Music and the Market: Song and Stock Volatility

Abstract: Popular music may presage market conditions because people contemplating complex future economic behavior prefer simpler music, and vice versa. In comparing the annual average beat variance of the songs in the US Billboard Top 100 since its inception in 1958 through 2007 to the standard deviation of returns of the S&P 500 for the same or the subsequent year, a significant negative correlation is observed. Furthermore, the beat variance appears able to predict future market volatility, producing 2.5 volatility points of profit per year on average.

Keywords: music; complexity; volatility; strategy; behavioral

Acknowledgments: I am grateful to Gabriele Lepori, Marcial Losada, participants at the 2011 Annual Meeting of the Academy of Behavioral Finance & Economics, and the anonymous referees for useful comments and suggestions, and to Larry Ribstein for a useful reference.

Philip Maymin
Assistant Professor of Finance and Risk Engineering
NYU-Polytechnic Institute
Six MetroTech Center
Brooklyn, NY 11201
(718)260-3175
phil@maymin.com

Music and the Market:
Song and Stock Volatility

Abstract: Popular music may presage market conditions because people contemplating complex future economic behavior prefer simpler music, and vice versa. In comparing the annual average beat variance of the songs in the US Billboard Top 100 since its inception in 1958 through 2007 to the standard deviation of returns of the S&P 500 for the same or the subsequent year, a significant negative correlation is observed. Furthermore, the beat variance appears able to predict future market volatility, producing 2.5 volatility points of profit per year on average.

1. Introduction

Social mood permeates both our choice of music and our financial markets. This paper explores the possibility of a link between properties of popular songs and measures of broad-based financial markets.

Externally imposed mood has often been found to affect the market. Certain weather-related investigations suggest wind (Keef and Roush, 2005) or sunshine (Hirshleifer and Shumway, 2003) directly affect mood and then indirectly the stock returns on the same day.

Music, however, is not imposed externally and unpredictably, but internally and through voluntary consumer purchases. Internally imposed mood has also been found to be related to the market: Lepori (2010) examines the results of endogenous mood on the market by comparing weekend comedy movie attendance with the corresponding weekend returns and finds that increased attendance corresponds with decreased returns.

In each case, a psychological explanation links mood with market returns, and neuroeconomics research has focused on isolating areas of the brain responsible for the particular moods (cf. Camerer, Loewenstein, and Prelec, 2005). Music may provide a better indicator of mood because there is no single music center in the brain (Sacks, 2008). Indeed, listening to and appreciating music seems to activity just about every part of the brain, and there is a long history of human evolutionary adaptation to music (Levitin, 2007).

Music directly affects mood in multiple ways. Bruner (1990) lists thirteen emotions ascribed to various musical components. For example, faster tempos generate animation, reed instruments generate melancholy, higher keys generate happiness, and complex harmonies produce agitation. Pearce, Ruiz, Kapasi, Wiggins, and Bhattacharya (2010) suggests that it is the unpredictability of music that drives our attention and appreciation.

Higher volatility proxies for unpredictability and complexity, both in music and in the market. In songs, we can use a measure called beat variance, literally the computed variance of the automatically extracted beats of the song, using a free online music analysis website called EchoNest.com. Simpler songs tend to have low beat variance while more complex songs have higher beat variance. Similarly, in market environments where pricing is difficult and complex, realized volatility is likely higher, while environments in which pricing appears straightforward will likely have lower realized volatility.

1.1. Variance

Previous studies that attempt to link some measure of mood with some measure of the market use the first moment: the average mood and the average return. By contrast, the measure we study here is the second moment, namely the variance, primarily because it reflects the underlying complexity, and also because it has the following three attractive properties.

First, the variance is a more precise measure than the average with much smaller statistical noise. As shown by Merton (1980), the estimate of the average market return for a given year does not improve with higher frequency, but the estimate of variance does. The same holds true for the average and the variance of the beat. Therefore, using variance lets us look at higher quality data.

Second, the variance does not itself vary as much; for example, when the market is volatile, it tends to remain volatile. Similarly for mood: when we are feeling emotionally volatile in one second, we are likely to continue to be volatile the next, whereas the average mood we have been experiencing has been fluctuating wildly.

Third, the variance is more representative of the character of the entire year or song than the average. A single month or a single substantial change in the beat of a song could have a significant impact on the average without reflecting the overall mood. For example, an eleven-month bear market followed by a December rally to end the year positive overall would not have felt like a bull market for most of the year to stock market participants. But a market that is volatile for 11 months and then steady would still be coded as volatile for the year, and indeed people would later remember it as a volatile year.

1.2. Complexity

This paper investigates the question of whether more complexity in music on Top 100 Billboard songs leads to less future complexity in the market through lower subsequent realized market volatility, because the complexity of popular music reflects the mood and the choices made by economic actors in light of their cognitive load with respect to future activity.

Specifically, people engaged in complex tasks prefer simpler music, and those engaged in simple tasks prefer complex music. Konečni and Sargent-Pollock (1976) find that people performing complex tasks have a tendency to choose less complex music. This choice is driven by the realization that listening to complex music while performing complex tasks is detrimental to performance. North and Hargreaves (1999) investigate the effects of listening to complex music while playing a car racing game and found that more complex music yielded worse results. Furnham and Allass (1999) find that more complex background music caused introverts to perform worse in observation and memory tests.

Dibben and Williamson (2007) report on a survey of 1,780 British drivers and their music listening habits. One of their results involve the kind of music people were listening to at the time of their last accident. They report the incidence of genres at the time of the last accident relative to overall incidence of genres. Unlike North and Hargreaves (1999) in which music was forced on drivers and more complex music resulted in more accidents, in the real world people can choose what music they listen to. Indeed, Dibben and Williamson (2007) report that music was playing in less than a quarter of accidents, and they infer therefore that "there is no direct link between the presence of music while driving and involvement in an accident." In other words, the musical choices were not themselves the cause of the risky driving, but rather reflected preferences by the drivers conditional on the complexity or risk of the driving task.

From our perspective, we can interpret the existence of accidents as being the result of riskier, more complex driving behavior. We would predict that drivers would have chosen simpler music because they were engaged in more complex tasks. Thus, we would predict that the music listened to at the time of the last accident would be simpler than usual. Dibben and Williamson (2007) report that this is indeed the case. Young drivers listened to dance/house music, which is typically low complexity, significantly more often at times of accidents than usual. In addition, they listened to indie/rock/punk music, which is typically high complexity, significantly less often at times of accident than usual. Middle aged drivers also listened to indie/rock/punk music less often at times of accident than they usually did. They were also less likely than usual to listen to classical music and chart-pop. Classical music has high complexity, but chart-pop would seem to be a counterexample to our prediction. However, the survey by Dibben and Williamson (2007) was conducted in 2005, which had the second highest average beat variance of any year since 1977. In other words, middle aged drivers were listening to the relatively complex chart-pop music of 2005 even less frequently during accidents than they otherwise would, conforming to the prediction that drivers in riskier situations would choose to listen to simpler music. Finally, elderly drivers also listened to the more complex classical music less frequently at the time of an accident. Overall drivers listened to simpler music more often, and complex music less often, when their more complex and riskier driving behavior resulted in accidents.

Just as drivers engaging in more complex and riskier driving activity choose simpler music, so too do we predict that people engaged in more complex and riskier economic activity would choose simpler music. These choices should then in turn lead to future market volatility, not because of the direct effects of the music, but because the musical choice was an early indication of their economic planning, just as musical choices by drivers did not cause the accidents, but because the choice of music reflected the riskiness of the behavior.

In determining future action, economic actors, including not only investors but also entrepreneurs, managers, executives, and politicians, can either contemplate complex possibilities, involving innovation and risk, or simple possibilities. Lee and Shields (2011) point out that particularly in a recession, the decision-making of virtually all economic actors are affected for a protracted period. Thus, the drivers of the financial markets are not merely financial traders, but all economic decision makers.

When they tend to contemplate more complex possibilities, they will prefer simpler music; further, contemplating more complex possibilities is linked to an increase in future risky activity, thus leading to future market volatility. In this way, popular preference for simple music today predicts turbulent market activity in the future, while popular preference for complex music today predicts relatively calmer market activity in the future. Finally, musical preferences explain more than can be explained simply by mean reverting market volatility: popular music decisions contain additional information.

1.3. Music and the Market

The literature both in psychology and in finance contain few references to any articles linking music and the financial markets. Previously, there was no easy way to collect songs nor an automatic way to categorize them. Even today, there is no single reliable source of digital music files nor a single accepted method for categorization. Of the few papers, described and distinguished below, that did research the link between music and the markets, none examined the role of the variance.

Zullow (1991) analyzed the lyrics of top 40 songs from 1955 to 1989 for depressive characteristics and found that it predicted, first, similar depressive characteristics in media stories one or two years later, and, second, predicted personal expenditure and GNP growth two years later. In contrast to a lyrical analysis, this paper analyzes the sound of the music, and instead of economic variables, considers financial variables. Most importantly, this paper considers the variance, both of the song and of the market, rather than the moving average of the level of the mood or changes in GNP.

Crain and Tollison (1997) analyze the time series of 921 unique weekly #1 Billboard songs from 1940 to 1988 and develop a model linking broad economic variables such as the real prime interest rate, real weekly earnings, the growth rate in real personal income, and the unemployment rate plus the inflation rate to characeristics of the songs, in particular its length, beats per minute, and key. They created their data set through the efforts of a trained musician who "obtained copies of the sheet music and listened to recordings of the songs with a metronome."

They conclude two things: first, they identify three major regime changes in the structure of the songs themselves, from 1940-1955, 1956-1964, and 1965-1988 and, second, that economic forces play a role in predicting the length and beats-per-minute of popular songs, with better results for length.

This paper differs from Crain and Tollison (1997) in several respects. The data set begins later (in 1958), lasts longer (to 2007), has more songs (five thousand), and instead of using a trained human musician, relies on an automated analysis of the music using a publicly available service. This paper also looks at the second moment of the beat (beat variance) rather than the length and the average beats per minute. Furthermore, below we explore the link between the beat variance and the volatility of the financial markets, as well as a trading strategy. These are all novel results.

There is some recent survey data linking music with personal finances. North and Hargreaves (2007) ask participants, along with their musical preferences and many other questions, if they own shares in a company, have more than one bank account, or have a credit card, and interpret "yes" responses as a proxy for access to financial resources. They report that "fans of DJ-based music, hip hop/rap, and dance/house had the lowest level of access to these financial resources whereas fans of adult pop and classical music had the highest." For our purposes, this finding helps bridge the possible gap between those who invest and those who listen to popular music: the results of North and Hargreaves (2007) suggest significant overlap between the two groups, especially in light of the discussion in Section 2 of changes in the Billboard Top 100 to reflect more pop songs.

Most recently, Pettijohn and Sacco (2009) linked characteristics in top Billboard songs with a measure of the market. Specifically, they found a relationship between song characteristics as determined by questionnaires on human raters listening to the music and the General Hard Times Measure, which they describe as a "standardized, global measure consisting of the US unemployment rate, change in disposable personal income, change in consumer price index, death rate, birth rate, marriage rate, divorce rate, suicide rate, and homicide rate." The kind of music characteristics the raters measured included the average pace of the song and whether it felt meaningful and was comforting or romantic.

This paper differs from Pettijohn and Sacco (2009) by using entirely objective and easily computable measures of the songs and the markets, and by focusing on the second moment rather than the first. Instead of asking if the average pace of the song is fast or slow, we ask if the pace changes a lot or a little. Instead of asking about broad economic trends, we ask if the broad US market bounces around a lot or a little. Additionally, we ask if the musical indicators cause or are caused by the financial indicators.

2. Data and Analysis

The Billboard Hot 100 is the industry standard popularity chart of US singles. It is issued weekly and also has a special year-end chart.

There are several limitations of the methodology (c.f. Whitburn, 2009). First, it may have been subject to various skews during the 1950s and 1960s as payola (illegal payment by record companies to secure radio broadcasts). Second, artists that have had popular albums with no particular popular single (such as Pink Floyd or Led Zeppelin) may be absent from the charts. Third, singles have become a less common form of song release over time, leading to recent songs more frequently remaining in the charts than in the past. Fourth, the popularity of the measure itself has led to strategizing by artists to time their debut to maximize the chances of topping the chart. Fifth, the methodology has undergone several changes over the years, from incorporating data from Nielsen SoundScan, to allowing songs unreleased as separate singles to join the chart, to incorporating data from the downloads of the songs.

Despite these limitations, the chart has historically served as the best measure of popularity. Here, we use it as a proxy for the songs people were listening to.

In 2005, Billboard launched a new chart called the Pop 100 to address concerns about the dominance of R&B and Hip Hop music on the Hot 100 chart. According to the methodology listed on Billboard's website, the primary difference between the Hot 100 and the Pop 100 is that the latter "confines its radio panel to mainstream top 40 stations." Therefore, the Hot 100, used here, is still probably the preferable chart for our purposes of measuring popular songs.

Archived charts are available for purchase directly from Billboard and copies can be found on various websites. Billboard also sells compilation CDs with the most popular songs of each year, but they typically only have ten of the top 100 songs of the year on the CD. Various websites provide all 100 hits from each year for download.

These 100 songs in MP3 format were accumulated for each of the 50 years from 1958 through 2007. For some reason, 1969 and 1995 each had 101 songs – these are included in the averages. Thus the total number of songs is 5,002.

A “music intelligence company” called The Echo Nest, officially launched in September 2008, provides an automatic service to analyze any song and output a "musical score for computers," describing the song's structure and musical content, including rhythm, pitch, and timbre, in an XML file.

Each of the 5,002 songs were uploaded via their API and the resulting XML files were downloaded. The files include a vast quantity of data both for the overall song and for automatically determined segments. The only data we are interested here is the overall beat variance of the song. Other calculated data fields include the loudness, key, time signature, beats, bars, sections, and tempo.[1]

For the market variance numbers, the annual volatility of the S&P 500 index for a particular year was calculated as the standard deviation of its daily returns multiplied by the square root of 252, the standard estimate of the number of trading days in a year. The S&P 500 is routinely chosen as being representative of the broad U.S. market index because it contains more stocks, including non-industrial companies, than the Dow Jones Industrial Average, and though it has far fewer stocks than the more comprehensive Russell 3000, its

---

[1] A regression across all 5,002 songs of the beat variance on measures of loudness, mean pitch, pitch variance, tempo, mean timbre, and timbre variance yields an R^2 of less than eleven percent, indicating that beat variance is a distinct musical feature that cannot be

components tend to be liquid and the index itself is widely tracked. Further, its volatility is nevertheless highly correlated with the others; indeed, the correlation between the contemporaneous monthly or annual rolling volatilities of any pair of those three indices over their overlapping history exceeds 97 percent.

There is a slight timing discrepancy between the calendar year and the Billboard chart release year: Billboard calculates its chart based on information from the beginning of December to the end of November. Thus, the market volatility was calculated on a matching time basis, so the market volatility for chart year "1958" is calculated using data from December 1, 1957 through November 30, 1958. Later sections discussing a trading strategy calculate annual volatility on a calendar basis.

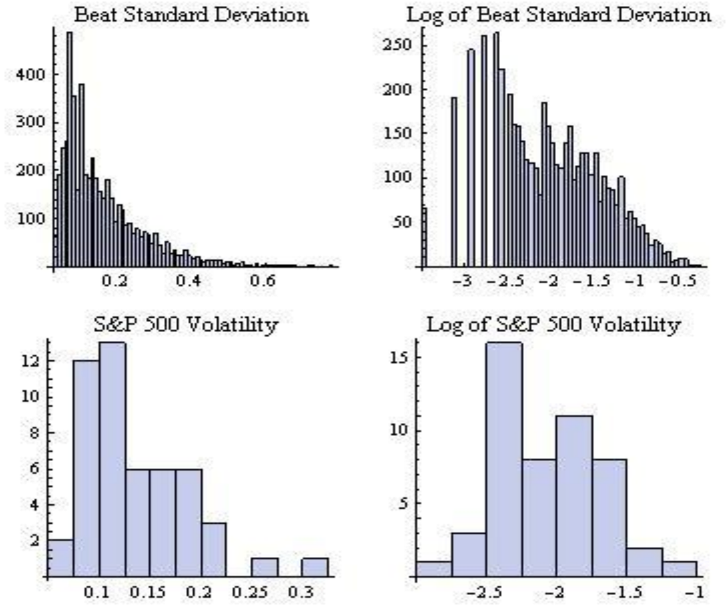
3. Results and Discussion

Table 1 lists summary statistics about the song data and the market volatility. Figure 1 compares the distribution of the song volatility and the market volatility. The histogram of the beat standard deviation (the square root of the beat variance) across all 5,002 songs shares similar features with the histogram of the annual market volatility across all 50 years: both exhibit a high frequency of relatively low volatility with tails to both sides. The histograms of the logarithms of the two numbers helps expand the left-hand-side tail where the numbers are otherwise packed into a small single histogram bar.

Table 1: Summary Statistics. This table lists the summary statistics for the beat variance and the market volatilities from 1958 through 2007. The song beat variance column summarizes across all songs while the annual beat variance column summarizes across the annual average beat variances.

	Beat Variance (Song)	Beat Variance (Annual)	Market Volatility
Minimum	0.0010	0.0131	0.0516
Median	0.0160	0.0351	0.1192
Mean	0.0368	0.0368	0.1353
Standard Deviation	0.0547	0.0143	0.0520
Maximum	0.5990	0.0765	0.3132

Figure 1: Histograms. The histograms on the left show the distribution of the beat standard deviations across all songs and the market volatility across all years, while the histograms on the right show the distributions for the natural logarithm of each, to help expand the left hand side tails. In both cases, the music and the market volatility seem to exhibit roughly similar distributions. The music volatility histograms have more granularity because there are 100 times more data points.



To better understand beat variance, we can examine which artists have consistently low or high numbers. Among artists with at least five appearances on the charts, Ray Charles had the highest average beat variance of 0.1298, ranging from a low of 0.0390 to a high of 0.2860 in his eleven hits. Other artists who averaged a high beat variance are Barbra Streisand, Bobby Vinton, Alicia Keys, and Alice Cooper. These artists sang relatively volatile songs. At the other end of the list, Billy Idol had the lowest average beat variance of 0.0034, ranging from a low of 0.0020 to a high of 0.0050. Other artists who averaged a low beat variance are Ace of Base, Genesis, Al Green, and Bobby Brown. These artists sang relatively stable songs.

The lowest beat variance of any song was 0.0010. Sixty-five songs achieved this lowest level, including A-ha's 1985 hit "Take On Me." There were no such songs prior to 1976, and only thirteen distinct years boasted more than one such song in its top 100. The counts, listed in table 2, show that 39 of the 65 low beat variance songs were hits from the 80s, specifically, 1982 through 1989. With the exception of 1981, 1990, and 1991, every year from 1979 through 2001 had at least one such low beat variance song on its top 100 chart. Only three years outside of that range had such songs: 1976, 2006, and 2007. A-ha's "Take On Me" occurred in the year with the most such low beat variance songs, two years before the high volatility market crash of 1987. Table 2 also reports the excess kurtosis of S&P 500 returns for the associated calendar year as a corresponding extreme measure of complexity for the market. The years with the highest number of the steadiest songs tended to occur around years with the fattest tails in market returns.

Table 2: Songs with the Lowest Beat Variance. This table lists the number of songs having the lowest beat variance in the Billboard Top 100 for each year from 1976 through 2007 as well as a visual boxplot. All but three of the songs with the lowest beat variance occurred from 1979 through 2001, peaking in 1985. Also reported is the excess kurtosis of market returns for the corresponding year.

Year	Steadiest Songs	Boxplot of Steadiest Songs	Excess Kurtosis	Boxplot of Excess Kurtosis
1976	1	■	0	
1977	0		0	
1978	0		2	■■
1979	1	■	2	■■
1980	1	■	0	
1981	0		0	
1982	2	■■	2	■■
1983	2	■■	0	
1984	2	■■	1	■
1985	11	■■■■■■■■■■■	0	
1986	6	■■■■■■	3	■■■
1987	4	■■■■	43	■■■
1988	5	■■■■■	7	■■■■■■■
1989	7	■■■■■■■	12	■■■■■■■■■■■■
1990	0		1	■
1991	0		2	■■
1992	1	■	0	
1993	7	■■■■■■■	2	■■
1994	2	■■	1	■
1995	1	■	1	■
1996	2	■■	2	■■
1997	3	■■■	6	■■■■■■
1998	1	■	4	■■■■
1999	2	■■	0	
2000	1	■	1	■
2001	1	■	1	■
2002	0		1	■
2003	0		1	■
2004	0		0	
2005	0		0	
2006	1	■	1	■
2007	1	■	1	■

The highest beat variance of any song was the 1958 Poni Tails hit "Born Too Late," which had a 0.5990 beat variance. Only two other songs had beat variances exceeding

0.5000: the 1992 hit "Just Take My Heart" by Mr. Big and the 1963 hit "What Will My Mary Say" by Johnny Mathis.

Figure 2 plots the annual market volatility and the negative average beat variance of the Billboard Hot 100. The negative of the beat variance helps facilitate visual confirmation of comovement: it is easier to spot two graphs moving together than opposite each other. We can also see from figure 2, as well as from table 2, that the beat variance does seem to be a relatively consistent measure, thereby supporting the implicit assumption of this study that social mood is stable over a period of a year.

Figure 2: Market Volatility and Negative Beat Variance. This figure shows the annualized volatility (shown in gray using the right axis) computed from daily S&P 500 returns for each year corresponding to the negative average beat variance of the Billboard Hot 100 songs (shown in black using the left axis). The negative beat variance is plotted to make it easier to spot the comovement between the two series.

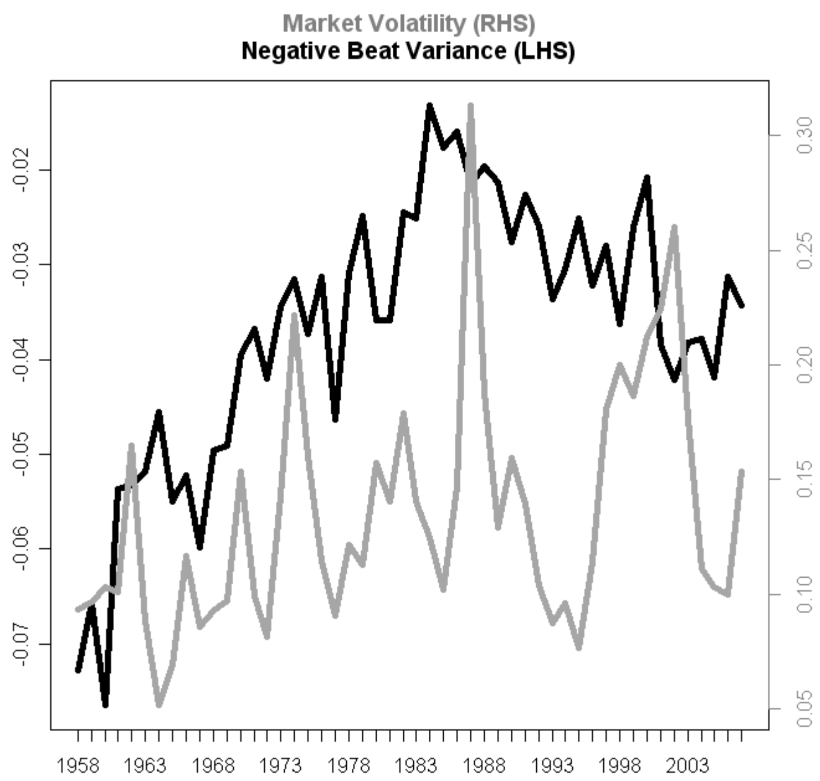
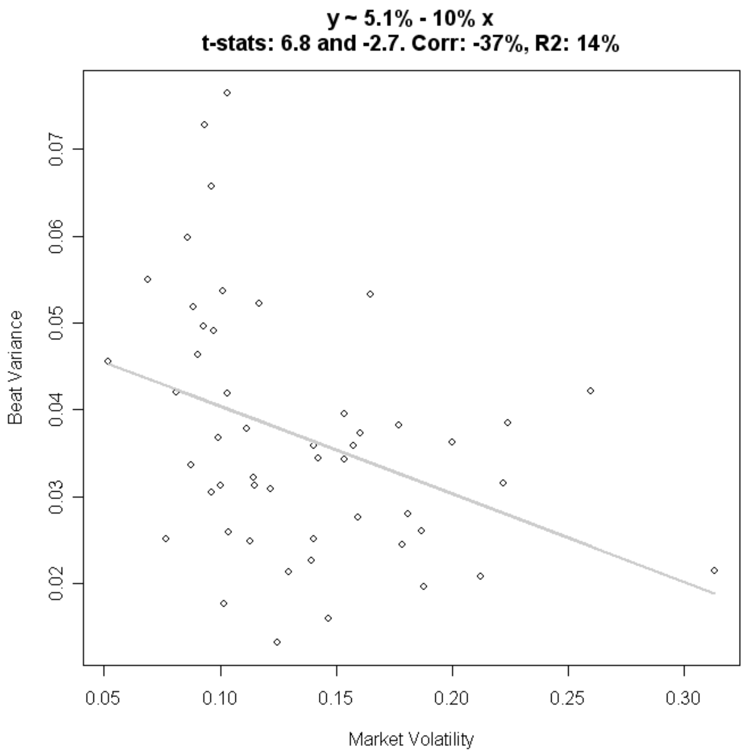


Figure 3 is a scatterplot of the average beat variance against the market volatility, with regression results.[2] The reported t-statistics are corrected using Newey and West (1994) as implemented and described by Zeileis (2004). The R^2 of the regression is 14% and the correlation between the average beat variance and the market volatility is negative 37%. The t-statistic on the coefficient of -2.7 is significant at the p = 1% level. The average beat variance decreases by about 10% of the increase in market volatility, with a "base" beat variance of 0.051 if the market volatility is effectively zero.

Figure 3: Market Volatility and Beat Variance. This figure shows the scatterplot of the annualized volatility computed from daily S&P 500 returns for each year corresponding to the average beat variance of the Billboard Hot 100 songs, as well as the best fit regression line with the regression results and Newey-West t-statistics. For a year where the market volatility averages 40%, the regression predicts an average beat variance of the top 100 songs will be 5.1% - 10% * 4.1% = 1.1%.



[2] The results are not driven solely by the 1987 crash. Regression results excluding 1987 are virtually identical.

3.1. Predictive Power

There is very little year-to-year overlap of the same songs. Out of the 50 years, 30 charts had no songs from the preceding year and another 10 had three or fewer songs from the preceding year. Only two years had eight songs repeated from the previous year; no years had more repeats than that.

In other words, the year-to-year averages of the beat variances are practically independent in terms of the underlying data. Nevertheless, they exhibit persistent autocorrelation. Table 3 lists the autocorrelation of the annual beat variance for a variety of lags. At a lag of one, i.e., the autocorrelation between each year's beat variance with that of the preceding year, the correlation is 0.84. It continues to remain high for lags up to eleven years. A Dickey-Fuller test for the presence of a unit root yields a statistic of -2.92, but this only corresponds to a p-value of 0.21, so we cannot reject the hypothesis that the time series of annual average beat variance has a unit root.

The market volatility is calculated using non-overlapping daily returns, so they are also practically independent. Yet they also exhibit autocorrelation. Table 3 also lists the autocorrelation of the market volatility for a variety of lags. At a lag of one, i.e., the autocorrelation of this year's volatility with last year's, the correlation is 0.55. But the correlation drops immediately after that and is never statistically significant again. A Dickey-Fuller test for the presence of a unit root yields a statistic of -4.45, corresponding to a p-value of less than one percent, so we can reject the hypothesis that the time series of market volatility has a unit root.

In other words, it seems as if people's musical preferences for beat variance persist for about a decade, making it reasonable to speak of the 80s as having a common type of music distinct from that of the 70s or the 90s yet having commonality between songs from 1980 and 1989. The regimes of market volatility, on the other hand, seem to last only a year.

The year-to-year consistency of the average beat variance may raise concerns of a spurious regression. However, the demeaned Phillips-Ouliaris (1990) statistical test for stationarity of the spread rejects the null hypothesis of non-cointegration at a p-value of less than one percent.

Table 3: Autocorrelations. This table lists the autocorrelations for the annual average beat variance and the market volatility for a variety of different lags, along with a Newey-West adjusted t-statistic (in brackets) of the corresponding regression to determine significance (bolded).

Lag	Market Volatility	Beat Variance
1	**0.5451 [5.17]**	**0.8353 [15.93]**
2	0.1562 [0.96]	**0.7967 [6.30]**
3	0.0695 [0.36]	**0.7565 [4.90]**
4	0.0881 [0.48]	**0.6883 [4.52]**
5	0.0128 [-0.08]	**0.6725 [4.45]**
6	0.1360 [-0.87]	**0.6158 [3.85]**
7	0.1275 [-0.62]	**0.5940 [3.43]**
8	0.0825 [-0.44]	**0.5807 [3.32]**
9	0.0720 [-0.80]	**0.5386 [2.99]**
10	0.0105 [0.06]	**0.4289 [2.34]**
11	0.1421 [0.71]	**0.3330 [2.18]**
12	0.3614 [1.97]	0.2547 [1.35]
13	0.4569 [2.90]	0.1620 [0.94]
14	0.3228 [1.66]	0.0598 [0.37]
15	0.1416 [0.54]	0.0109 [-0.06]

Table 4 lists the results of a Granger (1969) causality test. This test checks whether past values of one variable provide statistically significant additional information about another variable, over and above the information that is in the second variable's own past history. Using Pfaff (2008), a vector autoregressive model with a constant term was separately estimated for market volatility and beat variance to determine the optimal lag based on four possible information criteria: Akaike, Hannan-Quinn, Schwarz, and forecast prediction error (FPE). Market volatility had an optimal lag of one under each criterion while beat variance had an optimal lag of one under Hannan-Quinn and Schwarz and an optimal lag of three under Akaiki and FPE. Lags up to five years of history are reported to provide a comprehensive picture. The only evidence of Granger-causality in either direction is a weak one (p-value of 8.9%) for a one-year lag that suggests that last year's average beat variance has some small predictive power about the future market volatility.

Table 4: Granger Causality Tests. This table lists the F-statistics and the p-values (in parenthesis) of a Granger causality test on average beat variance and market volatility for a variety of lags. A low p-value means we can reject the null hypothesis that there is no causation. The first tests whether market volatility Granger-causes beat variances; in all cases, the null hypothesis is not rejected, and so we conclude that past market volatility does not Granger-cause future beat variance. The second column tests whether the beat variance Granger-causes market volatility; in all cases after the first two years, the null hypothesis is not rejected, and so we conclude that past beat variance does not Granger-cause future market volatility. The first two years, however, are weakly rejected.

Lag	Market Volatility → Beat Variance	Beat Variance → Market Volatility
1	0.15 (69.97%)	3.02 (8.90%)
2	0.01 (99.06%)	2.04 (14.24%)
3	0.65 (58.46%)	1.40 (25.68%)
4	0.83 (51.19%)	1.07 (38.47%)
5	0.95 (45.90%)	1.46 (22.85%)

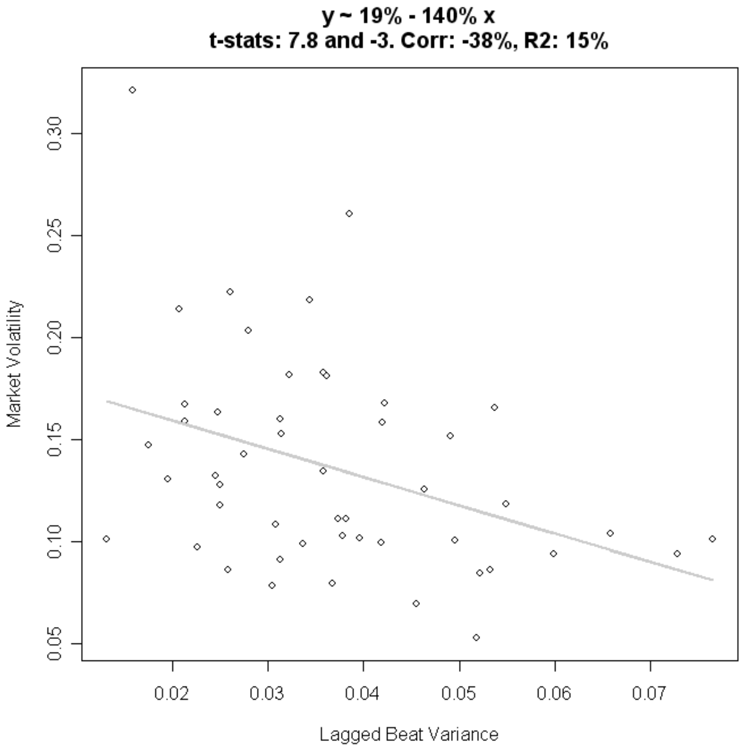
For the remainder of this section, we will consider whether this year's average beat variance on popular songs can tell us anything about next year's market volatility as we consider a simple trading strategy and examine its profitability.

First, we reverse the regression of figure 3 to use market volatility as the independent variable and beat variance as the dependent variable. Additionally, we now use calendar year market volatility instead of matching to the November-November cycle on which the Billboard Hot 100 chart is based. Using the calendar year market volatility ensures that it is calculated strictly after the December 31 release date of the year-end charts. The R^2 and correlation are of course unchanged but now the Newey-West adjusted t-statistic on the coefficient is -3.1, significant at the p = 0.42% level.

Figure 4 shows a scatterplot of market volatility on *lagged* annual beat variance. The R^2 is 15% and the correlation is -38%, up mildly from the non-lagged regression. The figure shows numbers only to a few significant digits; the precise regression result is:

$$\textit{Market Volatility next year} = 0.1873 - 1.3914 * \textit{average Beat Variance this year} \quad (1)$$

Figure 4: Market Volatility and Lagged Beat Variance. This figure shows the scatterplot of the annualized volatility computed from daily S&P 500 returns for each calendar year corresponding to the year after the chart release and calculation of the average beat variance of the Billboard Hot 100 songs, as well as the best fit regression line with the regression results and Newey-West t-statistics.



It seems as if the past characteristics of popular music can predict the future volatility of the equity market. Table 5 reports the regression results for lags ranging from one year (as shown in figure 4) to five years. The relation ceases being statistically significant starting with a five year lag. Therefore, it appears to be a stable relation at least for a period of several years.

Table 5: Lagged Prediction of Market Volatility. This table lists the correlation, regression coefficients, and Newey-West adjusted t-statistics for a regression of S&P 500 annual volatility on the average beat variance of the Billboard Hot 100 songs from prior years, ranging from one year prior to five years prior:

S&P 500 Annual Volatility at time *t* = Intercept + Coefficient * Billboard Hot 100 Beat Variance at time *t*–Lag

Lag	Intercept	Coefficient	Correlation
1	19% [7.8]	-139% [-3.0]	0.38
2	19% [7.2]	-140% [-2.6]	0.38
3	19% [6.9]	-145% [-2.7]	0.40
4	18% [6.8]	-113% [-2.3]	0.31
5	18% [5.5]	-116% [-1.8]	0.32

Table 6 summarizes the regression results, including two controlling for lagged market volatility. When controlling for the prior year's market volatility, the regression is done in two stages: first, the current market volatility is regressed on the lagged market volatility, and next, the unexplained residual from that first stage is regressed on either the current or the lagged beat variance. This approach allows us to examine the incremental impact of incorporating beat variance into our measurements. In all regression specifications, the coefficient on beat variance, whether lagged or not, is statistically significant.

Table 6: Regression Results. This table lists the regression coefficients and Newey-West adjusted t-statistics (in parenthesis) for a regression of S&P 500 annual volatility on the average beat variance of the Billboard Hot 100 songs from either the same or the prior year, and either controlling or not for the lagged market volatility. When controlling for lagged market volatility, the regression is done in two stages to assess the impact of beat variance on the portion of market volatility remaining unexplained after accounting for the previous value of the market volatility.

(1) Market Volatility at t = Intercept_1 + Coefficient * Beat Variance at time t + Residual
(2) Market Volatility at t = Intercept_1 + Coefficient * Beat Variance at time t–1 + Residual
(3) Market Volatility at t = Intercept_1 + Coefficient_1 * Market Volatility at time t–1 + Residual_1
Residual_1 = Intercept_2 + Coefficient_2 * Beat Variance at time t
(4) Market Volatility at t = Intercept_1 + Coefficient_1 * Market Volatility at time t–1 + Residual_1
Residual_1 = Intercept_2 + Coefficient_2 * Beat Variance at time t–1

Regression	(1)	(2)	(3)	(4)
Intercept 1	0.185 [8.41]	0.187 [7.82]	0.071 [4.27]	0.071 [4.27]
Intercept 2			0.026 [1.89]	0.027 [1.76]
Beat Variance	-1.36 [-3.15]		-0.73 [-2.52]	
Lagged Beat Variance		-1.39 [-3.01]		-0.73 [-2.32]
Lagged Market Volatility			0.49 [4.39]	0.49 [4.39]

An additional way of examining the impact of adding lagged market volatility as a control is in the context of the profitability of a trading strategy. Becker, Clements, and White (2006) demonstrated that the implied volatility index on the S&P 500 was not efficient during the period 1990 to 2003, and concluded that "an improved forecasting model could potentially be found." In that context, this strategy can be considered a candidate for such an improved forecasting model. Reitz (2006) argues that technical trading is only useful when the volatility is low and the rate is driven by hidden fundamentals. Although our strategy may be viewed as a type of technical strategy, it is not an oscillator model of moving averages, but rather a quantitative model that takes as input a seemingly separate source of cultural data.

The base strategy essentially predicts market volatility, and so would have required options or variance swaps to realize. Unfortunately, there is no such market data available going back to 1958. Instead, we can approximate by supposing that we can purchase future volatility at the current volatility level.

The strategy works as follows. On December 31, 1958, the first Billboard Hot 100 is released and we can calculate the average beat variance of the songs; it is 0.07284. From the lagged regression results of figure 4, we predict from equation 1 that the market volatility for the coming calendar year will be 0.1873 - 1.3914 * 0.07284 = 0.0860.[3] We compare that with the realized volatility of the previous calendar year, in this case, 0.0897, to conclude that next year's volatility is predicted to be lower. Therefore we sell one year calendar volatility at 0.0897, expecting a profit. In fact, the market volatility ends up at 0.0941, so this first year is a loss of 0.0941 - 0.0897 = 0.0044, or about half of a "volatility point." It turns out 1959 was one of the sixteen years with a negative profit.

---

[3] In principle we could also incorporate lagged market volatility for even better results, but because the purpose of this study is to explore the possible direct links between music and market volatility, we focus only on the strategy of equation (1).

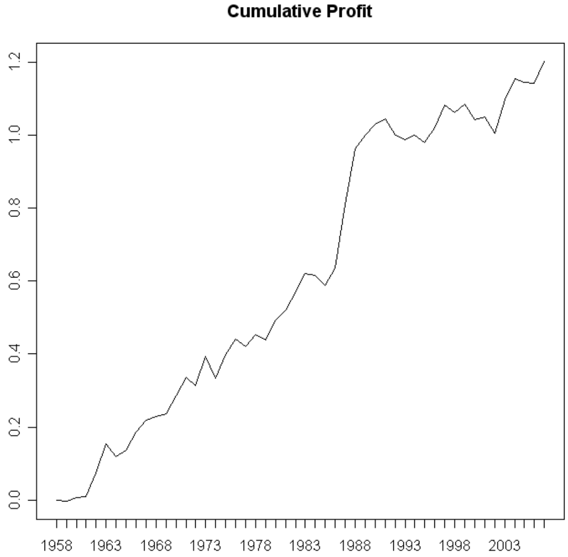
Table 7 summarizes the overall profitability of this strategy. The average profit is 2.5 volatility points per year. The standard deviation of profits is 4.7 volatility points but that is primarily profit volatility, such as the 17.5 volatility points of profit during 1987. The standard deviation of the sixteen realized losses is only 1.7 volatility points, resulting in a Sortino ratio (the equivalent of the Sharpe ratio but using only downside volatility) of 1.4.

Table 7: Summary of Profitability. This table summarizes the profitability of a strategy that buys next year's market volatility at the assumed price of this year's volatility whenever the volatility predicted by equation 1 exceeds the current volatility, and sells otherwise.

Worst Year	-0.0604 in 1974
Second Worst Year	-0.0465 in 2002
Median	0.0222
Mean	0.0246
Standard Deviation	0.0472
Upside Standard Deviation	0.0379
Downside Standard Deviation	0.0171
Second Best Year	0.1541 in 1988
Best Year	0.1746 in 1987

Removing the 1987 data point from the observation still results in a 2.1 volatility point profit on average. Figure 5 shows the accumulation of profit over time. Though 1987 and 1988 were the two best years, the strategy is still profitable even without those years, resulting in a 1.9 volatility point average annual profit.

Figure 5: Profit. This graph shows the cumulative profit of buying volatility for one year when the predicted level exceeds the current level, and selling otherwise. At the end of 2007, the total profit is 120.35 volatility points.



Though profit from the strategy is expressed in volatility points, it is essentially the same as a return, since it only needs to be divided by the number of volatility points required as collateral for either the hypothetical variance or volatility swap or the portfolio of options held to approximate a constant volatility exposure. The regression of the strategy returns on the market excess return, on the Fama-French three-factor model, and on the three-factor model plus momentum is shown in Table 8. None of the factor loadings are significant, but the alpha is both significant and approximately equal to the average annual profit.

Table 8: Factor Model Regression Results. This table shows the results of regressing the annual return of a strategy that buys next year's market volatility at the assumed price of this year's volatility whenever the volatility predicted by equation 1 exceeds the current volatility, and sells otherwise, on, respectively, the market excess return (CAPM), Mkt-Rf; on the three Fama-French factors (FF3F), *Mkt-Rf*, *SMB*, and *HML*; and on the three Carhart four-factor model, namely the Fama-French factors plus momentum (FF3F+MOM), *Mkt-Rf*, *SMB*, *HML*, and *MOM*.

	CAPM	FF3F	FF3F+MOM
Intercept	0.0216 [2.75]	0.0231 [3.11]	0.0446 [2.73]
Mkt-Rf	0.0005 [1.27]	0.0005 [1.33]	0.0003 [1.14]
SMB		-0.0003 [-0.79]	-0.0004 [-1.03]
HML		-0.0002 [-0.49]	-0.0008 [-1.69]
MOM			-0.0015 [-1.83]

This strategy implicitly used information from the future because the regression results of equation 1 were calculated using the entire sample period. A more realistic trading strategy would calculate the regression results each year using only past data. The drawback of such a strategy is the inability to trade for the first few years as the required number of years for the first regression must first pass.

How many years back should we look? Given the autocorrelations of table 3 discussed above, the commonplace notion of music being grouped into decades, and a trader's natural impulse to prefer relatively recent history, it is reasonable to use a lookback window of ten years.

The first time a trade is made is on December 31, 1968. A regression is run between the ten annual market volatilities for the years 1959 through and including 1968 on the ten annual (lagged) average beat variances for the years 1958 through 1967. The market volatility for next year is forecast using that regression result and the average beat variance for 1968. If the forecast volatility exceeds the realized volatility of 1968 (used as a proxy for the implied volatility for 1969), then we purchase volatility for one year, and our profit is the difference between the 1969 volatility and the 1968 volatility. Symmetrically, we sell if the forecast is below the realized.

Perhaps the profit from this strategy comes from mean reversion in the volatility itself. To distinguish that hypothesis, an alternative strategy is to regress on lagged market volatility instead of lagged beat variance. Finally, we can also regress on both the lagged market volatility and the lagged beat variance.

The results of these three strategies are presented in Table 9. Perhaps not surprisingly given the rolling nature of the regressions, the loss of the initial data to bootstrapping, and the number of years of data, none of the profits are statistically significant.

However, regressing market volatility only on beat variance does yield a profit of nearly three-quarters of a volatility point on average. Furthermore, using both beat variance and volatility as compared to just volatility alone increases the average profitability by a fifth. Finally, the rolling beat variance strategy was the only one that would have purchased volatility at the end of 2007. It would have made more than 25 volatility points of profit in 2008.

Table 9: Rolling Trading Strategy Results. This table shows the results of three different trading strategies. Each strategy regresses the past ten years of market volatility on either (1) the lagged beat variance, (2) the lagged market volatility, or (3) both the lagged beat variance and the lagged market volatility. The regression is then used in conjunction with the current beat variance and volatility, as needed, to forecast future market volatility. Buy volatility if the forecast exceeds the current level (used as a proxy for the one-year implied) and sell otherwise. Below, the mean and its standard error of each strategy are listed. Then the third row indicates whether the strategy would have bought or sold volatility in 2008. The final two rows list the mean and its standard error of each strategy through 2008, given the 41% realized market volatility for 2008.

	Market Volatility	Beat Variance	Both
Mean Profit	1.10%	0.72%	1.32%
Standard Error	0.87%	0.88%	0.86%
2008 Prediction	Sell Volatility	Buy Volatility	Sell Volatility
Mean Profit, incl. 2008	0.44%	1.33%	0.66%
Standard Error, incl. 2008	1.08%	1.06%	1.08%

4. Conclusion

When economic actors consider more complex and riskier economic activity in the future, they prefer to listen to simpler music. Indeed, there appears to be a negative relation between music volatility and market volatility. In tumultuous financial times, people prefer steadier music, and in stable financial times, people prefer tumultuous music. Furthermore, it appears as if musical tastes come first, as predicted; they have some ability to predict future market volatility.

The link between music and trading has not been studied in much depth, partially due to a difficulty in obtaining quantitative data. This paper shows not only that there is a link between song and stock volatility, but that the causality appears to go in an unexpected direction; namely, this year's popular music seems to predict next year's market volatility. The reason is that musical preferences are chosen by economic actors to counteract the complexity of their future planning.

Future research could include replicating these results for other markets when the requisite data for their charts and songs become available. In addition, when available, weekly popular music charts could be compared to weekly estimates of market volatility.

From the perspective of musicians deciding what types of songs to perform, the recent market volatility suggests that people will prefer steadier music, much like the 1980s. From the perspective of investors deciding what volatility to forecast in the future, the recent musical preferences suggest that people are considering complex economic activity in the future, which may lead to relatively higher market volatility in the future.

5. References

Becker, Ralf, Adam E. Clements, and Scott I. White (2006), "On the informational efficiency of S&P500 implied volatility," *The North American Journal of Economics and Finance*, **17**:2, 139-153.

Bruner, Gordon C. II (1990), "Music, Mood, and Marketing," *The Journal of Marketing* **54**:4, 94-104.

Camerer, Colin, George Loewenstein, and Drazen Prelec (2005), "Neuroeconomics: How neuroscience can inform economics," *Journal of Economic Literature* **43**:1, p-64.

Carhart, Mark M. (1997), "On persistence in mutual fund performance," *The Journal of Finance* **52**:1, 57-82.

Crain, W. M. and Tollison, R. D. (1997), "Economics and the architecture of popular music," *Journal of Economic Behavior and Organization* **901**, 185-205.

Dibben, Nicola and Victoria J. Williamson (2007), "An exploratory survey of in-vehicle music listening," *Psychology of Music* **33**, 571-589.

Fama, Eugene F. and French, Kenneth R. (1993). "Common Risk Factors in the Returns on Stocks and Bonds," *Journal of Financial Economics* **33**, 3–56.

Furnham, Adrian and Kathryn Allass (1999), "The influence of musical distraction of varying complexity on the cognitive performance of extroverts and introverts," *European Journal of Personality* **13**:1, 27-38.

Granger, C. W. (1969), “Investigating causal relations by econometric models and cross-spectral methods”, *Econometrica* **37**, 521-533.

Hirshleifer, David and Shumway, Tyler (2003), “Good day sunshine: Stock returns and the weather,” *The Journal of Finance* **58**:3, 1009-1032.

Keef, Stephen P. and Roush, Melvin (2005), “Influence of weather on New Zealand financial securities,” *Accounting and Finance* **45**:3, 415-437.

Konečni, Vladimir J. and Dianne Sargent-Pollock (1976), “Choice between melodies differing in complexity under divided-attention conditions,” *Journal of Economic Psychology: Human Perception and Performance* **2**:3, 347-356.

Lee, Kevin and Kalvinder K. Shields (2011), “Decision-making in hard times: What is a recession, why do we care and how do we know when we are in one?” *The North American Journal of Economics and Finance* **22**:1, 43-60.

Lepori, Gabriele M. (2010), “Positive Mood, Risk Attitudes, and Investment Decisions: Field Evidence from Comedy Movie Attendance in the U.S.,” *Working paper* available at http://ssrn.com/abstract=1690476.

Levitin, Daniel J. (2007). This is Your Brain on Music: The Science of a Human Obsession. USA: Plume/Penguin.

Merton, R. C. (1980), “On estimating the expected return of the market”, Journal of Financial Economics **8**, 323-361.

Newey, W. K. and West, K. D. (1994), “Automatic lag selection in covariance matrix estimation,” *Review of Economic Studies* **61**, 631-653.

North, Adrian C. and David J. Hargreaves (1999), “Music and driving game performance,” *Scandinavian Journal of Psychology* **40**, 285-292.

North, Adrian C. and David J. Hargreaves (2007), “Lifestyle correlates of musical preferences: 3. travel, money, education, employment and health,” *Psychology of Music* **35**, 473-497.

Pearce, Marcus T., Maria Herrojo Ruiz, Selina Kapasi, Geraint A. Wiggins, and Joydeep Bhattacharya (2010), “Unsupervised statistical learning underpins computational, behavioural, and neural manifestations of musical expectation,” *NeuroImage*, **50:**1, 302-313.

Pettijohn, Terry F., II, and Donald F. Sacco, Jr. (2009), “Tough times, meaningful music, mature performers: popular Billboard songs and performer preferences across social and economic conditions in the USA,” *Psychology of Music* **37**:2, 155-179.

Pfaff, Bernhard (2008). “VAR, SVAR and SVEC Models: Implementation Within R Package vars," *Journal of Statistical Software*, **27**:4. URL: http://www.jstatsoft.org/v27/i04/

Phillips, P.C.B. and S. Ouliaris. (1990), “Asymptotic Properties of Residual Based Tests for Cointegration,” *Econometrica,* **58**:1, 165–193.

Reitz, Stefan (2006), “On the predictive content of technical analysis,” *The North American Journal of Economics and Finance* **17**:2, 121-137.

Sacks, Oliver (2008). *Musicophilia*. USA: Vintage.

Simon, Herbert A. (1955), “A Behavioral Model of Rational Choice,” *The Quarterly Journal of Economics* **69**, 99-118.

Whitburn, Joel (2009). *Top Pop Singles*, 12th edition. USA: Record Research, Inc.

Zeileis, A. (2004), “Econometric computing with HC and HAC covariance matrix estimators”, *Journal of Statistical Software* **11**, 1-17.

Zullow, Harold M. (1991), “Pessimistic rumination in popular songs and newsmagazines predict economic recession via decreased consumer optimism and spending”, *Journal of Economic Psychology* **12**:3, 501-526.