IT-Enabled Broadcasting in Social Media: An Empirical Study of Artists' Activities and Music Sales

Hailiang Chen, Prabuddha De, and Yu Jeffrey Hu

Krannert School of Management
Purdue University, West Lafayette, IN

December 2011

Abstract

With the emergence of social media and Web 2.0, broadcasting in the online environment has evolved into a new form of marketing due to the much broader reach enabled by information technology. This paper examines the organizational use of social media – specifically, artist-generated content – and quantifies the impact of artists' broadcasting activities on a leading social media site for music, MySpace, on music sales. We employ a panel vector autoregression (PVAR) model that allows us to treat our main variables as endogenous and to investigate the inter-relationship between broadcasting promotions and music sales. We find empirical evidence that the market reach measured by the size of an artist's virtual friend network plays a crucial role in moderating the effect of IT-enabled broadcasting activities on sales. For the artists with many friends, broadcasting activities on MySpace have a significant impact on music sales; this, however, is not true for the artists with only a few friends. Our results continue to hold after controlling for artists' popularity, changes in their network sizes over time, and the impact of user-generated content.

Keywords: Broadcasting, Social media, Internet marketing, Music sales, Panel VAR model.

1. Introduction

Television and radio programs have long been used as the traditional channel to reach potential audience for marketing purposes. In the past few years, social media websites such as Facebook, MySpace, and Twitter have become an important component of people's daily life. IT-based broadcasting has been widely adopted by companies to market their products on social networking and microblogging sites. For instance, a *Wired* article (Silver 2009) notes, "Twitter, at first a place to tell everyone what you ate for breakfast, is now a place to promote yourself, your company or your product" (paragraph 13). Companies can broadcast information regarding their products to their Facebook "friends" or embed marketing messages in their "tweets" (a.k.a., status updates) to their Twitter "followers". Because these social media sites reach millions of users who often spend hours and hours on these sites, they can serve as a very powerful marketing channel for companies. For instance, according to a news report in *InformationWeek*, from 2007 to June 2009, Dell has generated a total of $2 million in direct sales of refurbished systems and $1 million in indirect sales of new systems from their Twitter presence @DellOutlet (Gonsalves 2009). Twitter as a microblogging service is becoming such an effective marketing tool that advertising startups like Ad.ly and SponsoredTweets have signed up thousands of Twitter users and paid them up to $10,000 per tweet for sending advertising messages to their Twitter followers (Learmonth 2010).

MySpace Music, known as the most popular social networking site for music, is a prominent example of how information technology can enable musicians to reach a huge audience at a relatively low cost. Over eight million artists and bands have set up their profiles on MySpace Music (Owyang 2008). F. Vincent, the author of the book titled "*MySpace for Musicians*" (Vincent 2010), puts it this way: "with so many potential pairs of eyes and ears at your fingertips, it is becoming a necessity for any musical artist – whether signed and selling or unsigned and hopeful – to have a profile on MySpace" (Vincent 2006, paragraph 2). MySpace Music offers many great free tools, such as bulletins and activity streams, and focuses on building a community of artists and music fans. Artists and bands can upload songs, show music videos, communicate with fans, and even sell MP3 downloads through the website. Moreover,

there are different kinds of commercial software designed to help musicians promote their music on MySpace Music (e.g., MySpace Friend Pro: http://myspacefriendspro.com).

Despite the abundant anecdotal evidence indicating that marketing on social media sites can be very effective in driving up product sales (e.g., Baker 2006, Gonsalves 2009, Vincent 2010), there is a lack of academic research that examines the dynamic relationship between companies' broadcasting activities and their product sales. Do companies' marketing activities on social media sites really have an impact on their product sales? How does such an impact vary across companies? What are the key factors that determine the success of companies' marketing activities on social media sites? These questions remain largely unanswered to date.

To answer the above questions, we use artists' marketing activities on MySpace Music as an example. We collect artists' activity stream data from MySpace Music and combine it with the sales rank data from Amazon. We address the potential endogeneity problem by employing a panel vector autoregression (PVAR) model (Holtz-Eakin et al. 1988) estimated by the Generalized Method of Moments (GMM) (Binder et al. 2005), and study the dynamic relationship between artists' activities and music sales. The activity stream data analyzed in this paper is quite unique – it is the time series data on artists' marketing messages broadcasted to their friend networks. The PVAR model applied to this data enables us to treat all of these time series variables as endogenous and to avoid making restrictive assumptions.

Our results suggest that artists' marketing efforts on MySpace Music do have a significant effect on their music sales. We also examine how market reach influences this effect across different artists, by dividing the sample into two equal-size groups according to the number of friends each artist has at the beginning of our study period. Interestingly, we find that the significant impact of MySpace activities on music sales applies only to those artists who have a relatively large friend network. For the artists with a relatively small friend network, this impact is statistically insignificant. In addition, we use the consumer search volume index from Google Trends as an indirect measure of artist popularity and, thus, are able to account for time-varying individual characteristics in dynamic panel models.

We conduct several additional analyses. We address the potential endogeneity issue of market reach

by incorporating the number of new friends in each period into the model. As different marketing resources are available to major labels and indie artists, we also compare these different types of artists in this context. To account for the impact of user-generated content (UGC) on sales, we also collect customer reviews data from Amazon and incorporate it into the model. Last, we test the case when quadratic terms are included in the regression for broadcasting activities, and find that the relationship between these activities and sales ranks can indeed be nonlinear.

This paper makes a number of contributions to the IS literature. First, we use a novel dataset on artist-generated content[1], rather than user-generated content (UGC). To date, IS research has studied consumers' activities on the Internet (e.g., customer reviews, ratings, and blogs) and the impact of such activities on product sales (see, for example, Dellarocas 2003, Chevalier and Mayzlin 2006, Chen et al. 2007, Forman et al. 2008, Dewan and Ramaprasad 2009, and Zhu and Zhang 2010). To the best of our knowledge, there is no existing research that has examined how companies' activities (or activities of the sellers) in the online environment can impact their product sales. Our paper bridges this gap in the literature by studying the organizational use of social media and providing a vivid example on how artists (and record companies) can promote their music on a social networking site. Second, the paper introduces a relatively new econometric methodology – the PVAR model, estimated using a GMM estimator – to the literature on social media. Because endogeneity is often an important concern for this type of research, the GMM estimation method employs lagged dependent variables as instruments and, thus, can be particularly useful in identifying the relationships, if any, in both directions. This paper also has important managerial implications. In light of our findings, companies should enhance their efforts to utilize social media as a marketing channel. In addition, they should pay close attention to their potential market reach

[1] Although artist-generated content and user-generated content share some common characteristics, there are several fundamental differences between them. First, the creator of artist-generated content is the producer and seller of the product, songs or albums in our context, while the creator of user-generated content is the buyer and consumer of the product. Second, artists can make decisions on what kind of content to be generated and how. User-generated content is, in general, less likely to be directly influenced by artists or companies who sell the product. Third, the mechanisms through which each makes an impact are different. Artist-generated content, such as the artist page, is an "official" channel for artists to publicize their profile information, new albums, recent activities, etc. to music fans. Given the artist's friend network, a posting (bulletin) or an action (friend update/activity stream) by her is broadcasted to potentially all of her friends in the network at the same time, which can be regarded as a "global" effect. However, user-generated content, such as a user posting, tends to have a "local" effect, which only renders an impact on a subgroup of friends in the network, the size of which depends on the influencing power of the posting (just like some consumer reviews are read by many people, while some receive almost no attention).

when planning marketing activities on social media sites – the size of the friend network or follower network is a key factor that determines the success of companies' social-media-based marketing efforts. Finally, companies should be careful about not over-marketing on social media sites.

The remainder of this paper is organized as follows. In Section 2, we review recent IS studies that examine the relationship between consumers' online activities and sales, as well as the relevant literature on the impact of marketing effects on sales. Section 3 first discusses how MySpace Music helps artists promote their music, and then gives a brief overview of the data. We present our empirical analysis and main results in Section 4. The robustness checks are carried out in Section 5. Section 6 concludes the paper and suggests some directions for future research.

2. Literature Review

Many papers have examined whether online word-of-mouth or user-generated content, such as consumers' reviews, ratings, and blogs, have an impact on sales. The earlier studies (see, for example, Dellarocas 2003, Chevalier and Mayzlin 2006) have found a significant relationship between online consumer reviews and product sales, such as more positive reviews lead to better sales and the impact of one-star reviews (most negative) is greater than the impact of five-star reviews (most positive). The more recent studies have taken more nuanced approaches toward examining such a relationship. For instance, Chen et al. (2007) study how the number of helpful votes on reviews and the reputation of reviewers influence the relationship between book ratings and book sales. Notably, the reviews with more helpful votes have a stronger impact on sales than the others. Forman et al. (2008) investigate the role of reviewer identity disclosure in affecting the relationship between consumer reviews and sales. Consumers tend to rate reviews with identity information of the reviewers more favorably, and prevelant disclosure of reviewer identity is associated with higher subsequent product sales. Li and Hitt (2008) study the time series of consumer reviews and find that early reviews are subject to self-selection biases and can have an influence on the long-term consumer behavior. Zhu and Zhang (2010) consider how product and consumer characteristics moderate the relationship between consumer reviews and product sales. They

find that the impact of reviews is stronger for less popular video games and games whose players have more Internet experience. In addition, researchers have started to pay attention to how consumer blogs can drive product sales. For instance, Dewan and Ramaprasad (2009) use the Granger Causality tests (Granger 1969) and two-stage least squares to study the causal relationship between blog buzz and music sales. All of these existing studies utilize data on user-generated content and address the impact of consumer behaviors on product sales.

Our study is related to the existing literature that studies how online content can influence product sales. However, it also differs greatly from this literature because we examine online content from an angle that has been largely ignored and study companies' activities (artist-generated content in this case) rather than consumers' activities. A few studies such as Dellarocas (2006) and Godes et al. (2005) have examined how firms can play a role in manipulating or controlling the user-generated content and online word of mouth. Dellarocas (2006) develops an analytical model to study the economic impact of firms' posting anonymous messages to Internet opinion forums in disguise as customers. Godes et al. (2005) summarizes four roles that firms can play in managing the social interactions among consumers: observer, moderator, mediator, and participant. Although firms' activities are considered in these studies, they are still conducted anonymously for the purpose of creating user-generated content. In this study, we are taking a step further and study how firms can directly generate online content in their own names to boost sales. In this regard, to the best of our knowledge, no other paper has tried to quantify the value of social media marketing[2].

This paper also draws upon the marketing literature that studies the effect of traditional advertising on sales. According to Dekimpe and Hanssens (1995), a marketing action can affect the sales performance of a brand or a firm in six ways: contemporaneous effects, carry-over effects, purchase reinforcement, feedback effects, firm-specific decision rules, and competitive reactions. In general, advertising often has an immediate effect on sales. A more subtle question to marketing managers is how long the cumulative

[2] A few recent studies also examine the value of social media sites in different contexts, such as Lin et al. (2009), which quantifies the value of social networks in online peer-to-peer lending markets.

effect of advertising persists. Early studies (e.g., Givon and Horsky 1990) in marketing have documented that the effect of advertising in one time period may be carried over, at least partially, into future periods. It can be argued that consumers remember past advertising messages, but this "goodwill" toward the advertised brand gradually decays because of forgetting and competitive advertising. Givon and Horsky (1990) and Horsky and Simon (1983) propose that advertising can also indirectly affect sales through purchase reinforcement. Advertising can encourage consumers to try a new product or a product they have not purchased for a long time; then, if they like their experience, they may purchase it again in the future. In addition, Simester et al. (2009) show that current advertising affects future sales through the two competing effects of brand-switching and inter-temporal substitution. These previous findings suggest that it is important to use appropriate models to capture the lagged effects when the effects of marketing activities on sales are studied.

Prior studies (see, for example, Bass 1969 and Hanssens 1980) also conclude that advertising efforts should not be treated as exogenous as they may be influenced by the current and past performance of sales. This phenomenon essentially calls for a research method that can treat both advertising and sales as endogenous variables. Failure to do this may result in a bias in estimation. In addition, competitive reactions could also have a major impact on the effectiveness of advertising. In the short run, marketing actions may prompt a positive sales response, but the long-run effect could be negligible depending on the nature of competitive reactions (Hanssens 1980).

We apply the insights from the studies mentioned above to the social media context, and employ a time series model that allows us to investigate the effects of artists' online marketing actions on sales and, at the same time, address the endogeneity problem. The model we use is a variant of the vector autoregression (VAR) model developed in the seminal work by Sims (1980). The VAR model has been widely used to analyze time series in macroeconomics, finance, and other fields (see, for example, Bessler 1984, Sims 1992, McCarty and Schmidt 1997). In the social networking context, Trusov et al. (2009) employ a VAR model to study the impact of word-of-mouth referrals and traditional marketing on the

number of sign-ups at a major social networking site[3]. In principle, VAR models can also be applied to study individual-level data. For example, Holtz-Eakin et al. (1988) is the first paper to estimate a Panel VAR model, revealing the dynamic relationship between individuals' hours of work and their wages. According to Pauwels et al. (2004), time series econometric models, such as VAR-based models, have made considerable contributions to the marketing literature.

3. Data

3.1 How Artists Promote Themselves on MySpace Music?

As a social networking site, MySpace allows bands and artists to set up profile pages that are different from the ones for normal users. This aims to provide a professional platform for artists to connect with their fans besides building their personal networks. On a typical artist profile page[4], aside from the basic profile information about the artists, there is usually a music player that allows visitors to play the songs in a playlist specified by the artists. The artists can also list the schedules of their upcoming shows or concert performances. Many artists also leave space for some recent blog entries and upload a few videos. Finally, a top friend list and some recent comments by friends appear at the end of the profile page.

The information provided on the profile page serves as only a starting point of online music promotion. MySpace Music also provides other tools to exploit the benefits of social networks. One of these tools is the bulletin board. Bulletins are posts that everyone in the friend list can see. The commercial software MySpace Friend Pro introduces its bulletin poster feature starting with the following, "Once you have enough friends on myspace.com, posting bulletins can be very very effective and unlike commenting and messaging you do not need to send individual bulletins to all your friends…People read bulletins as long as the topic is catchy and somewhat relevant."[5] Artists can post bulletins on a timely basis to tell their friends what they are doing right now, ask them to listen to their new songs, and invite them to post comments. Once a user becomes an artist's friend, then that user will automatically receive

[3] The authors (Trusov et al. 2009) do not disclose the site name in their paper.
[4] For an example of an artist profile page, please visit http://www.myspace.com/onerepublic.
[5] http://myspacefriendspro.com/bulletin.php.

the artist's bulletins and be able to see them until they expire in 10 days.

Another major promotional tool is the friend updates that inform music fans of the artists' latest MySpace activities (MySpace recently renamed "friend updates" to "activity stream"). The benefit of the friend updates lies in the fact that any activities by the artists are automatically updated to their friends if they choose to receive these updates. MySpace, like other social networks, also periodically sends emails to users to notify them of the recent friend updates. This feature, thus, enables artists to spread out a message to their fans very quickly and efficiently in online social networks, especially considering that the number of friends can easily reach hundreds of thousands for popular artists. There are several different types of activities that can show up in an activity stream. Among them, "add new blogs", "add new photos", "upload new track", and "add new friends" are the most common ones. By default, MySpace users can see their friends' updates in the past month after they login.

To summarize, the fundamental objective of promoting music is to keep music fans interested in them over time. To achieve this, MySpace Music provides the bulletin board and activity stream to automatically spread artists' messages across the network of their friends. Previous research has not studied the marketing impact of these automatic news feeds in virtual social networks. Among other things, this research aims to examine the roles of these emerging promotional tools.

3.2 Data Description

We use two data sources in this research, namely, the MySpace activity stream data and the Amazon sales rank data. MySpace is in partnership with Amazon allowing users to directly buy MP3 albums and songs from Amazon by following links on the MySpace pages. Amazon tracks the sales ranks of music artists that sell music (either CDs or digital albums/songs) on Amazon.com[6]. The activity stream on MySpace Music records all the promotional activities by artists, which are then broadcasted to their' friends. To collect this information, we subscribed to a selected sample of artists as their friends and retrieved the daily updates of their MySpace activities. We also counted the number of friends for each

[6] The rank data we collected is at the artist level, not the album or song level. Artist sales rank is the rank of Amazon.com sales from an artist's all CDs and digital albums/songs.

artist every day.

In this study, we use MySpace over other social networking or microblogging sites primarily for the following reason. MySpace operates an entire division (MySpace Music) dedicated to serving music artists and fans, which can not be found elsewhere. Millions of artists are attracted to this website for the basic reason that MySpace is the most popular – and perhaps coolest – website to promote their music. In this regard, marketing on MySpace for artists has become a large scale phenomenon (Vincent 2006, 2010). In addition, to better compete with other sites, MySpace aims to become an online community around entertainment and focuses on areas such as music, video, and games (Steel 2009). Given these observations, we believe that MySpace Music is the ideal choice to study how artist-generated content in social media can influence sales in the music industry.

Our sales rank data comes from the second largest music retailer, Amazon[7]. Unlike other top players such as iTunes and Wal-Mart, Amazon pursues a balanced strategy of selling both physical and digital music. We assess whether the Amazon sales ranks are correlated with the overall music market sales by comparing the Amazon music charts with the well-known Billboard charts. We retrieved the Billboard Hot 100 Songs and Amazon Top MP3 Songs on the same day and calculated the correlation between the sales ranks of the songs that appear on both lists. These two sales ranks turned out to be highly correlated, with a coefficient of 0.758. Thus, we conclude that artists' Amazon sales ranks are representative of their overall music sales ranks.

Another important problem that needs to be dealt with is how to select a sample of artists. Since there are over eight million artist profiles on MySpace Music and the list of all artists was not available to us, we selected the sample of artists through the lists of Amazon's daily Top MP3 Songs for all genres and 22 different individual genres (23 lists per day)[8]. We used the top song list instead of the top artist list to avoid selecting only the popular artists[9]. We selected three days (August 30, September 10, and

[7] According to a press release from the NPD group (2010), Amazon had 12 percent of the overall music market (including both the CD and digital formats) in the first quarter of 2010.
[8] The list of these 22 genres is available at the following link: http://www.amazon.com/gp/bestsellers/dmusic/digital-music-track.
[9] The top artist list is based on the total sales of all digital music; usually it is very difficult for obscure artists to get into this list, but it is much easier for them to show up in the top song list. In light of this, this we decided to use the top song list.

September 20) in 2008 and downloaded 23 lists per day. Originally, 5,146 artists appeared in the lists, and both major labels and indie (i.e., independent or unsigned) artists were included. A program was written to search these artists' information on MySpace Music and find their corresponding profile pages automatically. To ensure accuracy, the artist name and the titles of uploaded songs on the profile page were used to match with the information on Amazon. There could be multiple MySpace profiles for the same artist name, but we took a conservative approach and selected the artists that had only one match on MySpace. Overall 1,604 artists' profile pages were exactly matched. Next, we tried to send a friend request to each of these 1,604 artists, and until October 18, 2008, successfully subscribed to 631 artists[10]. The actual data collection started on October 19, 2008 and ended on May 30, 2009; thus, we have the data for 32 weeks in total.

Table 1: Descriptive Statistics

Variable	# Obs.	Mean	Std. Dev.	Median	Min	Max
SalesRank	20,192	5,261	9,232	2,116	2	111,093
Bulletins	20,192	0.67	1.94	0	0	48
FriendUpdates	20,192	0.57	1.06	0	0	12
NetworkSize	631	105,834	169,922	40,485	55	1,133,894

The main time series variables constructed for our analyses are artists' weekly average sales ranks, the number of bulletins, and the number of friend updates in each week. To measure the market reach of each artist, we define the number of friends at the beginning of the study period as *NetworkSize* and keep track of the number of new friends for each artist over time. Table 1 shows the descriptive statistics of these variables for all artists. The number of observations for the first three variables is 20,192, and the panel data set is strongly balanced. The sales ranks for artists range from 2 to 111,093, indicating that both popular and unpopular artists are included in the sample. We also find that a wide range of genres are represented in the sample. The average number of bulletins in each week is only 0.67, but the maximum number could reach 48 (i.e., in the extreme case, artists post almost 7 bulletins per day). Similarly, the average number of friend updates in each week is also small (0.56), but the maximum could reach 12. The

[10] If the friend request is accepted by the artist, then the person initiating the request becomes a friend of that artist and is able to observe his or her promotional activities.

median network size is 40,485. The average is even larger (105,834), and the maximum number of friends is more than one million in our sample.

4. Empirical Analysis

We estimate a panel vector autoregression (PVAR) model that examines the dynamic interactions among artists' sales ranks and promotional actions. To avoid making any assumptions that may be unreasonable, we adopt the reduced form of VAR models in which each dependent variable is endogenous and is a linear function of its own past values, the past values of all other dependent variables, a set of exogenous variables, and an error term[11]. The panel structure of the data allows us to control for unobserved individual heterogeneity. Toward that end, we introduce individual-specific effects to the model. Thus, we specify the following baseline model:

$$\mathbf{y}_{it} = \begin{pmatrix} \ln(SalesRank_{it}) \\ \ln(1+Bulletins_{it}) \\ \ln(1+FriendUpdates_{it}) \end{pmatrix} = \sum_{s=1}^{p} \mathbf{\Phi}_s \cdot \begin{pmatrix} \ln(SalesRank_{it-s}) \\ \ln(1+Bulletins_{it-s}) \\ \ln(1+FriendUpdates_{it-s}) \end{pmatrix} + \begin{pmatrix} \delta_{1t} \\ \delta_{2t} \\ \delta_{3t} \end{pmatrix} + \begin{pmatrix} f_{1i} \\ f_{2i} \\ f_{3i} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1it} \\ \varepsilon_{2it} \\ \varepsilon_{3it} \end{pmatrix}, \quad (1)$$

where $\mathbf{y}_{it} = (\ln(SalesRank_{it}), \ln(1+Bulletins_{it}), \ln(1+FriendUpdates_{it}))'$ is a three-element column vector for artist *i* at time *t*, containing the log transformation of the dependent variables[12]; $\mathbf{\Phi}$'s are 3×3 matrices of slope coefficients for endogenous variables; $\boldsymbol{\delta}_t = (\delta_{1t}, \delta_{2t}, \delta_{3t})'$ is a column vector of time dummies that control for any time effects such as seasonality; $\mathbf{f}_i = (f_{1i}, f_{2i}, f_{3i})'$ is a column vector of unobserved individual effects, characterizing artists' time-invariant attributes; $\boldsymbol{\varepsilon}_{it} = (\varepsilon_{1it}, \varepsilon_{2it}, \varepsilon_{3it})'$ is a three-element vector of errors, satisfying the assumption that $E(\varepsilon_{mit}) = E(\varepsilon_{mit}\varepsilon_{mis}) = 0$ for $m = 1,2,3$ and $t \neq s$, i.e., we assume that there is a lack of serial correlation; and *p* is the number of lags.

4.1 Model Identification and Estimation

Since the lagged dependent variables $\mathbf{y}_{i,t\text{-}1}, \mathbf{y}_{i,t\text{-}2}, ..., \mathbf{y}_{i,t\text{-}p}$ in equation (1) are correlated with the

[11] Recursive or structural forms are possible, but they require one to impose restrictions on causality or contemporaneous relationships between variables (Stock and Watson 2001).

[12] As these three variables have very different means and standard deviations, log transformation is used to improve model fit. Since *Bulletins* and *FriendUpdates* have zero observations, we add 1 to them before the log transformation. We have also tested different numbers such as 0.5 and 0.1, and the results remain the same.

average error term $\overline{\varepsilon}_i$ in the within-group estimator (i.e., the least-squares estimator after subtracting the individual means of the observations), the within-group estimator for the fixed effects model will be biased for this type of dynamic panel model (Nickell 1981, Arellano 2003). We estimate the model by the Generalized Method of Moments (GMM) (see Hansen 1982, Hamilton 1994, or Hayashi 2000 for a detailed review of this method). A consistent GMM estimator that uses lagged dependent variables and lagged differences of dependent variables as instruments has been developed in the econometrics literature. Estimation methods for single equations in dynamic panel models are discussed in a series of papers such as Arellano and Bond (1991), Ahn and Schmidt (1995, 1997), Arellano and Bover (1995), Blundell and Bond (1998), etc. Binder et al. (2005) summarize the standard and extended GMM estimation procedures for Panel VAR models with only one lag. The difference between these two procedures is that the extended GMM estimator utilizes more instruments than the standard GMM estimator by imposing the additional assumption that changes in the instrument variables are uncorrelated with the fixed effects. The standard GMM estimation method, on the other hand, requires only one assumption, namely, there is a lack of serial correlation in the error terms. In this paper, we estimate the proposed Panel VAR model using the standard GMM estimator following Binder et al. (2005). It is also shown in Binder et al. (2005) that the standard GMM estimator can break down if the model contains unit roots. We examine the stationarity of our panel dataset using the Levin, Lin, and Chu (2002) and Harris and Tzavalis (1999) tests and find that there is no unit root and all the panels are stationary.

To apply the standard GMM estimator, we first specify the model with a reasonably long length of lags and conduct an initial specification test using the Sargan test (Sargan 1958, 1959). After finding an accepted specification, we carry out the lag selection procedure through the Sargan difference test to select a lag length that provides the best balance of accuracy and efficiency. To address the concern that too many instruments may be used, one can again use the Sargan difference test to test the validity of subsets of instruments for a chosen lag. The details of this entire process and the derivation of the

standard GMM estimator for PVAR models with multiple lags are provided in the appendix[13].

Table 2: Panel VAR Coefficient Estimates (N=631, T=32)

Independent Variable	Dependent Variable		
	ln(SalesRank)	ln(1+Bulletins)	ln(1+FriendUpdates)
$\ln(\text{SalesRank}_{i(t-1)})$	0.549**	-0.031*	-0.008
	(0.014)	(0.014)	(0.014)
$\ln(1+\text{Bulletins}_{i(t-1)})$	-0.040**	0.225**	0.037**
	(0.009)	(0.010)	(0.010)
$\ln(1+\text{FriendUpdates}_{i(t-1)})$	-0.013	-0.0003	0.103**
	(0.008)	(0.009)	(0.009)

Notes: numbers in parentheses are standard errors; ** and * denote significance at 1% and 5%, respectively.

As shown in the appendix, we find that one lag is sufficient for all three equations in the Panel VAR model. The estimation results for one lag are presented in Table 2. Our main goal is to examine the relationship between artists' marketing actions and their sales ranks; we are less interested in the inter-relationship between the bulletins and friend updates variables. We first look at how MySpace promotions affect artists' music sales. The coefficient for *ln(1+Bulletins)* at lag 1 in the *ln(SalesRank)* equation is negative (-0.040) and statistically significant at the 1% level, indicating that the dependent variable, sales rank, will decrease (i.e., sales will go up) next week when the number of bulletins in this week increases. However, the coefficient for *ln(1+FriendUpdates)* is negative and insignificant at the 5% level. From an economic perspective, a 20% increase in the number of bulletins (more accurately, the number of bulletins plus one) this week is associated with a 0.8% decrease in the sales rank in the following week. We also try to estimate these impacts in terms of sales volumes, thereby quantifying the economic value of broadcasting promotions in social media. Without the actual sales data, we resort to the news and press releases from Nielsen SoundScan and use the method adopted by Brynjolfsson et al. (2003) to translate sales ranks into sales units. They fit sales and the sales rank to the following regression relationship:

$$\ln(Sales) = \alpha + \beta \ln(Rank) + \varepsilon \,, \qquad (2)$$

where α determines the location of the sales curve and β determines the shape of the curve. To illustrate the idea, we focus only on the weekly digital track sales of an artist. We assume the β

[13] To the best of our knowledge, there are no commercial statistical or econometric software packages that implement the estimation functions for this type of model. We have, therefore, written a Matlab program to estimate our model.

coefficient to be -0.7383, same as the shape parameter estimated by Chellappa and Chen (2008) using the Billboard Hot Digital Tracks chart, and then obtain an estimate of 12.6 for α according to the sales data published in the Nielsen Company 2009 Year-End Music Industry Report (2010). These estimates indicate that the topmost artist has a digital track sale of 296,559 units per week and that an artist with a rank of 1,000 has a digital track sale of 1,808 units per week. Based on these estimates, we show the impacts of posting additional bulletins on artists' digital track sales in Table 3. Depending on the sales rank of an artist, the increase in sales from the MySpace promotions can range from a few to a couple of thousands each week.

Table 3: Estimated Impact of 20% Increase in the Number of Bulletins on Weekly Digital Track Sales

Sales Rank	**1**	**10**	**100**	**1,000**	**10,000**	**100,000**
Weekly Digital Track Sales	296,559	54,176	9,897	1,808	330	60
Impact of Bulletins (0.8%)	2,372.5	433.4	79.2	14.5	2.6	0.5

Notes: This table presents the economic impacts of the coefficient estimates in Table 2. The same method can be applied to the coefficient estimates shown in other tables such as Tables 4, 5, 6, and 7.

In addition, the PVAR model estimates also reveal whether artists' marketing actions on MySpace Music can be affected by music sales. The coefficient for the sales rank (in logarithm) in the *ln(1+Bulletins)* equation in Table 2 is negative (-0.031) and statistically significant at the 5% level; however, the coefficient for the sales rank (in logarithm) in the *ln(1+FriendUpdates)* equation is insignificant at the 5% level. This suggests that artists' promotional activities are less likely to be affected by their sales ranks. In the following discussions, we focus more on how MySpace promotional activities affect sales ranks and do not report the results of the second and third equations, unless necessary.

4.2 Examining the Long-term effects

To examine the long-term behavior of the Panel VAR model, Impulse Response Functions (IRFs) are often employed to describe the effect of one unit increase in one variable on the future values of all variables in the system. The assumption here is that this error returns to zero in subsequent periods and all other errors are equal to zero (Stock and Watson 2001). By this setting we can learn whether a shock to one variable will have a permanent or transitory effect on any of the three variables, and if the effect is transitory, how long it will take to dissipate. To illustrate how the impulse responses are calculated, we

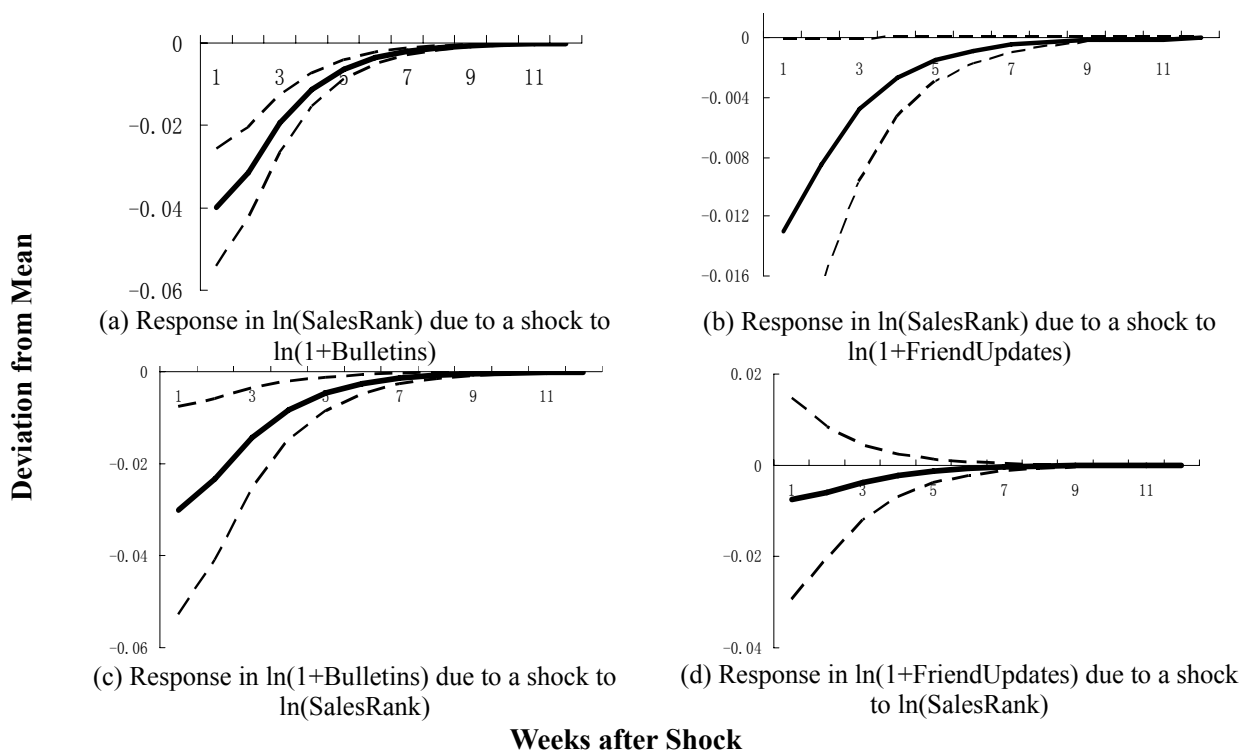
rewrite equation (1) as its vector moving average (MA) representation after dropping the subscript *i* (Hamilton 1994):

$$\mathbf{y}_t = \mathbf{c} + \boldsymbol{\eta}_t + \boldsymbol{\Psi}_1 \boldsymbol{\eta}_{t\text{-}1} + \boldsymbol{\Psi}_2 \boldsymbol{\eta}_{t\text{-}2} + \ldots \tag{3}$$

where $\mathbf{c}$ is a three-element vector of constants; the $\boldsymbol{\Psi}$'s are 3×3 matrices that can be recursively determined given the matrices $\boldsymbol{\Phi}$'s; and $\boldsymbol{\eta}$ is a vector white noise process with $\boldsymbol{\eta}_t = (\eta_{1t}, \eta_{2t}, \eta_{3t})'$ such that $E(\boldsymbol{\eta}_t) = \mathbf{0}$, $E(\boldsymbol{\eta}_t \boldsymbol{\eta}_t') = \sigma^2 \mathbf{I}$, and $E(\boldsymbol{\eta}_t \boldsymbol{\eta}_s') = \mathbf{0}$ for $t \neq s$. Then the (j,k)-th element, ψ_{jk}^s, of the matrix $\boldsymbol{\Psi}_s$ represents the impulse response of y_j with respect to the sth lagged innovation η_k

$$\frac{\partial y_{j,t+s}}{\partial \eta_{k,t}} = \frac{\partial y_{j,t}}{\partial \eta_{k,t-s}} = \psi_{jk}^s, \; j,k = 1,2,3. \tag{4}$$

Figure 1: Impulse Response Functions



(a) Response in ln(SalesRank) due to a shock to ln(1+Bulletins)

(b) Response in ln(SalesRank) due to a shock to ln(1+FriendUpdates)

(c) Response in ln(1+Bulletins) due to a shock to ln(SalesRank)

(d) Response in ln(1+FriendUpdates) due to a shock to ln(SalesRank)

Note: Dotted lines are the 5th and 95th percentiles.

Figure 1 presents the impulse response functions along with the 90% confidence intervals generated from Monte Carlo simulations[14]. We are particularly interested in how the sales rank responses to a shock

[14] We generate a random draw of the coefficient matrices $\boldsymbol{\Phi}$'s in equation (1) using the coefficient estimates and their

to the broadcasting activities (Figures 1a and 1b) and how the broadcasting activities response to a shock to the sales rank (Figures 1c and 1d). Figures 1a and 1b illustrate that an unexpected one-unit increase in the variable *ln(1+Bulletins)* (or *ln(1+FriendUpdates)*) is associated with a 4% decrease (or 1.3% decrease) in the logarithm of sales rank at t=1 (i.e., increase in sales). The effect of bulletins on sales rank is significantly different from zero as we go from week 1 to week 6. Both effects gradually reduce to zero after 9 weeks as the carry-over effect eventually dies out. Figures 1c and 1d show the different responses of broadcasting activities to a positive sales rank shock. We observe a slightly significant decrease in bulletins and an insignificant impact on friend updates going from week 1 to week 5. Overall, IRFs provide a graphical representation of how the system evolves over time, and its short-term result at t=1 is consistent with the interpretation of the coefficient estimates in Table 2.

4.3 How Market Reach Moderates Marketing Effects

Now we consider how the market reach measured by the network size could moderate the relationship between artists' promotional activities and music sales. The underlying reasoning is that the effect of promotions depends on not only how hard artists promote themselves on social networking sites but also their network size, which determines the target size of their promotional activities. With a small network, artists' promoting messages can be received by only a few people, thus limiting the potential benefits of these broadcasting efforts. However, with a large network, a moderate amount of effort in broadcasting may attract a huge amount of attention from music fans and eventually lead to a lot of music purchases by them.

To study this moderating effect of network size, we divide the sample into two equal-size groups based on the number of friends for each artist at the beginning of the study period. We then estimate the Panel VAR model for each group using one lag for the *ln(SalesRank)* equation and compare the results together with the whole sample in Table 4. To evaluate the consistence of the GMM estimator, we also compare our results with the within-group estimates from the two-way fixed effects models (artist and

---

variance-covariance matrix and calculate the impulse responses in equation (4). We repeat this procedure ten thousand times and obtain the 5th and 95th percentiles to construct the confidence interval.

time effects). As shown in Table 4, the coefficient estimates of the within-group estimator in columns (1)-(3) are in general much lager than the corresponding ones of the GMM estimator in columns (4)-(5); thus, the within-group estimator for this dynamic panel model tends to overestimate the impact of MySpace promotional activities on music sales. However, the signs of the coefficients are all in the right directions.

Table 4: Coefficient Estimates for the ln(SalesRank) Equation

	Within-Group Estimator			GMM		
		Network Size			Network Size	
	All Artists (1)	≥Median (2)	<Median (3)	All Artists (4)	≥Median (5)	<Median (6)
$\ln(\text{SalesRank}_{i(t-1)})$	0.679**	0.732**	0.596**	0.549**	0.666**	0.428**
	(0.005)	(0.007)	(0.008)	(0.014)	(0.015)	(0.017)
$\ln(1+\text{Bulletins}_{i(t-1)})$	-0.060**	-0.063**	-0.049**	-0.040**	-0.052**	-0.027
	(0.006)	(0.008)	(0.011)	(0.009)	(0.010)	(0.014)
$\ln(1+\text{FriendUpdates}_{i(t-1)})$	-0.036**	-0.049**	-0.022*	-0.013	-0.035**	0.001
	(0.006)	(0.009)	(0.009)	(0.008)	(0.010)	(0.011)
Number of Artists	631	316	315	631	316	315

Notes: Individual and time fixed effects are included in the within-group estimator; numbers in parentheses are standard errors; ** and * denote significance at 1% and 5%, respectively.

Focusing only on the GMM estimates, we have the following observations. First, when we study the artists with a relatively large network size (i.e., larger than or equal to the median network size), the coefficients for both bulletins and friend updates in column (5) are more negative (compared with the coefficients in column (4) when studying the whole sample) and statistically significant at the 1% level. In particular, the impact of friend updates increases from 1.3% for the whole sample to 3.5% for the artists with a relatively large network size, while the impact of bulletins increases from 4.0% to 5.2%. This implies that the marketing effect of bulletins and friend updates on sales is much larger for the artists with more friends. Second, for the artists with a relatively small network size (i.e., smaller than the median), the coefficients for both bulletins and friend updates in column (6) are no longer significant even at the 5% level. The opposing results exhibited by these two groups reveal that the network size is an important factor in determining the relationship between artists' online promotional activities and music sales. We also divide the sample at the 25th and 75th percentiles and estimate the models for the top and

bottom 25% of the data; the results are very similar to those presented here, ensuring that our conclusion is not affected by how we divide the sample.

4.4 Controlling for Time-varying Individual Characteristics

Fixed effects models only enable us to control for time-invariant individual characteristics. If there are some time-varying individual characteristics that could influence both the sales rank and MySpace promotions, then we infer a relationship that is nonexistent in reality. For example, one important factor that is usually unobservable, but could play a major role in determining an artist's music sales, is his or her popularity among consumers over time. This can be attributed to many things, such as promotions in traditional channels, concert tours, new album releases, or even celebrity scandals. Since there is no direct measure of artist popularity, we resort to the Google Trends data, which measures the volume of consumer searches on specific artists. This data source represents the level of consumer interest in different artists over time; so, it can be a good indirect measure of artist popularity. Consumer searches on artists are also highly correlated with news events and media coverage of artists; therefore, by including the Google Trends measures, we also control for these previously unobservable factors that could have an influence on music sales.

Google Trends provides data starting in January 2004, and the data is publicly available. We use the name of the artist as the search term and download the weekly search volume index for each artist. There is no absolute count of consumer searches on a specific artist name as the data is scaled based on the average search traffic of the term entered. However, Google Trends allows one to compare the popularity of two or more search terms over time. Hence, given the same benchmark, it will not matter if we only have the data on scaled search traffic. To collect the search volume indices on all the artists in our sample, we, thus, need to choose a benchmark artist so that the data for all artists is scaled based on the same search traffic and, hence, comparable. We first select the popular female artist Beyoncé as our benchmark and collect the search volume index data using this benchmark.

One potential hazard with this data source is that it is censored. The minimal value available for

display is 0.01 (note that Google scales the search volumes over time so that the average search traffic for Beyoncé is 1.0); so, any value smaller than 0.01 but still larger than zero may be displayed as zero. This creates a censorship problem as the data for obscure artists may be inaccurate because we can not distinguish observed zero values from true zero values. To counter this problem, we limit our sample to only those artists who always have a positive search volume index compared to Beyoncé over the study period. This ensures that the search volume data is accurate for each artist. We are able to find 297 artists that satisfy this constraint. To be conservative, we also exclude the artists whose names are too general, such as Bond, CSS, Ivy, Panda, etc. because such general names may be confused with other terms, making the search volume index inaccurate (i.e., we cannot be sure if the user is searching for these artists or other unrelated things). This leaves us with 278 artists in total. The correlation between the search volume index and the log of the sales rank is -0.278 for these 278 artists, which implies that highly ranked artists tend to attract more consumer searches.

Table 5: Panel VAR Coefficient Estimates with SearchIndex (N=278, T=32)

Independent Variable	**Dependent Variable**			
	ln(SalesRank)	ln(1+Bulletins)	ln(1+FriendUpdates)	SearchIndex
$\ln(\text{SalesRank}_{i(t-1)})$	0.643**	-0.036*	-0.017	-0.006**
	(0.014)	(0.018)	(0.018)	(0.002)
$\ln(1+\text{Bulletins}_{i(t-1)})$	-0.052**	0.207**	0.057**	0.006**
	(0.011)	(0.014)	(0.014)	(0.002)
$\ln(1+\text{FriendUpdates}_{i(t-1)})$	-0.035**	0.032*	0.120**	0.004*
	(0.011)	(0.014)	(0.014)	(0.002)
$\text{SearchIndex}_{i(t-1)}$	-0.427**	0.007	-0.018	0.366**
	(0.085)	(0.098)	(0.098)	(0.015)

Notes: numbers in parentheses are standard errors; ** and * denote significance at 1% and 5%, respectively.

To incorporate the search volume index into the basic model, we treat this new variable *SearchIndex* as endogenous and estimate a Panel VAR model with four equations, i.e., the dependent variable vector in Equation (1) becomes $\mathbf{y}_{it} = (\ln(SalesRank_{it}), \ln(1+Bulletins_{it}), \ln(1+FriendUpdates_{it}), SearchIndex)'$. To illustrate all the relationships between different variables, we report the GMM coefficient estimates for all four equations in Table 5. We have the following findings. First, in the *ln(SalesRank)* equation, the coefficient for *ln(1+Bulletins)* is -0.052 and significant at the 1% level, while the coefficient for

ln(1+FriendUpdates) is -0.035 and also significant at the 1% level. This means that even after controlling for consumer interest on artists, the impact of broadcasting activities on sales ranks is still significant. Second, the coefficient for search volume index in the *ln(SalesRank)* equation is also negative and statistically significant, indicating that more consumer searches are associated with larger music sales. Third, the coefficients for *ln(SalesRank)* in the *ln(1+Bulletins)* and *ln(1+FriendUpdates)* equations remain consistent with the previous result that the sales rank may influence the number of bulletins, but is unlikely to have an impact on the number of friend updates. Fourth, in the *SearchIndex* equation, the coefficient for *ln(SalesRank)* is negative (-0.006) and statistically significant at 1%, suggesting that an increase in the sales rank of the artist (i.e., sales decrease) this week is associated with a decrease in the volume of consumer searches next week. Relating to the coefficient for search volume index in the *ln(SalesRank)* equation, we conclude that music sales and consumer searches influence each other. Finally, the coefficients for the bulletins and friend updates variables in the *SearchIndex* equation are positive and significant, implying that intense promotional activities in social networks can drive consumers' search volumes.

To address the concern that Beyoncé may be too popular to be used as a benchmark, we also download the data using Nickel Creek, an obscure music group with a unique name, as a benchmark. Our sample then increases to 320 artists, after excluding those with general names as before. The results remain consistent with those presented in Table 5 and, thus, are omitted here.

5. Robustness Checks and Additional Analyses

In this section, we address a few potential issues to check the robustness of our results and conduct some additional analyses to gain further insights. First, we incorporate the number of new friends as an endogenous variable into the baseline model and see if our main results still hold. Second, we examine the relationship of the main variables for two types of artists, major labels and indie artists. Third, we study the impact of artist-generated content and user-generated content on the sales rank at the same time. Finally, we test a model specification in which the sales rank is a nonlinear function of broadcasting

activities.

Table 6: Panel VAR Coefficient Estimates with NewFriends

Independent Variable	Dependent Variable			
	ln(SalesRank)	ln(1+Bulletins)	ln(1+FriendUpdates)	ln(NewFriends)
$\ln(\text{SalesRank}_{i(t-1)})$	0.554**	-0.025	-0.007	-0.033
	(0.013)	(0.013)	(0.013)	(0.023)
$\ln(1+\text{Bulletins}_{i(t-1)})$	-0.042**	0.223**	0.037**	0.001
	(0.009)	(0.009)	(0.010)	(0.015)
$\ln(1+\text{FriendUpdates}_{i(t-1)})$	-0.013	-0.0004	0.102**	0.014
	(0.008)	(0.009)	(0.009)	(0.013)
$\ln(\text{NewFriends}_{(t-1)})$	-0.013*	0.012	0.022**	0.143**
	(0.006)	(0.007)	(0.007)	(0.011)

Notes: numbers in parentheses are standard errors; ** and * denote significance at 1% and 5%, respectively.

5.1 Endogeneity of Network Size

In general, each artist's network size gradually increases over time. When studying the moderating effect of network size, we use the network size at the beginning of the study period so that it is exogenous for future periods. However, a change in the network size itself may be endogenous as artists continuously add new friends over time and, thus, influence the relationship between MySpace activities and music sales. Because the variation in the network size variable is very small, it is impossible to estimate its effect in a fixed effects model. We, therefore, test the endogeneity of network size by adding the number of new friends for each artist in each period, *NewFriends*, as an endogenous variable to the basic Panel VAR model and estimate the new model using GMM. The median change in network size per week is very small, only 16 friends; the 75th percentile is 54; but there exist some extreme outliers that drive the mean to 99. Table 6 reports the coefficient estimates for all four equations. We highlight two important results here. First, the coefficients in the *ln(SalesRank)* equation do not change much compared to previous estimates. The coefficient for *ln(NewFriends)* is negative (-0.013) and statistically significant at the 5% level, implying that an increase in the network size has a negative impact on the sales rank (i.e., a positive impact on sales). Second, in the *ln(NewFriends)* equation, the coefficients for *ln(SalesRank)*, *ln(1+Bulletins*) and *ln(1+FriendUpdates)* are all statistically insignificant. This suggests that the sales rank and social networking activities do not have a direct impact on the network size, at least in the short

run. Therefore, we conclude that influential promotions do not lead to an abrupt change in the network size in the short term and network size can be assumed as exogenous in the analysis.

Table 7: Major Labels vs. Indie Artists

	All (1)	Major label (2)	Indie (3)
$\ln(\text{SalesRank}_{i(t-1)})$	0.549** (0.014)	0.625** (0.017)	0.499** (0.018)
$\ln(1+\text{Bulletins}_{i(t-1)})$	-0.040** (0.009)	-0.047** (0.012)	-0.051** (0.013)
$\ln(1+\text{FriendUpdates}_{i(t-1)})$	-0.013 (0.008)	-0.021* (0.010)	-0.009 (0.012)
# Artists	631	333	247

Notes: numbers in parentheses are standard errors; ** and * denote significance at 1% and 5%, respectively; column (1) is the same as column (4) in Table 4.

5.2 Major Labels vs. Indie Artists

Besides the moderating effect of the network size, it is also important to know how the impact of these broadcasting activities varies across different types of artists (major label vs. independent or indie). Artists signed to major record labels are primarily advertised through mass marketing, expensive ad campaigns and radio airplay, while indie artists heavily rely on the Internet to promote or even sell their music. Thus, it will be interesting to know whether different types of artists are using the social media sites in the same way and whether the effects of social media marketing on music sales are identical across different types of artists[15]. We can identify the type of artists (major label or indie) through the information revealed on their MySpace profile pages. In our sample of 631 artists, 333 are major labels and 247 are indie artists, with the remaining 51 artists labeled as "Other" on MySpace Music. We estimate the basic model for these two types of artists and present the coefficient estimates for the *ln(SalesRank)* equation in Table 7. As shown in the table, the coefficients for the bulletins and friend updates variables in column (2) are both significant, implying that both bulletins and friend updates have a significant impact on the sales rank for major label artists. The significance of coefficients in column (3) for indie artists follows the same pattern as those in column (1) for all artists; however, the coefficient for the bulletins

[15] Examining the group of indie artists alone further addresses the potential concern of omitted variables due to the unavailability of data on traditional promotions, as indie artists generally do not have the resources to market themselves in the traditional channels.

variable in column (3) is more negative (-0.051) than the corresponding coefficients (-0.040 and -0.047) in columns (1) and (2). This indicates that indie artists enjoy the most benefit from bulletin messages. Besides the difference that bulletins are words written by artists while friend updates are automatically generated news feeds by MySpace Music, we conjecture that bulletins on MySpace Music are similar as "tweets" on Twitter, in which the personal touch can be more easily established for indie artists than friend updates, especially in a smaller social network.

5.3 Impact of User-generated Content

As extensively documented in the literature, user-generated content (UGC) is considered to have a significant impact on product sales. It would be interesting to study artist-generated content and UGC together and see how each influences sales. To do this, we collect the customer reviews data from Amazon, which is probably the most widely used data source for customer reviews. For each artist, we identify the digital albums (including single-song albums) that are released by them. By default, Amazon shows all customer reviews on the same album, no matter whether the format is physical CD or digital album. In this way, we can gather all the reviews through either physical CDs or digital albums. Amazon also allows users to post customer reviews on songs (those belong to a certain album), but the number of songs belonging to an artist is too large for us to handle; so, we only focus on customer reviews on albums (including single-song albums). For each week, we count the total number of *new* customer reviews on any album of an artist (variable *Reviews*). Among those new reviews, we also count the number of positive reviews (variable *PosReviews*) as well as the number of negative reviews (variable *NegReviews*), which are determined in the following way: if the new review rating is larger than the current aggregate rating of the artist, we treat it as a positive review; if the new review rating is smaller than the current aggregate rating of the artist, we treat it as a negative review. We work directly with these two variables instead of the variable *Reviews* and the average rating of new reviews, because there are too many missing observations for the rating since there is no new customer review for an artist in a week three fourths of the time. Also for this reason, the aggregate rating for an artist does not change much over time and can not be estimated in a fixed effects model.

Table 8: Panel VAR Coefficient Estimates with Amazon Review Data

Independent Variable	Dependent Variable				
	SalesRank	Bulletins	FriendUpdates	PosReviews	NegReviews
$\ln(\text{SalesRank}_{i(t-1)})$	0.541**	-0.026*	-0.001	-0.082**	-0.077**
	(0.012)	(0.013)	(0.013)	(0.011)	(0.009)
$\ln(1+\text{Bulletins}_{i(t-1)})$	-0.042**	0.222**	0.036**	0.029**	0.004
	(0.009)	(0.009)	(0.010)	(0.008)	(0.007)
$\ln(1+\text{FriendUpdates}_{i(t-1)})$	-0.013	-0.001	0.099**	0.019*	0.015*
	(0.008)	(0.009)	(0.009)	(0.008)	(0.006)
$\ln(1+\text{PosReviews}_{(t-1)})$	-0.080**	-0.001	-0.003	0.122**	0.073**
	(0.008)	(0.009)	(0.009)	(0.008)	(0.007)
$\ln(1+\text{NegReviews}_{(t-1)})$	-0.087**	0.015	0.031**	0.092**	0.114**
	(0.010)	(0.011)	(0.012)	(0.010)	(0.008)

Notes: Dependent variables are also log-transformed in the same way as the independent variables and the logarithm operators are omitted to save space; numbers in parentheses are standard errors; ** and * denote significance at 1% and 5%, respectively.

To incorporate the UGC into the system, we add the two variables *PosReviews* and *NegReviews* (with log transformation) as endogenous variables and estimate a Panel VAR model with five endogenous variables. Table 8 presents the estimation results. Previous results about the relationship between broadcasting activities and the sales rank still hold. As to the impact of UGC on sales, both coefficients for *ln(1+PosReviews)* and *ln(1+NegReviews)* in the *ln(SalesRank)* equation are negative (-0.080 and -0.087, respectively) and statistically significant at the 1% level, which implies that the number of either positive or negative reviews is negatively associated with the sales rank. Perhaps more interestingly, the magnitude of each of these two coefficients are almost twice as large as the coefficient for *ln(1+Bulletins)*, suggesting that the impact of the number of customer reviews on sales is twice as large as the impact of the same number of bulletins on sales.

5.4 A Nonlinear Specification

One concern is that artists may over advertise in social media because of the low cost. Then, using only a linear representation may not be sufficient enough to model the complex relationship between artists' broadcasting activities and their music sales. To specify a nonlinear relationship, we use the raw numbers of bulletins and friend updates and add two quadratic terms, the squares of the numbers of bulletins and friend updates, as endogenous variables to the basic model. We can still fit a linear regression, although the sales rank is now nonlinear in the variables of social networking activities. The

estimation results are presented in Table 9. We find that the coefficient estimate for the quadratic term $Bulletins^2_{i(t-1)}$ is positive (0.0003) and statistically significant at the 5% level. However, the coefficients of both $FriendUpdates_{i(t-1)}$ and $FriendUpdates^2_{i(t-1)}$ are insignificant. This suggests that the sales rank can be a convex function of the number of bulletins, and that there exists an optimal level of promotion using bulletins on MySpace Music. Setting the first order condition with respect to the number of bulletins to zero, we find that the optimal number of bulletins is roughly 27 in one week ($\frac{0.016}{2 \times 0.0003} = 26.7$). Thus, this number is the threshold above which the sales rank starts to increase (i.e., sales start to decrease). In our sample, we do observe that some artists are over advertising as the maximum number of bulletins per week is 48, which is far more than the optimal threshold.

Table 9: Coefficient Estimates for the ln(SalesRank) Equation

	GMM Estimator	
	(1)	(2)
$\ln(SalesRank_{i(t-1)})$	0.548**	0.558**
	(0.013)	(0.012)
$Bulletins_{i(t-1)}$	-0.008**	-0.016**
	(0.003)	(0.004)
$FriendUpdates_{i(t-1)}$	-0.008*	-0.008
	(0.004)	(0.006)
$Bulletins^2_{i(t-1)}$		0.0003*
		(0.0002)
$FriendUpdates^2_{i(t-1)}$		0.0001
		(0.001)

Notes: numbers in parentheses are standard errors; ** and * denote significance at 1% and 5%, respectively; column (1) is the same as column (4) in Table 4.

6. Conclusion and Discussion

Because of its fast speed and low cost, broadcasting in social media has evolved into a new form of internet marketing recently. On MySpace, artists regularly update their profile pages to keep music fans interested in their music. Their activities are automatically broadcasted to their friends by the bulletin board and friend updates features. Using the activity stream data from MySpace Music and the sales rank data from Amazon, we employ a panel vector autoregression model to quantify the effect of MySpace marketing on music sales. The proposed Panel VAR model utilizes time series data and is able to address the endogeneity problem among marketing activities, sales ranks, and other focal variables such as artist

popularity. We find that marketing activities on social networking sites yield significant benefits only for the artists with many friends. To translate these effects into actual sales, we estimate that one additional bulletin by an artist can lead to an increase of a few thousand digital track sales each week.

This study provides important managerial implications. First, our results demonstrate that artists' marketing activities on MySpace have a significant impact on the sales of music. Although we have only studied MySpace and the music industry, the models we use and the insights we obtain can be easily applied to many other social media sites and industries. In light of our findings, companies across industries should carefully plan their marketing activities via social media, besides such activities in traditional channels. Second, broadening the market reach, such as building up the fan base and enlarging the size of the friend network, is the key to the success of social media marketing. Not all firms that engage in such activities will be successful in their marketing efforts. The same amount of effort may yield very different payoffs depending on the potential market reach confined by the size of the friend network or follower network. Firms should focus on attracting more people to join their networks at early stages of implementation and gradually enlarge the network size over time for successful social media marketing. To get a rough estimate of how broad the market reach should be to enjoy the benefits, we divide the sample into four equal-size groups by splitting it at the 25th, 50th, and 75th percentiles of the number of friends for each artist. We find that the coefficient estimates on MySpace activities for the first two groups are statistically insignificant and that those for the last two groups are statistically significant. Thus, we estimate that firms should at least aim for a network size of between 40,485 (50th percentile) and 118,322 (75th percentile). Third, we also find that the relationship between marketing activities on MySpace and sales ranks can be nonlinear. Companies should, thus, be careful that they do not overuse social media technology just because it is relatively inexpensive to adopt.

There are a number of ways for future researchers to extend the findings in this paper. Since MySpace is generally the most popular social media site for artists to market their music, we have chosen it to be our data source. However, it is likely that artists have a significant presence at other social media sites such as Facebook, Bebo, and Twitter as well. Activities on these sites could also affect artists' sales

ranks on Amazon. Due to data limitations, we are not able to study the interactions between music sales and the overall marketing intensity across different social media sites. Future research is warranted in this area. The following are some interesting research questions for artists, record companies, as well as academics: How to effectively manage marketing activities across different social media sites? How to allocate limited advertising budgets across different channels? Similarly, this study uses the search index from Google Trends to partially control for artists' marketing campaigns carried out in traditional channels. If a researcher is able to obtain the actual data on traditional marketing activities, it would be interesting to study how marketing activities in traditional channels and those in social media outlets should be coordinated, as well as to compare the effectiveness of social media marketing and that of traditional marketing. Such results would provide more insights to artists and record companies on how to manage the tradeoffs between different types of marketing channels.

In summary, social media marketing is more of a two-way communication in nature, rather than a one-way flow of information. Its success lies in the engagement and interaction with customers. The goal of this study is to provide insights on the effectiveness of social media marketing and address the business implications of broadcasting in social media. As more applications emerge in the future, further research is necessary to understand this new and interesting phenomenon.

Appendix

I. GMM Estimation of Panel VAR models

To illustrate the estimation of a Panel VAR model in general, we rewrite the Equation (1) in the main text in the matrix form with *m* endogenous variables and *p* lags. Let $\mathbf{y}_{it}$ be an $m \times 1$ vector of endogenous variables for individual *i* at time *t*. The model can be simplified as

$$\mathbf{y}_{it} = \mathbf{\Phi}_1 \mathbf{y}_{i,t\text{-}1} + \mathbf{\Phi}_2 \mathbf{y}_{i,t\text{-}2} + \ldots + \mathbf{\Phi}_p \mathbf{y}_{i,t\text{-}p} + \boldsymbol{\delta}_t + \mathbf{f}_i + \boldsymbol{\varepsilon}_{it} \tag{A1}$$

As the lagged dependent variables $\mathbf{y}_{i,t\text{-}1}, \mathbf{y}_{i,t\text{-}2}, \ldots, \mathbf{y}_{i,t\text{-}p}$ may be correlated with the fixed effect, we can eliminate the individual effect $\mathbf{f}_i$ by first-differencing equation (A1),

$$\Delta\mathbf{y}_{it} = \mathbf{\Phi}_1 \Delta\mathbf{y}_{i,t\text{-}1} + \mathbf{\Phi}_2 \Delta\mathbf{y}_{i,t\text{-}2} + \ldots + \mathbf{\Phi}_p \Delta\mathbf{y}_{i,t\text{-}p} + \Delta\boldsymbol{\delta}_{it} + \Delta\boldsymbol{\varepsilon}_{it} \tag{A2}$$

for $t = p+2, 3, \ldots, T$, and $\Delta\mathbf{w}_{it} = \mathbf{w}_{it} - \mathbf{w}_{it-1}$ for a vector $\mathbf{w}_{it}$. Based on the assumption of the lack of serial correlation in the error term $\boldsymbol{\varepsilon}_{it}$, we can derive the following moment conditions:

$$E[\Delta\boldsymbol{\varepsilon}_{it} \cdot \mathbf{y}'_{i,t-s}] = \mathbf{0} \qquad \text{for } s = p+1, p+2, \ldots, t-1, \tag{A3}$$

In other words, dependent variables that are lagged $p+1$ or more periods qualify as valid instruments for equation (A2). The orthogonality conditions can then be written as

$$E[(\Delta\mathbf{y}_{it} - \sum_{j=1}^{p} \mathbf{\Phi}_j \Delta\mathbf{y}_{i,t\text{-}j} - \Delta\boldsymbol{\delta}_{it}) \cdot \mathbf{q}'_{i,t}] = \mathbf{0} \qquad \text{for } t = p+2, p+3, \ldots T, \tag{A4}$$

where $\mathbf{q}_{it}$ is the $[m(t-p-1)+1] \times 1$ vector of instruments defined by

$$\mathbf{q}_{it} = (\mathbf{y}'_{i1}, \mathbf{y}'_{i2}, \ldots, \mathbf{y}'_{i,t-p-1}, 1)'. \tag{A5}$$

We can rewrite equation (A4) in stacked form as

$$E[\mathbf{Q}'_i(\Delta\mathbf{Y}_i - \sum_{j=1}^{p} \Delta\mathbf{Y}_{i,\text{-}j} \mathbf{\Phi}_j)] = \mathbf{0}, \tag{A6}$$

where $\mathbf{Q}'_i$ is a matrix of dimension $\{\frac{m(T-p)(T-p-1)}{2} + T - p - 1\} \times (T-p-1)$ constructed from the vectors of instruments,

$$\mathbf{Q}_i' = \begin{pmatrix} \mathbf{q}_{i,p+2} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{q}_{i,p+3} & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{q}_{i,T} \end{pmatrix}, \tag{A7}$$

and $\mathbf{\Delta Y}_i$ and $\mathbf{\Delta Y}_{i,-j}$'s are $(T-p-1)\times m$ matrices defined as

$$\mathbf{\Delta Y}_i = (\mathbf{\Delta y}_{i,p+2}', \mathbf{\Delta y}_{i,p+3}', ..., \mathbf{\Delta y}_{i,T}')', \quad \mathbf{\Delta Y}_{i,-1} = (\mathbf{\Delta y}_{i,p+1}', \mathbf{\Delta y}_{i,p+2}', ..., \mathbf{\Delta y}_{i,T-1}')', \ldots$$

and $\mathbf{\Delta Y}_{i,-p} = (\mathbf{\Delta y}_{i,2}', \mathbf{\Delta y}_{i,3}', ..., \mathbf{\Delta y}_{i,T-p}')'$.

By solving the first-order condition of (A6), the standard GMM estimator following Binder et al. (2005) is expressed as

$$\hat{\boldsymbol{\varphi}}_{GMM} = (\mathbf{S}_{\mathbf{ZX}}'\mathbf{A}^{-1}\mathbf{S}_{\mathbf{ZX}})^{-1}\mathbf{S}_{\mathbf{ZX}}'\mathbf{A}^{-1}\mathbf{S}_{\mathbf{Zy}}, \tag{A8}$$

where

$$\mathbf{S}_{\mathbf{ZX}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{Z}_i'\mathbf{X}_i, \quad \mathbf{S}_{\mathbf{Zy}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{Z}_i'\mathbf{y}_i, \quad \mathbf{A} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{Z}_i'\mathbf{\Omega}\mathbf{Z}_i,$$

$$\mathbf{Z}_i' = \mathbf{Q}_i' \otimes \mathbf{I}_m, \quad \mathbf{X}_i' = [\mathbf{\Delta Y}_{i,-1} \quad \mathbf{\Delta Y}_{i,-2} \quad \ldots \quad \mathbf{\Delta Y}_{i,-p}] \otimes \mathbf{I}_m, \quad \mathbf{y}_i = \mathrm{vec}(\mathbf{\Delta Y}_i'),$$

and $$\mathbf{\Omega} = \mathbf{V} \otimes \mathbf{I}_m = \begin{pmatrix} 2 & -1 & 0 & & & \\ -1 & 2 & -1 & & 0 & \\ & & & \ddots & & \\ & 0 & & -1 & 2 & -1 \\ & & & 0 & -1 & 2 \end{pmatrix} \otimes \mathbf{I}_m.$$

A consistent estimate of the variance matrix of $\hat{\boldsymbol{\varphi}}_{GMM}$ can be obtained as

$$\widehat{Var}(\hat{\boldsymbol{\varphi}}_{GMM}) = (\mathbf{S}_{\mathbf{ZX}}'\mathbf{A}^{-1}\mathbf{S}_{\mathbf{ZX}})^{-1}\mathbf{S}_{\mathbf{ZX}}'\mathbf{A}^{-1}\mathbf{\Psi}\mathbf{A}^{-1}\mathbf{S}_{\mathbf{ZX}}(\mathbf{S}_{\mathbf{ZX}}'\mathbf{A}^{-1}\mathbf{S}_{\mathbf{ZX}})^{-1}, \tag{A9}$$

where $\mathbf{\Psi} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{Z}_i'\hat{\mathbf{\Omega}}\mathbf{Z}_i$, $\hat{\mathbf{\Omega}} = \frac{1}{N}\sum_{i=1}^{N}\hat{\mathbf{e}}_i\hat{\mathbf{e}}_i'$, and $\hat{\mathbf{e}}_i = \mathrm{vec}([\mathbf{\Delta Y}_i - \sum_{j=1}^{p}\mathbf{\Delta Y}_{i,-j}\hat{\mathbf{\Phi}}_{j,GMM}]')$.

Note that $\mathbf{\Omega}$ is a weighting matrix that can be replaced with some alternatives, such as $\mathbf{\Omega}_1 = \frac{1}{N}\sum_{i=1}^{N}\hat{\mathbf{e}}_i\hat{\mathbf{e}}_i'$ or $\mathbf{\Omega}_2 = \hat{\mathbf{e}}_i\hat{\mathbf{e}}_i'$, where $\hat{\mathbf{e}}_i = \mathrm{vec}([\mathbf{\Delta Y}_i - \sum_{j=1}^{p}\mathbf{\Delta Y}_{i,-j}\hat{\mathbf{\Phi}}_j]')$, and the $\hat{\mathbf{\Phi}}_j$'s are initial

coefficient estimates by specifying an arbitrary weighting matrix $\mathbf{\Omega}_a$. In this case, the variance matrix of the estimator is simplified to $\frac{1}{N}(\mathbf{S}'_{\mathbf{ZX}}\tilde{\mathbf{A}}^{-1}\mathbf{S}_{\mathbf{ZX}})^{-1}$, where $\tilde{\mathbf{A}}$ varies depending on the chosen weighting matrix.

II. Specification Test

The Sargan test (Sargan 1958, 1959) is one of the most common specification tests applied in this context (see Dahlberg and Johansson 2000). The Sargan statistic Q is the objective function evaluated at the estimated parameters when forming the overidentifying restriction test. Under the null hypothesis of valid instruments, this statistic follows a chi square distribution with a degree of freedom (df) equal to the number of instruments minus the number of parameters.

In our context, we first choose a lag length of five for each of the endogenous variables in each equation and conduct an initial specification test. An even longer length of lags can be used, which should have virtually no influence on the final specification after the lag selection procedure, but we choose five lags particularly due to the fact that bulletin messages on MySpace would disappear from users' pages after ten days and friend updates after one month. So, we think a length of five weeks would be a long enough period to capture the carry-over effects of the broadcasting activities. We carry out the Sargan Test for each equation and the results are shown in Table A.1.

Table A.1: Test of initial model specification: m=3, p=5, T=32, N=631

Dependent Variable	Q (Sargan Statistic)	Degree of freedom	p-value
ln(SalesRank)	117.57	1038	1.000
ln(1+Bulletins)	215.41	1038	1.000
ln(1+FriendUpdates)	211.92	1038	1.000

For each of the three equations, we observe a p-value of close to 1, so we accept the null hypothesis that a lag length of five is a correct specification and the instruments are valid.

III. Lag Selection

The next question we want to answer is whether a model with five lags has achieved the best balance

of accuracy and efficiency. With three endogenous variables in the system, we are actually estimating 15 coefficients for lagged dependent variables in each equation. To test whether the model is over parameterized, we can exclude one lag at a time and see if it is at the expense of accuracy. The Sargan difference statistic is formed by estimating both the restricted and unrestricted models and then calculating the difference of the two Sargan statistics. Under the null hypothesis, this difference statistic follows a chi square distribution with $df = df_R - df_U$. If the p-value is very small (<0.01), then we reject the null hypothesis and use the unrestricted model with the larger number of lags.

Table A.2: Test of Lag Length

Dependent Variable	Lag reduction	Sargan Difference Statistic	Degree of Freedom	p-value
ln(SalesRank)	5→4	10.2	84	1.000
	4→3	12.9	87	1.000
	3→2	18.0	90	1.000
	2→1	28.0	93	1.000
	1→0	901.9	96	0.000
ln(1+Bulletins)	5→4	19.7	84	1.000
	4→3	29.4	87	1.000
	3→2	27.3	90	1.000
	2→1	49.7	93	0.999
	1→0	1778.2	96	0.000
ln(1+FriendUpdates)	5→4	20.8	84	1.000
	4→3	13.9	87	1.000
	3→2	22.5	90	1.000
	2→1	36.1	93	1.000
	1→0	2130.3	96	0.000

In Table A.2, we report the results of the lag selection procedure. Starting with the *ln(SalesRank)* equation, we see that the lag length can be reduced from five to four as the Sargan difference statistic is only 10.2 with a degree of freedom of 84 and the p-value is close to 1. Continuing on, we find that the lag length can be further shortened to one, but not from one to zero (the Sargan difference statistic is 901.9 and the null can be rejected at any reasonable significance level). Turning to the *ln(1+Bulletins)* and *ln(1+FriendUpdates)* equations, we observe a similar pattern as the *ln(SalesRank)* equation. These results indicate that for each of the three equations, one lag is the best specification that ensures both parsimony

and validity of the instruments.

For a given lag length, one also has the option of using only a subset of the instruments. With a T equal to 32 in our sample, the number of available instruments for $p = 1$ is 1395, which is huge. To be cautious, we have also checked the results when we use only a subset of all instruments. The results are very similar to those presented in the main text of the paper.

References

Ahn, S.C., P. Schmidt. 1995. Efficient estimation of models for dynamic panel data. *Journal of Econometrics* **68**(1), 5-27.

Ahn, S.C., P. Schmidt. 1997. Efficient estimation of dynamic panel data models: Alternative assumptions and simplified estimation. *Journal of Econometrics* **76**(1-2), 309-321.

Arellano, M. 2003. *Panel Data Econometrics.* Oxford University Press, Oxford.

Arellano, M., S. Bond. 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies* **58**(2) 277-297.

Arellano, M., O. Bover. 1995. Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics* **68**(1) 29-51.

Bass, F. M. 1969. A simultaneous equation regression study of advertising and sales of cigarettes. *Journal of Marketing Research* **6**(3) 291-300.

Baker, B. 2006. MySpace music marketing: How to promote & sell your music on the world's biggest networking web site. http://www.bob-baker.com/myspace/music.html.

Bessler, D. A. 1984. An analysis of dynamic economic relationships: an application to the U.S. hog market. *Canadian Journal of Agricultural Economics* **32**(1) 109-124.

Binder, M., C. Hsiao, M. H. Pesaran. 2005. Estimation and inference in short panel vector autoregressions with unit roots and cointegration. *Econometric Theory* **21**(4) 795-837.

Blundell, R., S. R. Bond. 1998. Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics* **87**(1) 115-143.

Brynjolfsson, E., Y. J. Hu, M. D. Smith. 2003. Consumer surplus in the digital economy: Estimating the value of increased product variety at online booksellers. *Management Science* **49**(11) 1580-1596.

Chellappa, R. K., C. Chen. 2008. On the temporal nature of sales-rank relationships of albums and digital tracks in the music industry: The relevance of Billboard charts post digitization. Conference on Information Systems and Technology. College Park, Maryland.

Chen, P., S. Dhanasobhon, M. D. Smith. 2007. An analysis of the differential impact of reviews and reviewers at Amazon.com. *Proceedings of the 28th International Conference on Information Systems,* Montreal, Quebec, Canada.

Chevalier, J. A., D. Mayzlin. 2006. The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research* **43**(3) 345-354.

Dahlberg, M., E. Johansson. 2000. An examination of the dynamic behaviour of local governments using GMM bootstrapping methods. *Journal of Applied Econometrics* **15**(4) 401-416.

Dekimpe, M. G., D. M. Hanssens. 1995. The persistence of marketing effects on sales. *Marketing Science* **14**(1) 1-21.

Dellarocas, C. 2003. The digitization of word of mouth: Promise and challenges of online feedback mechanisms. *Management Science* **49**(10) 1407-1424.

Dellarocas, C. 2006. Strategic manipulation of Internet opinion forums: Implications for consumers and firms. *Management Science* **52**(10) 1577-1593.

Dewan, S., J. Ramprasad. 2009. Chicken and egg? Interplay between music blog buzz and album sales, *PACIS 2009 Proceedings,* Paper 87.

Forman, C., A. Ghose, B. Wiesenfeld. 2008. Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Information Systems Research* **19**(3) 291-313.

Givon, M., D. Horsky. 1990. Untangling the effects of purchase reinforcement and advertising carryover. *Marketing Science* **9**(2) 171-187.

Godes, D., D. Mayzlin, Y. Chen, S. Das, C. Dellarocas, B. Pfeiffer, B. Libai, S. Sen, M. Shi, P. Verlegh. 2005. The firm's management of social interactions. *Marketting Letters* **16**(3/4) 415-428.

Gonsalves, A. 2009. Dell makes $3 million from Twitter-related sales. *InformationWeek*, June 12.

Granger, C. W. J. 1969. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* **37**(3) 424-438.

Hamilton, J. 1994. *Time series analysis*. Princeton University Press, Princeton, NJ.

Hansen, L. P. 1982. Large sample properties of generalized method of moments estimators. *Econometrica* **50**(4) 1029-1054.

Hanssens, D. M. 1980. Market response, competitive behavior and time-series analysis. *Journal of Marketing Research* **17**(4) 470-485.

Harris, R. D. F., E. Tzavalis. 1999. Inference for unit roots in dynamic panels where the time dimension is fixed. *Journal of Econometrics* **91**(2) 201-226.

Hayashi, F. 2000. *Econometrics*. Princeton University Press, Princeton, NJ.

Holtz-Eakin, D., W. Newey, H. S. Rosen. 1988. Estimating vector autoregressions with panel data. *Econometrica* **56**(6) 1371-1395.

Horsky, D., L. S. Simon. 1983. Advertising and the diffusion of new products. *Marketing Science* **2**(1) 1-17.

Learmonth, M. 2010. New Twitter rules clamp down on ad networks. Advertising Age, May 24.

Levin, A., C.-F. Lin, C.-S. J. Chu. 2002. Unit root tests in panel data: Asymptotic and finite-sample properties. *Journal of Econometrics* **108**(1) 1-24.

Li, X., L. M. Hitt. 2008. Self-selection and information role of online product reviews. *Information Systems Research* **19**(4) 456-474.

Lin, M., N. R. Prabhala, S. Viswanathan. 2009. Can social networks help mitigate information asymmetry in online markets? *Proceedings of the 30th International Conference on Information Systems,* Phoenix, Arizona.

McCarty, T. A., S. J. Schmidt. 1997. A vector-autoregression analysis of state-government expenditure. *American Economic Review* **87**(2) 278-282.

Nickell, S. 1981. Biases in dynamic models with fixed effects. *Econometrica* **49**(6) 1417-1426.

Nielsen Company 2009 Year-End Music Industry Report. 2010. 2009 top ten selling digital artists. (Available at http://blog.nielsen.com/nielsenwire/wp-content/uploads/2010/01/Nielsen-Music-2009-Year-End-Press-Release.pdf.)

Owyang, J. 2008. Social network stats: Facebook, MySpace, reunion (Jan, 2008).

http://www.web-strategist.com/blog/2008/01/09/social-network-stats-facebook-myspace-reunion-jan-2008/.

Pauwels, K., I. Currim, M. G. Dekimpe, D. M. Hanssens, N. Mizik, E. Ghysels, P. Naik. 2004. Modeling marketing dynamics by time series econometrics. *Marketing Letters* **15**(4) 167-183.

Sargan, J. D. 1958. The estimation of economic relationships using instrumental variables. *Econometrica* **26**(3) 393-415.

Sargan, J. D. 1959. The estimation of relationships with autocorrelated residuals by the use of the instrumental variables. *Journal of the Royal Statistical Society, Series B (Methodological)* **21**(1) 91-105.

Silver, C. 2009. Organized chaos: Viral marketing, meet social media. Retrieved February 23, 2010, http://www.wired.com/epicenter/2009/10/organized-chaos-2/.

Simester, D., Y. J. Hu, E. Brynjolfsson, E. Anderson. 2009. Dynamics of retail advertising: Evidence from a field experiment. *Economic Inquiry* **47**(3) 482-499.

Sims, C. A. 1980. Macroeconomics and reality. *Econometrica* **48**(1) 1-48.

Sims, C. A. 1992. Interpreting the macroeconomic time series facts: The effects of monetary policy. *European Economic Review* **36**(5) 975-1000.

Steel, E. 2009. MySpace tries to recover its cool. *Wall Street Journal*, October 15.

Stock, J. H., M. W. Watson. 2001. Vector Autoregressions. *Journal of Economic Perspectives* **15**(4) 101-115.

The NPD Group. 2010. The NPD Group: Amazon ties Walmart as second-ranked U.S. music retailer, behind industry-leader iTunes [Press release]. Retrieved Feb. 23, 2010, http://www.npd.com/press/releases/press_100526.html

Trusov, M., R. E. Bucklin, K. Pauwels. 2009. Effects of word-of-mouth versus traditional marketing: Findings from an Internet social networking site. *Journal of Marketing* **73**(5) 90-102.

Vincent, F. 2006, July 1. MySpace for Musicians. Retrieved Feb. 23, 2010, http://emusician.com/tutorials/emusic_myspace_musicians/.

Vincent, F. 2010. *MySpace for Musicians*. Course Technology PTR.

Zhu, F., X. M. Zhang. 2010. Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics. *Journal of Marketing* **74**(2) 133-148.