

# Bandwagon and Reputation Effects in the Popular Music Charts

by

Anders Grimsø Moe and Peter E. Earl\*

School of Economics  
University of Queensland  
St Lucia, Brisbane  
Queensland 4072, Australia

\*Corresponding author: e-mail [p.earl@uq.edu.au](mailto:p.earl@uq.edu.au); phone +61 7 3365 6598

## Abstract

Recorded music is often an experience good, especially in the case of albums and is often consumed in a social context. Reputation and bandwagon effects may thus be expected to have a major influence on a recording's sales. This paper draws a distinction between network bandwagon effects (based on interaction within a social network) and market bandwagon effects (driven by market-level signals of behaviour). We use data from the top 20 singles and top 40 albums charts in Norway to investigate the significance of these market bandwagon and reputation effects. A simple model is proposed in which predicts the highest chart position reached by a recording as a function of its initial entry point, the difference between its first and second weeks in the charts, and whether or not the artists has had previous chart success. We find strong evidence consistent with a bandwagon effect, but not for the impact of reputation.

Keywords: recorded music industry, bandwagon effects, reputation, music charts

JEL classification codes: D12, Z11

## **1 Introduction**

Pre-recorded popular music provides an excellent example of a product whose demand goes through a lifecycle rather than being steady unless prices or incomes change. A variety of lifecycle sales profiles are evident. Most of the recordings that are released fail to win significant sales but a few take off very rapidly and sell in huge quantities before their markets become saturated or they fall out of fashion, after which they may continue to generate a much lower volume of sales as 'back catalogue' products. Occasionally, as occurs sometimes in the movie business (see De Vany, 2004), there are recordings whose initial sales are not spectacular but which enjoy a growing market presence as word gets around and they achieve some kind of 'cult' status over a long period amongst a particular type of buyer.

The existence of such lifecycles poses a major challenge to record companies and retailers: how can they estimate demand so that their products do not go out of stock temporarily due to them being unexpectedly successful or do not end up having to be remaindered with a large discount due to far too many units being manufactured? This problem is particularly acute for those recordings that make it into the top-20 sales charts since sales tend to be skewed sharply in favour of those who make it to the upper reaches of the charts. The problem would be compounded by bandwagon effects (Leibenstein, 1950) such that the success of a recording in one period is for whatever reason a function of its observed sales success in the previous period. If bandwagon effects are present and a record company fails to supply enough copies to record stores (either due to it not making enough to satisfy their orders, or due to the record stores under-ordering and then having to wait for further supplies) this does not simply result in sales being deferred

until the unexpectedly popular recording is back in stock. Rather, if it goes out of stock, this will harm its position in the sales charts for that week and reduce demand for it in subsequent weeks. Similarly, if consumers use the presence of highly discounted recordings as a signal that the artist in question is falling out of favour, over-production that is followed by discounting may reduce total revenue from that recording and from subsequent recordings by the same artist.

This problem is mitigated somewhat by the ability of record companies to switch their production plants rapidly between different recordings and to use strategies such as partial vertical integration into the retailing of recorded music (as with HMV and Virgin stores) as means of gathering better market intelligence. Even so, risks associated with the combination of sales spikes, skewed distributions of sales towards the upper end of the charts, and bandwagon effects would be reduced if firms in this industry could model how recordings moved up the sales charts. This paper attempts to show what might be done with very simple models that focus on bandwagon and reputation effects. It uses data from the Norwegian popular music charts. The risk of stock-outs becomes much less of an issue for suppliers if music can also be purchased online and, because the recorded music market is moving increasingly towards downloading, via retailers such as iTunes and Amazon, we use data from just before this transition begins to take hold.

The rest of the paper is organized as follows. Section 2 explores the nature of the choice problem facing buyers of pre-recorded music and how this is likely to give rise to bandwagon effects, superstars and a major role for reputation as a determinant of sales success. Section 3 reviews existing literature in this area. Section 4 presents the model and hypotheses to be tested, while section 5 discusses the data set and section 6 outlines

our estimates of the model as applied to the singles charts. The model's robustness and consistency are explored in section 7, and it is applied also to the album charts. Section 8 is a concluding discussion.

## **2 The process of choosing pre-recorded music**

When choosing between rival providers of pre-recorded popular music, consumers face major challenges from which orthodox rational choice theory has tended to abstract. A typical record store has thousands of different products in stock and could order in many thousands more if customers requested particular items. Prior to the advent of music downloading services that offer partial previews, consumer had access to listening facilities at some stores, but they had little hope of discovering what more than a tiny fraction of the products in stock sounded like. Whilst pre-recorded music is in principle a search good (Nelson, 1970), for practical purposes most of the recordings on offer historically have been experience goods: unless the consumer had heard them in their entirety via electronic media (which was only likely with singles that had previously been successful, or were currently in the charts or were new releases that are on play-lists of disk jockeys on the radio), or through interacting socially, there was the risk of being disappointed due to quality uncertain issues being impossible to resolve in the record store due to lack of time and adequate listening facilities. While the popular music products themselves are typically quite simple compared with esoteric classical music that takes many hearings to appreciate properly, the information environment facing the buyer of popular music is clearly one that imposes bounded rationality. Even with the

advent of iTunes and Amazon download previews, there simply is not enough time to listen to everything that is on offer.

The consumer's problem is not only that there are there major barriers to making globally rational choices of recorded music; the tendency for music to be consumed socially means that the wrong kind of choice may be a source of considerable embarrassment. Departing from the kinds of choices one's peers are making invites inquisition for it challenges the wisdom of their choices (Earl, 1983). The deviant record buyer risks being labelled as having weird (and, by implication, bad) tastes unless he or she has cultivated the kind of reputation for being knowledgeable in this area or being 'hip' in Holbrook's (1995, chapter 10) sense of having the knack of being ahead of the pack in making accurate guesses about what is *going to be* popular. Music consumption is much less of a trial if one chooses products by recording artists with well-established reputations or which for whatever reasons are being bought by one's peers despite not having well-established track records. Imitation of someone else's choice does not challenge the quality of that choice; indeed, if those whose choices are imitated aspire to be seen as fashion leaders, they will welcome being copied even though it may put them under pressure to buy more recorded music that does not yet have widespread currency, in order to maintain their fashion-leader status (cf. Chai, Earl and Potts, 2007).

Sticking with familiar artists has self-reinforcing advantages due to information economies. As Adler (1985, p. 212) noticed, superstars are advantaged because everyone knows something about them. Familiarity limits the start-up costs of discussing their work in a social setting and advantages them via reduced search costs at the time of purchase. The superstars' advantages pose a hurdle for more talented artists seeking to

obtain market share: being rumoured only to be a little better may not be enough to ensure that potential customers will go to the trouble of finding out what they are *really* like and on-selling their merits to their peers. This is an example of what Rothschild (1973) calls a ‘two-armed bandit problem’: the failure to incur the costs of experimenting to discover more about the probability distribution of the quality of music recorded by other artists means that at the next choice point their knowledge of probable qualities of rival artists’ recordings is exactly the same as it was before unless they have been given new information in the interim.

Market shares in this kind of choice environment thus may have very little to do with being cheaper (often, as with first-release movies, newly released music recordings sell for identical prices) or objectively better (because consumers lack the knowledge required to rank all performances even if they could agree on what constitutes quality). Rather, it seems likely to be driven by factors that overcome the problems associated with pre-recorded music being a socially consumed experience good: what we might expect to sell in large quantities is what is low risk due to it being from an artist with an established reputation, or because it is being heard frequently via electronic media (with the performer also being written about or interviewed for electronic and print media) or because it is being consumed already by others.

The last factor is the bandwagon effect, but two variants of it need to be distinguished here. The first we call a ‘network bandwagon effect’, since it involves social network effects in which individuals can observe whether or not members of their network have purchased a particular product and their probability of purchasing it is a function of the number of their network members who have purchased it and their

motivation to copy particular members. The motivation to copy others needs to be included here since they will feel under greater social pressure to conform with some people that they know than with others (cf. the ‘normative beliefs’ component of the model of choice offered by Fishbein and Ajzen, 1975) and may see some members of their network as better judges than others in this context (Earl and Potts, 2004). The second kind of bandwagon effect we call a ‘market bandwagon effect’ since it works via the market as a whole: it is not based on individuals’ particular sets of social network connections but on the knowledge that individuals have about the behaviour of the entire population of which they are members.

The network bandwagon effect may be particularly significant in determining whether a recording makes it into ‘the charts’ in the first place and at what level it enters, but may also affect sales in the weeks after the first entry. However, because this paper’s focus is on how much might be learned from a very simple approach to modelling, we only focus here on the market bandwagon effect. In respect of the latter the obvious crucial factor is very straightforward to observe, namely, how far a recording has made it into the popular music sales ‘charts’, for this will play a major role in determining how often it is broadcast on commercial media and it will also indicate to potential buyers whether there is a ‘bandwagon’ on to which they may safely jump.

### **3. Previous empirical work on success in the popular music charts**

Empirical work on the market for pre-recorded popular music has largely been inspired by the theoretical contributions of Rosen (1981), Adler (1985) and MacDonald (1988) on how superstars benefit from increasing returns to scale and knowledge advantages over non-stars of possibly greater talent, and on the supply of would-be stars. Previous studies

of the behaviour of the market for pre-recorded popular music have been based on the US (Chung and Cox, 1994, Bradlow and Fader, 2001, Fox and Kochanowski, 2004, Giles, 2006, 2007), using data from the *Billboard* Hot 100 chart, or the UK (Strobl and Tucker, 2000), using data from the *New Musical Express* Top 40 albums chart. As Rosen's seminal theoretical analysis of the economics of superstars predicted, extreme skewness in the distribution of chart success is evident. This skewness can be characterised by variants of Lotka and Yule distributions, much as is reported for the distribution of earnings amongst movies and movie stars in the work of De Vany (2004). From this research it appears that there are indeed increasing returns for record companies if they concentrate on creating, nurturing and promoting superstar acts so long as these performers do not use the threat (or actuality) of defection to a rival company or of starting their own record label as a means of capturing the rents that they generate. The most successful of the superstars win vastly more gold and platinum awards than lesser artists and their albums stay in the charts for far longer. The star performers also seem to suffer have been less vulnerable than more minor acts to losing revenue due to MP3 file-sharing amongst consumers (Bhattacharjee, Gopal, Lertwachara, Marsden, and Telang, 2005), and file sharing seems to be related to a shortening of the lifecycles of albums (Gopal, Lertwachara and Marsden, 2007), though because of the experience good problem and multiple outputs of individual performers it is by no means clear that piracy via file-sharing necessarily works against the profitability of record companies (Bhattacharjee, Gopal, Lertwachara and Marsden, 2006).

The present study differs from these works in its focus on the possibility of predicting how high in the charts a recording is likely to get in because of the artist's



reputation and bandwagon effects. The previous studies have been more concerned with modelling the frequency and duration of chart success and hence are less applicable for helping record companies deal with their production planning in respect of the height of the sales spike that might be expected once a recording enters the charts.

The studies that come closest to the present one are those of Strobl and Tucker (2000) and Bradlow and Fader (2001). Strobl and Tucker rather briefly examines the relationship between the number of weeks an album stayed in the UK Top 50 and its entry level. Strobl and Tucker found that the point of first entry affected how long an album stayed in the charts, with entry straight into the top-10 reducing an album's time in the charts by 159 per cent. We suggest that this finding should be read as implying cases where a very powerful reputation effect applies: there is a rush to buy the eagerly-awaited album, rather than its sales being driven by a bandwagon effect, much in the way that blockbuster movies are often shown on a huge number of screens but for a shorter period of time than movies whose sales build through word-of-mouth. (The Harry Potter movies are an example of this.) Strobl and Tucker also found that an initial position below 39<sup>th</sup>-place tended to reduce by 78.6 per cent the length of time an album spent in the UK Top 50 charts. They suggested this was consistent with bandwagon and snowballing effects (which seems likely, given that fewer stores are likely to be displaying prominently the lower reaches of the Top 50). However, their work does not attempt to predict how high an album would rise in the chart. Bradlow and Fader, by contrast, note that a variety of different lifecycle profiles can be found in the *Billboard* Hot 100 charts and they recognizes that the history of an artist's chart success may affect the latent popularity of their recordings, and hence they add to the *Billboard* data set two covariates—whether or

not the artist has previously had a hit, and whether or not the song is from a movie soundtrack—and then use Bayesian econometrics to model the time series of the top-100 in terms of generalized gamma curves based on the latent ‘worth’ of each recording that figures in the charts. This latent ‘worth’ is modelled as a function of the past accumulation profile of sales since the ‘birth’ of the recording. To predict the shape of the curve traced by each song’s chart ranking history during 1993, they used five measures, in addition to the two covariates, namely, debut rank, weeks from debut to peak, peak rank, total number of weeks in the chart and rank on last week in the chart.

#### **4. A Simple Bandwagon Model**

The model that we propose is very simple and focuses only on the determination of a recording’s peak ranking in the charts. It is based on the discussion in section 2. It has two main explanatory variables, the Bandwagon Effect and Reputation, with the recording’s Highest Position Reached in the sales ranking as the dependent variable. By focusing on rankings rather than levels of sales, we can employ readily accessible data and can avoid complications associated with the impact of seasonal factors (especially the Christmas period) on discounting. Given their past sales achievements, record company executives might be expected to have reasonably reliable rules of thumb concerning seasonal variations and likely sales associated with a particular chart position at a particular point in the year. What they may have lacked, however, is a means of predicting how far up the charts a recording would go once it had entered.

To model the bandwagon effect, two independent proxy variables are used, namely, First Entry Position and Position Change from 1<sup>st</sup> to 2<sup>nd</sup> Week. These two variables each represent separate dimensions of the bandwagon effect.

According to bandwagon theory, as the number of people using or recognizing the popularity of a product increases, the more additional people will buy it as consumers gain additional utility from its recognition by others. The First Entry Position variable can thus be an estimate of the signal of the music's popularity, recognition, or fashion strength, and thus, according to the theory, affect purchase behaviour. Position Change from 1<sup>st</sup> to 2<sup>nd</sup> Week, will capture the relative change in popularity or, so to speak, how fast the bandwagon is moving. While the First Entry Position will be the first indication the consumer gets of the record's popularity, the Position Change from 1<sup>st</sup> to 2<sup>nd</sup> Week may be seen as a good indication of the record's popularity trend.

An estimate of Reputation is the final independent variable. This can be estimated by the artist's previous performance on the chart. Here, we use 'Previous Top 10 or not' as a dummy variable as a proxy for the presence or absence of a reputation effect. It is likely that the consumer's uncertainty about an artist will have decreased significantly if a previous recording by the artist reached the top 10 on the chart, as it will have been not only purchased by many but also been exposed through radio, TV and other media. In addition, artists commonly have a stable fan base, inelastic in their buying behaviour of the respective artist's new releases, which limits the relative effect of other factors determining buying behaviour, such as the 'quality' of the music and bandwagon effects/fashion trends in terms of their favourite artist.

Taken together, these elements give us the following simple multiple regression model to estimate:

$$y = \beta_1 + \beta_2x_2 + \beta_3x_3 + \beta_4x_4 + e \quad (1)$$

in which:

$y$  = Best Position Reached

$\beta_1$  = Constant (Intercept)

$\beta_2 \beta_3 \beta_4$  = Coefficients of the respective variables

$x_2$  = First Entry Position

$x_3$  = Position Change from 1<sup>st</sup> to 2<sup>nd</sup> Week

$x_4$  = Previous Top 10 Single or Not (Dummy, 1 for yes, 0 for no)

$e$  = the error term to represent the influence of all other variables affecting the dependent variable.

This model is estimated via a multiple regression. The influences on  $y$  by the variables represented by  $e$  are predicted to be too small to warrant modelling as additional independent variables. This assumption is based on the theoretical discussion given earlier, outlining the dominant effect from reputation and bandwagons in determining consumer demand for records.

In relation to the model and the discussion above, the following hypotheses can be made:

**H<sub>0A</sub>** : The individual variables ( $x_i$ ) have *no influence* on the dependent variable  $y$  ( $H_{0A}$ :

$$\beta_i = 0)$$

**H<sub>1A</sub>** : The individual variables ( $x_i$ ) *have influence* on the dependent variable  $y$  ( $H_{1A}$ :  $\beta_i \neq$

$$0)$$

**H<sub>0B</sub>** : The variables ( $x_i$ ) has *no influence* combined on the dependent variable  $y$  ( $H_{0B}$ :  $\beta_1$

$$= \beta_2 = \beta_3 = \beta_4 = 0)$$

**H<sub>1B</sub>** : The variables ( $x_i$ ) has influence combined on the dependent variable  $y$  ( $H_{1B}$ :  $\beta_1, \beta_2, \beta_3, \beta_4 \neq 0$ )

In addition, several predictions can be stated about the expected signs, according to the theory, of the different coefficients of the independent variables:

*First Entry Position*: The initial bandwagon effect has positive effect on sales: the smaller the number of the first entry (a lower number is better, indicating a higher ranking in the chart), the smaller is the number for Best Position Reached.

*Position Change from 1<sup>st</sup> to 2<sup>nd</sup> Week*: As people see a positive trend towards purchasing the record, more people will judge it safe to jump on the bandwagon. This implies a positive relationship between the change, measured by subtracting the first week's position from the second week's position (so the bigger the rise, the bigger the negative number), and the Best Position Reached.

*Previous Top 10 or not*: Reputation is predicted to have a positive relationship to sales, thus one would expect a negative relationship between this parameter and the Best Position Reached on the chart for the respective record.

The model is estimated as a sample regression function (SRF), with the objective of reflecting the population regression function (PRF). Ordinary Least Squares (OLS) is used here to calculate the respective coefficient estimates for the variables. This method

will ensure that the estimates of the coefficients are the best, in terms of guaranteeing the residual sum of squares (RSS) is as small as possible ( $\sum e_i^2$ ). This will give several desirable statistical properties, summed up in the Gauss-Markov property of Best Linear Unbiasedness Estimators (BLUE)—i.e., the estimator is a linear function of the sample observations, it is unbiased in that the estimator on average coincide with the true value of the population parameter, and it is the best in that it is the estimate of the parameter with the lowest variance (see for example Gujarati (1995) for more details). This is, however, subject to a set of assumptions including: (1) the explanatory variables are uncorrelated with the error term  $u$ ; (2) the expected value of  $u$  is zero ( $E(u) = 0$ ), or in other words that it has no effect on the dependent variable *on average*; (3) the variance of each separate  $u$  is the same (homoscedacity); (4) there is no correlation between the  $u$ 's (no autocorrelation, or algebraically covariance  $(u_a, u_b) = 0$  with  $a \neq b$ ); and (5) no exact linear relationship exists between the explanatory variables (no perfect multicollinearity). All these assumptions are intended to ensure the validity and quality of the analysis. Their applicability to this model and its data set is explored below in the section covering model estimation and hypotheses testing. The model is summed up in Table 1

**Table 1: Summary of Variables and their Properties**

Dependent	Independent	Label	Coefficient Label	Parameter	Measurement	Minimum/Maximum	Mean	Coefficient of Variation <sup>1</sup> (StDev)
Best Chart Position		<b>y</b>	-	Predicted Sales/Success	Given by a range between 1 and 20, with 1 being the best	1 / 20	8.1	0.73 (5.86)
	Constant/ Intercept	-	$\beta_1$					
	First Entry Position	<b>x<sub>2</sub></b>	$\beta_3$	Initial Bandwagon	Same as above	1 / 20	12	0.5 (6)
	Position Change from 1 <sup>st</sup> to 2 <sup>nd</sup> Week	<b>x<sub>3</sub></b>	$\beta_4$	Trend	Given by subtracting the 2 <sup>nd</sup> week position from the 1 <sup>st</sup> , meaning a better position will be reflected by a negative number	0 / 18 (change, + or -)	3.6 (change, + or -)	1.01 (3.68)
	Previous Top 10 or not	<b>x<sub>4</sub></b>	$\beta_4$	Reputation	Represented by a dummy variable (1 if true 0 if not), meaning the coefficient will only be relevant for true outcomes <sup>2</sup>	0 / 1 (Dummy)	0.48	N.A.

We also experimented with several other variables to represent the reputation effect, including the total number of previous hits on chart, previous highest position for single from artist and previous highest position for album from artist. These variations will be further discussed in the hypotheses testing section below.

<sup>1</sup> Is given by the ration of the Standard Deviation (StDev) to the mean. This can be used to determine the variables variation from the mean. Though there is no hard-and-fast rule about the value of it, a low value such as 0.05 would imply that the StDev is only 5% of the mean, and thus, does not vary enough to represent any significance to the dependent variable.

<sup>2</sup> True means the artist has a reputation as judged by the parameter, while untrue values will be an artist without any reputation represented by a 0

## 5. Data Source

In contrast to previous research using data from the US and UK charts, the data to test the model were retrieved from the official Norwegian Music Chart, 'VG-Lista', or VG-Chart ([www.vg.no](http://www.vg.no)), and are based on a top 20 Single Chart. This is the chart published by the largest newspaper in Norway (*Verdens Gang*), and the basis of several pop-chart shows presented on TV and radio. Its Internet site contains a search engine that can be used to obtain performance data on every single and album that appeared on the chart from 1958 for singles and from 1967 for albums. The VG-Chart is based on record sales from 100 of the highest-selling retail outlets in Norway, and is constructed through a cooperative agreement between *Verdens Gang*, the Norwegian record music trade association (GGF), and NRK, Norway's national broadcasting network and its largest radio and TV company.

A key requirement in Leibenstein's (1950) analysis is that, for consumers to become part of a bandwagon process, they need a means of observing which products other consumers are choosing. In the case of the demand for pre-record popular music in Norway, the VG-Chart provides just such a means and it can reasonably be assumed that in Norway most consumers of music pay attention to the VG-Chart, either directly by looking it up or indirectly through radio or TV. The VG-Chart's dominant position in the Norwegian music scene makes data from it suitable as a basis for testing the model.

The sample used in this study comprises cross-sectional data for all the artists represented on the VG-Chart during year 2000. By going back to 2000, we can insulate our findings from the impact of music downloading sites, with their preview capabilities, on the experience good aspect of music choice. There are a few exclusions, namely,



music from soundtracks, compilations between artists, and records re-entering the chart after more than one year out of the chart. The 170 artists that remain make up a significant sample, well in excess of the 30 that Ramanathan (1998, p. 574) suggests are needed as a minimum for reliable estimates.

## 6 Model Estimation

The model was estimated by Ordinary Least Squares using Shazam Professional Edition. This gave the following sample regression function (SRF) for the singles chart:

$$y = -2.04 + 0.83x_2 + 0.66x_3 + 1.11x_4 + e$$

This function gives the conditional mean value of the dependent variable  $y$ , conditional on the given values of the independent variables  $x_2$ ,  $x_3$  and  $x_4$ . The constant  $\beta_1$  is estimated to  $-2.04$ , and represents the average value of the dependent variable  $y$  when the parameters  $\beta_2$ ,  $\beta_3$  and  $\beta_4$  equal to zero. Since it is a multiple regression, the coefficients of each of the parameters are partial regression coefficients. This means that each of them reflects the partial effect on the mean value of  $y$ , *ceteris paribus* (i.e. the other independent variables are held constant). The coefficient estimates ( $\beta_i$ ) of the independent variables,  $x_2$ ,  $x_3$  and  $x_4$ , above are calculated to be 0.83, 0.66 and 1.11 respectively, each showing a positive linear relationship to  $y$ . The dummy variable  $x_4$ , means that the corresponding coefficient 1.11 will only be included if its value is one, while a 0 will represent a second category, which is not Previously Top 10 in this case. The error term,  $e$ , represents all factors other than the independent variables, and thus is a random unsystematic component affecting the preciseness of the regression.

Having estimated the model, we conducted a number of tests to check its statistical robustness for purposes of hypothesis testing. In hypotheses testing, there is necessary to add one assumption to the classic linear regression model and the estimated coefficient above: the error term ( $e$ ) follows the normal distribution with mean zero and variance  $\sigma^2$ . Alternatively, since the  $\sigma^2$  is commonly not known, the Student's T-distribution can be used. This assumption is based on the Central Limit Theorem (CLT), which postulates that 'if there is a large number of identical distributed random variables, their distribution tends to follow the normal distribution as the number of such variables increases indefinitely' (Gujarati, 1995, p. 161). Further, since the error term follows the normal distribution, a property of this distribution is that any linear function of such a variable is itself normally distributed. It is therefore common for the error term to be assumed to be normally distributed, on stronger and weaker grounds, and thus that the coefficients of the OLS will also be normally distributed. In the present analysis, however, this is problematic, since the variables are based on chart position ranging from 20 to 1, thus they are discrete and limited values. This will mean the hypotheses testing will only be based on an *approximation* of the distribution. .

The t-statistics (and the p-values) for the independent variables estimates  $\beta_1$ ,  $\beta_2$  and  $\beta_3$  are all significant at a 99% significance level, while  $\beta_4$  is significant at a 90% significance level, with all tests two-sided (the estimated t-, p- and f-values and the respective critical values, are provided in Table 2). The other variables tested for the latter parameter (reputation), as mentioned above, were all insignificant.

Testing the effect of all the independent variables collectively, the F-statistic shows the coefficients to be significance at a 99% level.

**Table 2: Summary of Results from the Singles Chart Model**

*Included observations: 170*

Variable	Coefficient	Std. Error	T-Statistic	P-Value
$\beta_1$	-2.0356	0.6700	-3.038	0.0028
$\beta_2$	0.82698	0.4403E-01	18.78	0.000
$\beta_3$	0.66039	0.5000E-01	13.21	0.000
$\beta_4$	1.1091	0.4961	2.235	0.027
$r^2$	0.7168	Durbin-Watson stat		2.1331
Adjusted $r^2$	0.7116	F-Statistic		140.024
S.E. of regression	3.196570	Critical T-Value (99%)		$\approx 2.6$
Sum squared errors	1621.0	Critical F-Value (99%)		$\approx 3.9$

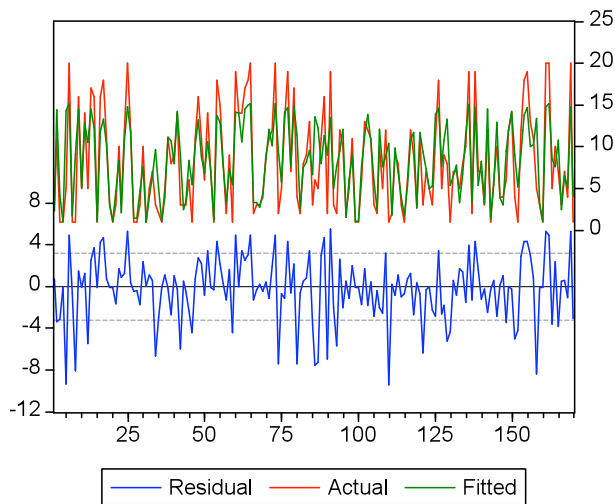
In terms of the hypotheses presented earlier, the significance of the coefficient estimates means the  $H_{0A}$  and  $H_{0B}$  can be rejected, thus the alternative hypotheses  $H_{1A}$  and  $H_{1B}$  can be accepted. In other words, all the independent variables, including First Entry Position, Position Change from 1<sup>st</sup> to 2<sup>nd</sup> Week and Previous Top 10 or not, have an effect on the dependent variable Best Position of Record, both individually and collectively.

In addition to the hypotheses testing, we conducted several tests of the other assumptions of the OLS estimates outlined in the formulation of the model. The two first assumptions, of no correlation between  $e$  and the independent variables and the expected value of the  $e$  is zero, are automatically filled respectively since the estimated regression is conditional, and hence non-stochastic, and since  $e$  is a assumed random its negative and positive expected impact on  $y$  will on average cancel out. For the other assumptions, some tests and general judgement are necessary to confirm their validity.

Perfect multicollinearity—i.e., cases of perfect linearity between the independent variables—would make it difficult to obtain unique estimates of the different parameters.

This is, however, seldom observed; rather, we get degree of multicollinearity, which will not compromise the BLUE quality of the variables, but may affect the quality of the estimates. Referring to the model, this could have been a problem with having two similar parameters,  $\beta_2$  and  $\beta_3$ , for the different dimensions of the bandwagon effect. There was, however, no sign of it in the regression output given in Table 2. In addition, the pairwise correlation between the independent variables, the subsidiary regressions or the Variance Inflation Factors (VIF) used to check for multicollinearity, showed little evidence of it.

Heteroscedasticity, or unequal variance, is common in cross-sectional data, and thus is relevant to test for in relation this model. In terms of the estimates being BLUE, this can give inefficient results, or in other words, the estimates might not have minimum variance, thus making the conclusions unreliable. In relation to the model, it may, for example, be that the best position for a record will vary more for entries at a higher position, than for one entering at a lower position as there may be differences in a record's fluctuation potential from its expected value. Both the Park test and the Glesjer test were applied to the model, both giving results of insignificant statistical relationships between the different functional forms of the error terms and the regression. This means there is no evidence of heteroscedasticity. In general, this procedure tests whether the  $e$  is systematically related to the independent variables. This can also be detected by examining a graphical output of the residuals  $e$  shown in Figure 1 (obtained via E-Views software) and look for patterns in the data.



**Figure 1: Residuals ( $e$ ) of the Regression for the Singles Chart Model**

No clear patterns can be observed in Figure 1, indicating the residuals are random. This is also relevant to dismissing autocorrelation, which also would mean that the residuals are correlated and thus would not be random. As with heteroscedasticity, autocorrelation will render the OLS estimates inefficient, and thus not BLUE. Another way of testing this is the Durbin Watson Statistic given in Table 2, which should be close to  $2^3$ , which it is, thus no autocorrelation is concluded to be present.

Finally, the adjusted  $r^2$  measures the goodness of fit of the estimated sample regression line, by giving the proportion of the total variation in the dependent variable explained by the independent variables, of the model. It is adjusted from  $r^2$  in the sense that it takes into account the degrees of freedom (d.f.) in the model so the model's goodness of fit will not increase automatically with the number of independent variables.

<sup>3</sup> More specifically this value should be in a certain interval to reject autocorrelation, given the number of observations and the explanatory variables excluding the intercept, this intervals approximately 1.8 and 2.2 for the model.

The model gives an adjusted  $r^2$  of about 71% (Table 2), which means it explains a good portion of the variation in the dependent variable  $y$ .

## 7. **Model Adequacy and Consistency**

To explore the applicability of the model it was also applied it to the Top 40 albums chart from the same source. In this case, however, the sample was of 164 observations, including all the artists represented on positions 40 to 1 on chart for the first six months of year 2000. Since we are dealing with albums, and not singles, two additional independent variables were added to the regression, including ‘the Number of Singles on Chart from the Album Before its Best Position’, and a dummy variable for ‘Best of Album or Not’. The former was included because of the effect singles might have on the album sales, as they are usually released before the album. Singles might have a positive effect on album sales due to their airplay increasing consumers’ knowledge about the album’s contents, but could also be expected to exert a negative effect if they were a substitute for buying the album. The latter variable was included since 14 percent of the chart consisted of ‘best of’ albums, which were considered ‘special cases’ of pure reputation effects. Negative signs would be expected for both of these additional independent variables since if they help the sales of an album this will result in a lower number for the position it reaches in the chart.

The other independent variables were expected to yield much the same result as the regression on the single chart, thus supporting its applicability. This is confirmed in the estimated regression:

$$y = -1.22 + 0.77x_2 + 0.47x_3 + 2.36x_4 - 2.78x_5 - 5.44x_6 + e$$

in which the additional variables are:

$x_5$  = Number of Singles on Chart from the Album Before its Best Position;

$x_6$  = Best of Album or Not (Dummy, 1 for yes 0 for no).

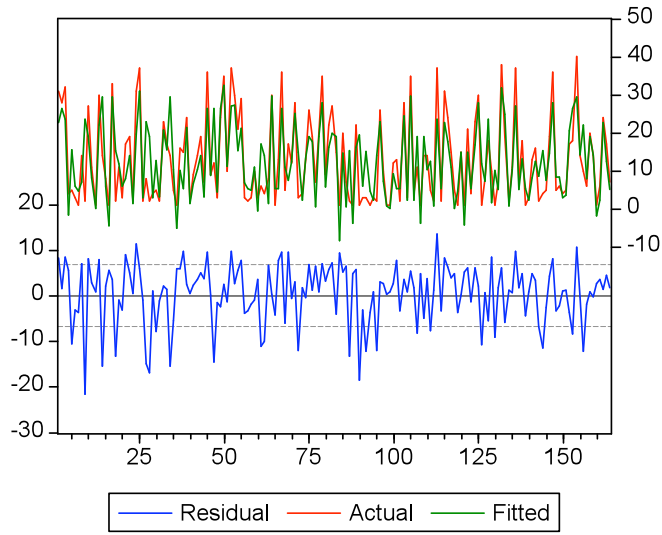
As regards to the testing of the model, the results were very similar to the singles chart regression, with the reputation parameter giving the same result, in terms of t-statistic, for the same variables. Both the new variables were significant. A summary of the results and a graphical plot of the residuals of the model are presented in Table 3 and Figure 2, respectively.

**Table 3: Summary of Results from the Albums Chart Model**

*Included observations: 164*

Variable	Coefficient	Std. Error	T-Statistic	P-Value
$\beta_1$	-1.2151	1.1190	-1.0859	0.2792
$\beta_2$	0.77030	0.39506E-01	19.498	0.0000
$\beta_3$	0.46558	0.57982E-01	8.0298	0.0000
$\beta_4$	2.3638	1.0558	2.2388	0.0266
$\beta_5$	-2.7780	0.74758	-3.7160	0.0003
$\beta_6$	-5.4446	1.5632	-3.4831	0.0006
$r^2$	0.7464	Durbin-Watson stat		2.1695
Adjusted $r^2$	0.7384	F-Statistic		93.018
S.E. of regression	5.6938	Critical T-Value (99%)		$\approx 2.6$
Sum squared errors	5122.2	Critical F-Value (99%)		$\approx 3.15$

**Figure 2: Residuals (e) of the Albums Chart Regression**



In addition, two separate regressions were run on the singles chart data using the same independent variables as the initial models, but with number of weeks on chart, and number of weeks on chart multiplied by the (inverse of) best position to reflect total sales. In terms of positive sales, both gave a positive relationship to the bandwagon effects, and a negative relationship to reputation. The regressions are given in Tables 4 and 5.

**Table 4: Summary of Results from the Model using Number of Weeks Multiplied by the Best Position on Chart as the Dependent Variable**

*Included observations: 170*

Variable	Coefficient	Std. Error	T-Statistic	P-Value
$\beta_1$	236.51	12.879	18.364	0.0000
$\beta_2$	-9.8	0.84642	-11.578	0.0000
$\beta_3$	-12.241	0.96104	-12.737	0.0000
$\beta_4$	-22.747	9.5366	-2.3852	0.0582
$r^2$	0.5775	Durbin-Watson stat		1.8238
Adjusted $r^2$	0.5699	F-Statistic		75.631
S.E. of regression	60.066	Sum squared errors		59893



**Table 5: Summary of Results from the Model using Number of Weeks in Chart as the Dependent Variable**

*Included observations: 170*

Variable	Coefficient	Std. Error	T-Statistic	P-Value
$\beta_1$	12.209	0.66266	18.424	0.0000
$\beta_2$	-0.39170	0.43552E-01	-8.9939	0.0000
$\beta_3$	-0.58755	0.49449E-01	-11.882	0.0000
$\beta_4$	-0.96000	0.49070	-1.9564	0.0521
$r^2$	0.5095	Durbin-Watson stat		1.7824
Adjusted $r^2$	0.5006	F-Statistic		57.479
S.E. of regression	3.0907	Sum squared errors		1585.7

## 8 Discussion

### 8.1 Original Hypotheses and Discussion of Unexpected Results

Both hypotheses  $H_{0A}$  and  $H_{0B}$  were rejected against hypotheses  $H_{1A}$  and  $H_{1B}$ , meaning the independent variables all have a significant influence on the dependent variable, both individually and jointly. The effect of bandwagons in the record market, and the charts' effect on this is supported by the fact that (a) First Entry ( $x_2$ ) is positively related to Best Position, meaning the lower number (= better) the entry position the lower (= better) the expected best position and vice versa; and (b) Position change ( $x_3$ ) is positively related to Best Position, meaning the larger the negative change in positions towards number one from week one to week two, the lower (i.e. better) is the expected Best Position.

However, while both the estimated coefficients  $\beta_2$  and  $\beta_3$  were positive as expected,  $\beta_4$  shows a positive relationship to  $y$ , which was not as expected and implies that past chart success by a particular artist hindered the highest position that the artist's

subsequent recording reached in the charts. This is, however, only on a 90% significance level, which is arguably not high enough.

Our predictions about the role of reputation were better supported when we separately regressed the Best Position on the Number of Previous Singles on Chart and Number of Previous Albums on Chart. In both cases there was the predicted negative relationship but it was statistically insignificant, just as including these variables in the regressions only has insignificant statistical impacts. The weak negative relationship between highest position and the number of previous recordings in the chart may at the same time deter one from concluding that our unexpected results with the reputation effect could be the result of diminishing returns to reputation. This might reasonably be expected to arise due to consumers experiencing diminishing marginal utility because artists they had previously favoured failed to do anything particularly novel on their later albums or because artists are prone eventually to experience declining popularity due to becoming seen as no longer 'cool' and appearing 'out of date' or out of creativity.

The unexpected sign of the reputation effect above should be seen in relation to the composition of the record charts, which mainly consist of established artists in the first place. This can also explain why it was difficult to get significant parameters for the reputation effect. Not only might the relationship be weakened because of diminishing returns in terms of demand, from increases in reputation but the effect of reputation will then be neutralized if most artists making it into the top 40 chart are sufficiently well known to benefit from this effect anyway. The latter possibility seems plausible in the light of Table 6, which shows the composition of artists on the charts and also suggests that the different kinds of artists enjoy success on the two kinds of charts.

**Table 6: Established artists in the Charts**

	ALBUM CHART		SINGLE CHART	
	Number	%	Number	%
<b>Prev top 10 album</b>	80	0.49	54	0.32
<b>Prev top 20 album</b>	90	0.55	58	0.34
<b>Prev top 10 single</b>	67	0.41	90	0.53
<b>Prev top 20 single</b>	70	0.43	82	0.48

Both of the added independent variable of the album chart regression have, as expected, negative relationships to the Best Position variable, indicating that the number of single releases ( $\beta_5$ ) and whether the album is a best of album ( $\beta_6$ ) are both good for the success of the album. The former estimate ( $\beta_5$ ) can be an indication of the reduced uncertainty of buying the album, as the consumer will know more of the songs, and thus the content of the album. The latter estimate ( $\beta_6$ ) can be seen as an indication of the pure reputation effect.

It should also be noted that although the experience good problem is the main reason for expecting reputation to be important, the unexpected result may partly reflect the fact that this problem is not as acute for the context of the study—namely, the performance of recordings that make it into the upper league of the sales charts—compared with the broader context of choosing records that have not yet made any inroads to this league (for example, having only reached somewhere between 21 and 100 in the top 100 singles (between 41 and 100 in the case of albums) , or yet to register in the top 100). Once a single has got a foothold in the top 20 charts, its ‘A-side’ will have been, and will currently be being, heard frequently on the radio or experienced via

recordings purchased by members of social networks. Albums will remain more of an experience good but can also be experienced vicariously in a social setting and assessed via published reviews. For recordings that have not yet hit the charts or received airplay, it still seems likely that reputation will be crucial in getting their sales started unless they are heavily promoted, given excellent reviews by respected music journalists and/or purchased by socially influential ‘hip’ consumers.

## 8.2 *Limitations of the study and some possible extensions*

A limitation of this study is that because of its limited and discrete variables the probability distribution is assumed an approximate to a normal distribution. This is an assumption, which should have been supported by econometric evidence, had it been available. The problem, and a challenge for potential extensions of the model, is to obtain the corresponding sales data to each of the position from 20 to 1, and use ranking models to get the probability distribution. This would require, however, that the sales figure (or range) corresponding to each position of the chart is static from week to week, something that it would be unwise to assume given the widespread reports of seasonality in the demand for recorded music, with a peak around Christmas. Given this, it could be worthwhile to try a model specification, based on Ordered Multiple Choice Models, like the following:

$$y_{it} = \text{sales by artist } i \text{ in period } t \text{ (unobserved)} = f(\text{lagged } y_{it}, x1_{it}, x2_{it}, \dots, xk_{it}) + e$$

$$r_{it} = \text{Rank of Artist } i \text{ in period } t \text{ (unobserved)}$$

This gives:

$$r_{it} = 1 \text{ if } y_{it} > y_{jt} \text{ for all } j \text{ not equal to } i$$

$r_{it} = 2$  if ...

Then,

Prob ( $r_{it} = 1$ )

Prob ( $r_{it} = 2$ )

After obtaining this, the likelihood-function could be found and by maximising it the probability distribution could be found. Alternatively, an expression for the expected value of  $r_{it}$  could be found, before OLS to estimate this expression. An alternative to the dependent variable could be to trace the entire lifecycle of each of the units of the sample week by week, and thus get a better estimate of actual sales. This would, however, be a very time demanding task consider the size of the sample. In addition, this still would not be exact, because of the problem of different sales numbers being recorded for the same position every week.

Despite its present limitations, the study at least offers empirical evidence of the bandwagon effect in the record market and gives insight into the reputation effect that record company executives might find useful. The strength of the study is further supported by a significant sample size, and its consistency shown when applied to the two different charts. Further research could also apply the model to other years, to check for the same results. Tracing the life cycle of records by using the data given by this chart could also be useful in future research. It might further be worthwhile to tag artists by their career lifecycle stage and their music genre and then see what kinds of clusters could be observed in chart performance. Finally, we note that it might be possible to obtain play-lists for radio stations and explore the extent to which the entry, and point of

first entry into the charts is shaped by the frequency with which it has been ‘plugged’ on air. Ethnographic research can also be conducted to examine the network bandwagon effects that have been ignored in the empirical part of this paper.

## References

- Adler, M. (1985). Stardom and talent. *American Economic Review*, 75, 208-212.
- Bhattacharjee, S., Gopal, R.D., Lertwachara, K., Marsden, J.R. and Telang, R. (2005). The effect of digital sharing technologies on music markets: A survivor analysis of albums on ranking charts. Working paper available at <http://www.heinz.cmu.edu/~rtelang/survival.pdf> (accessed 14 October 2008).
- Bhattacharjee, S., Gopal, R.D., Lertwachara, K. and Marsden, J.R. (2006). Consumer search and retailer strategies in the presence of online music sharing. *Journal of Management Information Systems*, 23(1), 129-159.
- Bhattacharjee, S., Gopal, R.D., Lertwachara, K. and Marsden, J.R. (2007). Stochastic dynamics of music album lifecycle: An analysis of the new market landscape. *International Journal of Human-Computer Studies*, 65(1), 85-93
- Bradlow, E.T. and Fader, P.S. (2001). A Bayesian lifetime model for the hot 100 Billboard songs. *Journal of the American Statistical Association*, 96, 368-381.
- Chai, A., Earl, P.E. and Potts, J. (2007). Fashion, growth and welfare: An evolutionary approach. In M. Bianchi (ed.) *The Evolution of Consumption: Theories and Practice. Advances in Austrian Economics*, 10 (pp. 187-207), Oxford: Elsevier.

- Chung, K.H. and Cox, R.A.K. (1994). A stochastic model of superstardom: an application of the Yule distribution. *Review of Economics and Statistics*, 76, 771-775.
- De Vany, A. (2004). *Hollywood Economics: How Extreme Uncertainty Shapes the Film Industry*, London and New York: Routledge.
- Earl, P.E. (1983). The consumer in his/her social setting: A subjectivist view. In Wiseman, J. (ed.) *Beyond Positive Economics?* (pp. 176-191) London: Macmillan.
- Earl, P.E. and Potts, J. (2004). The market for preferences. *Cambridge Journal of Economics*, 28, 619-633.
- Fisbein, M.A. and Ajzen, I. (1975). *Belief, Attitude, Intention and Behavior*. Reading, MA: Addison-Wesley.
- Fox, M.A. and Kochanowski, P. (2004). Models of superstardom: An application of the Lotka and Yule distributions. *Popular Music and Society*, 27(4), 507-522.
- Giles, D.E. (2006). Superstardom in the US popular music industry revisited. *Economics Letters*, 92(1), 68-74.
- Giles, D.E. (2007). Increasing returns to information in the US popular music industry. *Applied Economics Letters*, 14(5-6), 327-331.
- Gujarati, D.N. (1995). *Basic Econometrics* (3rd edition), New York: McGraw-Hill.
- Holbrook, M.B. (1995). *Consumer Research: Introspective Essays on the Study of Consumption*, Thousand Oaks, CA: Sage.
- Leibenstein, H. (1950). Bandwagon, snob, and Veblen effects in the theory of consumers' demand. *Quarterly Journal of Economics*, 64, 183-207.

- MacDonald, G.A. (1988). The economics of rising stars. *American Economic Review*, 78, 155-166.
- Nelson, P. (1970). Information and consumer behavior. *Journal of Political Economy*, 78, 311-329.
- Ramanathan, R. (1998). *Introductory Econometrics with Applications*, Singapore: South-Western.
- Rosen, S. (1981). The economics of superstars. *American Economic Review*, 71, 845-858.
- Rothschild, M. (1973). Models of market organization with imperfect information: A survey. *Journal of Political Economy*, 81(6), 1283-1308.
- Strobl, E.A. and Tucker, C. (2000). The dynamics of chart success in the UK pre-recorded popular music industry. *Journal of Cultural Economics*, 24, 113-134.