000	.031	
<b>Blind Melon:</b>	1.09	.63	.23	<i>.00</i>	<i>.00</i>	0.61	1.21	.33	.020	.011	
<i>Soup</i>	-1.04	.33	.76	<i>-.00</i>	<i>-.00</i>				.011	.023	
<b>Charles &amp; Eddie:</b>	.64	.69	.20	.42	.30	<i>-.05</i>	.65	.13	.049	.025	
<i>Chocolate Milk</i>	<i>-.24</i>	.17	.79	.77	.59				.025	.137	
<b>Dink:</b>	.09	.96	.03	.22	.02	<i>.56</i>	<i>.44</i>	<i>-.04</i>	.026	.005	
<i>Dink</i>	<i>-.73</i>	.23	.83	<i>-.40</i>	<i>.34</i>				.005	.926	
<b>Everclear:</b>	<i>-.10</i>	.91	.11	<i>.01</i>	<i>-.17</i>	.73	.59	<i>-.21</i>	.013	.002	
<i>Sparkle &amp; Fade</i>	<i>.02</i>	<i>-.01</i>	1.01	<i>.03</i>	<i>-.37</i>				.002	.024	
<b>John Hiatt:</b>	.42	.74	.23	<i>.00</i>	<i>-.00</i>	1.68	.72	-1.01	.012	.007	
<i>Walk On</i>	.43	.11	.81	<i>-.00</i>	<i>-.01</i>				.007	.030	
<b>Richard Marx:</b>	<i>.09</i>	1.00	<i>-.02</i>	<i>.19</i>	<i>-.10</i>	1.18	.60	-.46	.012	.005	
<i>Paid Vacation</i>	<i>.11</i>	.10	.88	<i>.18</i>	<i>.58</i>				.005	.063	
<b>Radiohead:</b>	.69	.87	.06	<i>-.00</i>	<i>-.05</i>	1.13	<i>.34</i>	-.54	.011	.001	
<i>The Bends</i>	2.31	<i>-.31</i>	1.04	<i>-.01</i>	<i>-.22</i>				.001	.121	
<b>Bonnie Raitt:</b>	<i>-.08</i>	.92	.09	<i>.04</i>	<i>.14</i>	2.19	1.00	-1.51	.018	.000	
<i>Longing In Their Hearts</i>	<i>.21</i>	.05	.92	<i>.07</i>	<i>.27</i>				.000	.010	
<b>Bonnie Raitt:</b>	1.53	.75	.10	<i>.01</i>	<i>.03</i>	2.19	1.13	-1.36	.012	.004	
<i>Road Tested</i>	.28	<i>-.08</i>	1.04	.35	.98				.004	.039	
<b>Bob Seger:</b>	<i>-.57</i>	.92	.15	<i>.00</i>	<i>-.00</i>	2.04	<i>.91</i>	<i>-1.33</i>	.015	.006	
<i>It's a Mystery</i>	.85	<i>.01</i>	.88	<i>-.00</i>	<i>-.00</i>				.006	.014	
<b>Smoking Popes:</b>	2.36	<i>.14</i>	.54	<i>.09</i>	<i>.47</i>	<i>-.44</i>	<i>.16</i>	1.21	.013	<i>-.007</i>	
<i>Born to Quit</i>	.48	<i>-.15</i>	1.06	<i>.17</i>	<i>.87</i>				<i>-.007</i>	.312	

**Table 2.** Parameter estimates from VARMA(1,1) for all albums in the dataset. Italicized figures denote estimates that are not statistically different from zero at a .05 confidence level.  $\beta_1$ ,  $\beta_2$ , and  $\beta_3$  denote parameters associated with the Christmas week, Christmas build-up, and Christmas decay effects, respectively.

A cursory look at the results of Table 2 reveals that strong autoregressive relationships between sales and airplay and their past values, but beyond that there is a great deal of diversity amongst the relationships between sales and airplay. Most albums show a dynamic feedback relationship between sales and airplay, i.e., the models have significant parameter estimates in the off-diagonals of the  $\Phi$  and  $\Theta$  matrices. Bonnie Raitt's *Longing in Their Hearts* is one example of an album with feedback effects that was discussed in the previous section. All but two albums (Beastie Boys's *Ill Communication* and Richard Marx's *Paid Vacation*) show airplay has a significant

effect on subsequent sales. Similarly, all but three albums show that sales can alter future airplay (Everclear's *Sparkle & Fade*, Bob Seger's *It's a Mystery*, and Smoking Popes' *Born to Quit*). However, the off-diagonal elements of the  $\Phi$  and  $\Theta$  matrices represent lagged effects. The correlation coefficient reveal contemporaneous relationships between radio airplay and album sales. Only three albums show significant contemporaneous relationships (Adam Ant's *Wonderful*, Blind Mellon's *Soup*, and Richard Marx's *Paid Vacation*). It would appear that for most albums radio airplay does not induce consumers to go directly to a retail store, but instead takes at least a week before we can measure any impact.

A closer examination of the autoregressive parameters shows that in most cases, album sales and radio airplay follow a common trend, i.e., the series are cointegrated. Cointegration involves the question of whether there are one or two trends in our bivariate process. For example, do sales exhibit a diminishing trend because airplay has a downward trend? Or do airplay and sales each follow their own diffusion process which evolve separately and independently. A cointegrated relationship can be determined from the  $\Phi$  and  $\Theta$  parameters. Past research of cointegrated series has focused primarily upon vector autoregressive (VAR) models. We can determine whether the series of a VAR(1) model are cointegrated by testing whether the eigenvalues of the AR matrix are one. A recently developed procedure by Lutkepohl and Claessen (1997) for analysis of cointegrated VARMA processes suggest inverting the MA component and approximating the model with a finite VAR model.

The largest eigenvalue from Adam Ant, Blind Melon, Dink, Everclear, Richard Marx, and both of Bonnie Raitt's albums are one, indicating cointegration. The estimate the remaining eigenvalue from each of the remaining series are all in the range of .94 to .96, which again results in behavior similar to cointegration. The remaining eigenvalue from each of these models are not statistically close to one, indicating that there is one common trend, and not two separate trends for album sales and radio airplay. Therefore a consistent finding from these models is that the downward trend in sales seems and airplay are influenced by a common underlying factor, and diffusion models need to reflect the fact that the apparent diffusion of one series is not occurring independently of the others diffusion process.

This analysis demonstrates that VARMA models have shown that they provide a flexible framework

through which to model complex and dynamic relationships between multiples series without having to possess any knowledge of the relationships that may be observed beforehand. In our dataset, the VARMA models indicate dynamic relationships between sales and airplay are common, although the intensity of these relationships can vary a great deal. As demonstrated with Bonnie Raitt's *Longing in Their Hearts* album in §4 we can use these models to answer the question: Should music labels pay for radio airplay? Obviously, the answer will differ for each type of relationship observed. If airplay does not impact sales response, either lagged or contemporaneously, it would be inefficient to allot resources to radio station promotions.

Album	2 million additional GRPs		2 million fewer GRPs	
	Expected Sales Gain	Incremental Profit	Expected Sales Loss	Incremental Profit
Adam Ant: <i>Wonderful</i>	3,239	17,110	(4,530)	(45,042)
Beastie Boys: <i>Ill Communication</i>	5,055	31,639	(7,168)	(66,145)
Blind Melon: <i>Soup</i>	5,004	31,228	(5,393)	(51,945)
Charles & Eddie: <i>Chocolate Milk</i>	829	(2,172)	NA	NA
Dink: <i>Dink</i>	383	(5,734)	NA	NA
Everclear: <i>Sparkle &amp; Fade</i>	2,834	13,870	(4,419)	(44,156)
John Hiatt: <i>Walk On</i>	11,533	83,467	(19,530)	(165,038)
Richard Marx: <i>Paid Vacation</i>	949	(1,209)	(1,000)	16,800
Radiohead: <i>The Bends</i>	1,973	6,985	NA	NA
Bonnie Raitt: <i>Longing In Their Hearts</i>	4,946	30,766	(5,261)	(50,885)
Bonnie Raitt: <i>Road Tested</i>	2,976	15,012	(3,625)	(37,799)
Bob Seger: <i>It's a Mystery</i>	13,288	97,508	(14,383)	(123,861)
Smoking Popes: <i>Born to Quit</i>	753	(2,779)	(984)	(16,673)
<i>Average</i>	4,135	24,284	(6,629)	(61,834)

**Table 3.** Expected incremental gain in album sales over a 20 week period from an additional 2,000,000 GRPs during the album's launch week. The three albums with results listed as not available (NA) have fewer than 2 million GRPs already allocated to them.

To quantify the sales impact of the current airplay schedule we compute the expected incremental album sales over the next 20 weeks from an additional two million GRPs during the first week. The results from these incremental sales are given in Table 3. We assume that the marginal income generated from selling an additional album is \$8 as in our analysis from §4 of Bonnie Raitt's *Longing in Their Hearts* and the cost of the additional two million GRPs is \$8,800. Notice that on average this would be a profitable proposition for this portfolio. This finding goes directly to answering the question: Should music labels pay for radio airplay? If this sample portfolio

is representative of the music label's entire portfolio the answer is clearly *yes*.

Even though it is profitable to increase the radio airplay for the total portfolio, this does not mean that it is profitable for every album. In fact, for four out of thirteen of the albums it would not be profitable to buy the additional radio airplay. Clearly the more profitable proposition is to pay for radio airplay only for those albums that would generate positive incremental income. Moreover, given that radio airplay is in fixed supply what is probably more relevant is to consider how to reallocate radio airplay across these albums. For example, an increase in the radio airplay of one album will require the decrease in the air time of another. If we assume that this is the complete set of albums competing for radio airplay and they were all released at the same time, profitability can be increased by adding two million GRPs of additional airplay to John Hiatt and Bob Seger and reducing it by two million for Richard Marx and Smoking Popes. This would have the net effect of increasing profits by \$147,500. Notice that the expected gains of increasing the other albums would not offset the expected losses of the others. Hence, not only is increased airplay important, but perhaps just as critical is control over how airplay is allocated.

## **6 Conclusions**

We can summarize the findings of our research with three basic findings:

1. We find that it could potentially be very profitable if music labels could pay to increase radio airplay. For the thirteen albums studied in this paper we found that 2 million additional GRPs could increase the average album by 4,135 units (see Table 3). If each album has a gross profit margin of \$8 and 2 million GRPs sell for \$8,800 then these incremental sales could increase profits by \$315,700 ( $=\$24,300$  average profit per album  $\times$  13 albums), which would be a handsome return. At the same time we understand that radio airplay is a limited resource. Increasing airplay for one album will necessarily decrease the airplay that is available for other albums. Our results also indicate there is an asymmetric response to radio airplay. A decrease in airplay of 2 million GRPs could similarly reduce profits by \$803,800 ( $=\$61,800$  average profit per album  $\times$  13 albums). Perhaps a better approach is to focus on reallocating radio airplay amongst the set of 13 albums, which would increase profits by

\$147,500. The heterogeneity in sales response among music albums to radio airplay indicates that it is important to select which albums should receive increased promotional effort in the form of higher radio airplay. A related question is how should these profits be allocated amongst the channel members, we have assumed that radio stations would sell airtime at a price similar to commercial airtime. However, there is no reason to expect that this would be true. Our results indicate that there is a potential channel coordination problem between music labels and radio stations due to legal constraints over sponsored radio programming.

2. Methodologically we have shown how VARMA models can be used to answer the substantive question of whether music labels should pay for radio airplay. Although we have just focused upon a bivariate model, VARMA models can easily be generalized to higher dimensions to include other variates such as word-of-mouth or distribution measures. Using VARMA models we were able to differentiate autoregressive effects on sales (airplay) from the effects of airplay on sales (and/or sales on airplay). VARMA methods also allow us to assess the feedback relationships that have been posited. We have found that most of the series in our sample exhibit some sort of feedback relationship. Across a portfolio of albums, there is significant heterogeneity among the albums in terms of the dynamics between sales and airplay. VARMA models are able to distinguish these relationships from those with only a unidirectional causal relationship, i.e., either sales affects airplay or airplay affects sales but not both. In the feedback cases, both sales and airplay are autoregressive in nature and have a significant impact on each other. This interaction is overlooked by univariate analyses which do not recognize and cannot allow for any feedback relationships.

The multivariate analysis also results in a different interpretation of the airplay impact on sales. Any univariate model with airplay affecting sales as an explanatory variable would only allow for a one period effect. That is, radio support in a given week would boost sales for that week and have no lingering effects. However, a feedback model which simultaneously estimates a system of airplay and sales equations would let airplay have an effect on sales that would perpetuate beyond just the one period, and likewise for sales. The feedback model would acknowledge a cyclical process in which additional radio support would increase sales which would reinvigorate airplay in later periods. This implies that one additional minute of radio exposure at time  $t$  would

boost sales for that week, increasing airtime in the following weeks, and thereby impacting sales in those future weeks as well. The single decision to play a song one more time has sales effects that would linger beyond that week. This process is recognized by the feedback model but cannot be accommodated by any univariate analyses. Our analysis in §4 and §5 demonstrates that radio airplay exposure adds value in terms of incremental sales. From this analysis, music labels can quantify how much they should pay for radio airplay exposure.

3. Our modeling approach points out a limitation of multivariate diffusion models when those diffusion processes share a common trend as apparently is the case with music albums. The Fourt and Woodlock model (Moe and Fader 2001) is quite susceptible to spurious regression problems. Bass and Srinivasan (2002) have also noted that some forms of the Bass model are quite susceptible to spurious regression when the covariates are smoothly trending. We have shown that the VARMA modeling approach can easily overcome this problem, even when it is not known a priori whether there is a common trend or independent trends. It also has the advantage of avoiding overdifferencing to stabilize the process. Furthermore, we have shown that under certain circumstances the forecasts of our proposed VARMA model with logarithmic transformations and diffusion models are the same.

There are also several limitations that are inherent in our study. First, we make no claims that our sample is representative of music albums. Hence, a larger dataset is necessary to more definitely answer the question of the value of radio airplay at an industry level. Although our approach could be readily implemented.

Another caveat we offer about stochastically modeling radio airplay is that currently, radio airplay is not under the control of music labels. If music labels begin to purchase time on radio stations, radio airplay may become deterministic instead of stochastic. In that case, our present model of radio airplay may be substantially altered and may no longer be valid. The processes and the fundamental structure of the industry may change. Currently, consumers value the gatekeeping role of radio stations and may respond to the radio exposure of an album. If music labels begin to control radio content, consumers may not respond to radio exposure in the same way. Airplay becomes more similar to advertising, and consumers have learned to critically evaluate advertisements as opposed to taking the message at face value.

A further complicating issue related to the pay-for-play radio airplay is how will radio stations modify their current radio airplay schedules. If an advertiser begins to purchase radio airplay, will radio's reduce the amount of airtime that they give to the albums being played in the pay-for-play airtime slots? This could seriously diminish the effectiveness of the pay-for-play programming. Again our dataset does not allow to answer these questions, and these answers to these questions may not be known until pay-for-play is put into practice. This also points out the symbiotic relationship that has developed between radio stations and music labels. Not only do music labels benefit from the free exposure from radio airplay, but radio stations benefit from the promotions that music labels spend upon their artists and the low royalty fees that radio stations are charged for playing albums. If pay-for-play begins to dominate airplay schedules, then music labels may substantially increase the royalty fees that radio stations must pay to air music, and the structure of the radio industry may be dramatically altered. Unfortunately these questions can only be addressed with experimental evidence.

A potential empirical limitation of our technique is that we have used all data points available when estimating the relationship between radio airplay and music album sales. Hence, our ability to assess which albums would gain from additional airplay is post hoc. Many music managers may wish to make pre-launch sales forecasts (Lee, Boatwright, and Kamakura 2002, Moe and Fader 2002). Both Lee et al. (2002) and Moe and Fader (2002) recommend hierarchical Bayesian procedures to make prelaunch predictions about parameters. We agree that a hierarchical Bayesian or random coefficients framework would be beneficial. Such an approach could readily be applied to our model. However, our small dataset limits our ability to adequately illustrate this methodology. For example, with 13 series the likelihood would be improper for the covariance matrix meaning that any specification would be highly dependent upon the prior.

Not only is there a desire to predict the relationship between sales and airplay before the album launch, but managers may also want to manipulate radio airplay early in the life of the album to better estimate the relationship between sales and airplay. For example, music labels may design a pulsing strategy for radio airplay to induce artificial variation in the process. This may increase the statistical information in the process and result in more precise parameter estimates. We believe experimentation may be an interesting area for future research.

## References

- Ansley, Craig F. (1979), "An Algorithm for the Exact Likelihood of a Mixed Autoregressive-Moving Average Process", *Biometrika*, 66 (1), 59:65.
- Arino, M.A. and P.H. Franses (2000), "Forecasting the levels of vector autoregressive log-transformed time series", *International Journal of Forecasting* 16 (1): 111-116.
- Bass, Frank M. (1969), "A New Product Growth Model for Consumer Durables", *Management Science*, 15 (January), 215-227.
- Billboard (1999), "Politicians call for FCC to examine label/radio deals", Jan 30, Vol 111, Issue 5, pg. 6.
- Box, G. E. P., Jenkins, G. M., and G.C. Reinsel, (1994), *Time Series Analysis: Forecasting and Control*, (3rd ed.), Englewood Cliffs, NJ: Prentice-Hall.
- Bradlow, E. T., and P. S. Fader (2001), "A Bayesian lifetime model for the 'Hot 100' Billboard Songs", *Journal of the American Statistical Association*, 96 (454): 368-381.
- Dekimpe, Marnik G. and Dominique M. Hanssens (1995a), "Empirical Generalizations About Market Evolution and Stationarity", *Marketing Science*, 14 (2, 2), G109-21.
- Dekimpe, Marnik G. and Dominique M. Hanssens (1995b), "The Persistence of Marketing Effects on Sales", *Marketing Science*, 14 (1), 1-21.
- Dekimpe, Marnik G. and Dominique M. Hanssens (1999), "Sustained Spending and Persistent Response: A New Look at Long-Term Marketing Profitability", *Journal of Marketing Research*, 36, 397-412.
- Dekimpe, Marnik G., Dominique M. Hanssens, and Jorge M. Silva-Risso (1999), "Long-Run Effects of Price Promotions in Scanner Markets", *Journal of Econometrics*, 89 (March/April), 269-91.
- Dekimpe, Marnik G., Jan-Benedict E.M. Steenkamp, Martin Mellens, and Piet Vanden Abeele (1997), "Decline and Variability in Brand Loyalty", *International Journal of Research in Marketing* 14 (5), 405-20.
- Dickey, David A. and Wayne A. Fuller (1979), "Distribution of the Estimators for Autoregressive Time Series with a Unit Root", *Journal of the American Statistical Association*, 74 (June), 427-431.
- Engle, R. F. and C.W.J. Granger (1987), "Cointegration and Error Correction - Representation, Estimation, and Testing", *Econometrica*, 55 (2): 251-276.
- Fourt, Louis A. and Joseph W. Woodlock (1960), "Early Prediction of Market Success for New Grocery Products", *Journal of Marketing* 25 (October), 31-38.
- Franses, P. H. (1994), "Modeling New Product Sales: An Application of Cointegration Analysis", *International Journal of Marketing* 11, 491-502.
- Funkhauser, G.R. (1973), "The Issues of the Sixties: An Exploratory Study in the Dynamics of Public Opinion,"

*Public Opinion Quarterly*, 37, 62-75.

- Granger, C.W.J. and P. Newbold (1986), *Forecasting Economic Time Series*, New York: Academic Press.
- Hannan, E. J., and Deistler, M. (1988), *The Statistical Theory of Linear Systems*, New York: Wiley.
- Hanssens, D.M., L.J. Parsons, and R.L. Schultz (1990), *Marketing Response Models: Econometric and Time Series Research*, Kluwer Academic Publishers, Boston.
- Hilmer, Steven C. And George C. Tiao (1979), "Likelihood Function of Stationary Multiple Autoregressive Moving Average Models," *Journal of the American Statistical Association*, 74 (September), 652:659.
- Holland, Kelley (1997), "Record Companies in Cahoots?", *Business Week*, New York.
- Iyengar, S. and D.R. Kinder (1987), *News that Matters: Agenda-Setting and Priming in a Television Age*, Chicago: University of Chicago Press.
- Kalbfleisch, John D. And Ross L. Prentice (1980), "Structure of Repeat Buying for New Packaged Goods," *Journal of Marketing Research*, 17 (August), 316:322.
- Lawless, J.F. (1982), *Statistical Models and Methods for Lifetime Data*, New York: John Wiley & Sons, Inc.
- Lee, Jonathan, Peter Boatwright, and Wagner A. Kamakura (2002), "A Bayesian Model of Empirical Generalizations for Pre-launch Sales Forecasting of Recorded Music", Working Paper, Indiana University, Kelley School of Business.
- Ljung, G. M., and G. E. P. Box (1978), "On a Measure of Lack of Fit in Time Series Models", *Biometrika*, 65, 297-303.
- Lutkepohl, Helmut and Holger Claessen (1997), "Analysis of Cointegrated VARMA Processes", *Journal of Econometrics*, 80 (2, October), 223-39.
- Mahajan, Vijay, Eitan Muller, and Frank M. Bass (1990), "New Product Diffusion Models in Marketing", *Journal of Marketing* 54, 1, 1-26.
- McCombs, Maxwell E. and Donald L. Shaw (1972), "The Agenda-Setting Function of Mass Media," *Public Opinion Quarterly*, 36, 176:187.
- Moe, Wendy and Peter Fader (2001), "Modeling Hedonic Portfolio Products: A Joint Segmentation Analysis of Music Compact Disc Sales", *Journal of Marketing Research*, 38 (3): 376-385.
- Moe, Wendy and Peter Fader (2002), "Using Advance Purchase Orders to Forecast New Product Sales", *Marketing Science*, forthcoming.
- Montgomery, A. L., V. Zarnowitz, R.S. Tsay, and G.C. Tiao (1998), "Forecasting the U.S. Unemployment Rate", *Journal of the American Statistical Association*, 93 (442): 478-493.
- New York Times (2002), "Charges of Payola Over Radio Music", May 25, B7.
- Radas, Sonja and Steven M. Shugan (1998), "Seasonal Marketing and Timing New Product Introductions," *Journal*

- of Marketing Research*, 35 (August), 296:315.
- Reinsel, Gregory C. (1995), *Elements of Multivariate Time Series Analysis*, New York: Springer Verlag.
- Sawhney, Mohanbir S. and Jehoshua Eliashberg (1996), "A Parsimonious Model for Forecasting Gross Box Office Revenues", *Marketing Science*, 15:2, 113-131.
- Tiao, G. C., and Box, G. E. P. (1981), "Modeling Multiple Time Series With Applications", *Journal of the American Statistical Association*, 76, 802-816.
- Tiao, G. C., and Tsay, R. S. (1989), "Model Specification in Multivariate Time Series" (with discussion), *Journal of the Royal Statistical Society, Set. B*, 51, 157-213.
- Tsay, R. S., and Tiao, G. C. (1984), "Consistent Estimates of Autoregressive Parameters and Extended Sample Autocorrelation Function for Stationary and Nonstationary ARMA Models", *Journal of the American Statistical Association*, 79,84-96.
- Tsay, R. S. (1988), "Outliers, Level Shifts, and Variance Changes in Time Series", *Journal of Forecasting* 7 (1): 1-20.
- Wall Street Journal (1998), "Radio's New Spin on an Oldie: Pay-for-Play," March 16, B1.
- Wall Street Journal (2002), "Music Labels Say It Costs Too Much to Get Songs on Radio", June 10, B1.