Anthropology Senior Theses   Department of Anthropology

Spring 2008

Stock Spam Emails: Proliferation and Impact

Ernest Baskin
University of Pennsylvania

Follow this and additional works at: http://repository.upenn.edu/anthro_seniortheses

Part of the Anthropology Commons

Stock Spam Emails: Proliferation and Impact

Disciplines
Anthropology

STOCK SPAM EMAILS: PROLIFERATION AND IMPACT

By

Ernest Baskin

In

Anthropology

Submitted to the

Department of Anthropology

University of Pennsylvania

Thesis Advisor: Dr. Greg Urban

2008

Stock spam is an ever increasing problem in the modern age of fast processors and large bandwidth. This paper gives a general overview of stock spam and then attempts to determine the changing distribution of stock spam over the next several years using the Negative Binomial Distribution and the Evolving Visits models and positioning the analysis in terms of cultural motion. Next, the paper looks at the case study of Emulex to draw conclusions of how false news affects the market. Finally, the paper looks at the new trends in the world of stock spam and tries to make predictions based on them.

Introduction

On February 25, 2001, the *New York Times Magazine* ran a feature story on Jonathan Lebed, a seemingly normal 15 year old(Lewis, 2001). However, this teenage computer whiz, had an important difference that set him apart from most students his age. As mentioned in the *New York Times Magazine* article, he had a case brought against him by the Securities and Exchange Commission (S.E.C.) and had settled for around 285,000 dollars. Lebed caught the eyes of the S.E.C. after he started taking large positions in penny stocks and then touting them on his website, Stock-dogs.com, and on various message boards such as Yahoo Finance. This publicity convinced naïve investors to believe that these stocks were poised to skyrocket. Unknowing investors brought in to the hype and bought shares of these stocks at which point Lebed sold his own shares. The money that that he was forced to give up to the SEC was far from all of the money that he earned. Over the course of the period in question, he had made over 800,000 dollars by buying low and selling high.

Lebed's story is a perfect lens to think about the impact of certain events on stock prices. How was a 15 year old living at home able to influence investors and then profit from their investments? More generally, the same questions can apply to company reports, earnings statements, and analyst reports. How do they influence the stock market as events that are not wholly representative of the true status of the entire company?

My proposal is to interpret the events and opinions influencing stock prices through the lens of metaculture. Metaculture, in its simplest form, "is culture about a culture"(Urban, 2001). In a way, it is the mechanism responsible for the movement of culture from person A to person B. In the case of earnings restatements and the stock market, this paper acts like a metacultural object in transferring cultural knowledge to you, its reader. In the same way, statements about a

company will affect the company's stock metaculturally. Metaculture has the power to accelerate certain bits of culture while slowing down the movements of other parts.

Inherently, stock prices measure the worth of a company using the intersection of supply and demand. In general, price is determined by the intersection of the prices that people are willing to buy and sell a particular stock. The finance literature has researched numerous corollaries to this phenomenon. Most important for us is the literature on market efficiency (see (Fama, 1991) for a review). The general consensus is that prices reflect all available information on the stock barring efficiency and trading costs. In addition, privately held information may not be perfectly reflected in stock prices. In this way, the publicly available information changes the supply and demand functions of the players in the market. The intersection moves to a new point and the price changes. Arbitrage, to the extent that it is economically feasible when taking trading costs into account, assures that stock prices reflect market information. Thus, since all knowable information is incorporated into stock prices, there is no way to efficiently beat the market in the long run(Malkiel, 1996). However, insider information is not necessarily priced in to the market and thus, acting on that information will ensure a profit for the active trader in the short run. Metaculture is the mechanism for this insider information to be spread to the market. The reason that people like Jonathan Lebed are able to so profoundly influence certain stocks is that their information acts much like insider information would and metaculturally primes the stock price. This occurs even though the information may not even be real to begin with.

This brings up an interesting problem in regards to stock prices. Market information is not necessarily equivalent to firm information. While a firm works on its annual 10-K, analysts offer earnings predictions. When merger talks are rumored, analysts may contribute to the rumor mill. Legally, analysts are not firm insiders so they cannot know the real news even though they

may work with the company and check their information. However, analysts maintain a large cache of cultural capital and thus, their predictions are trusted as a proxy to actual insider news. These predictions are aggregated and incorporated into stock prices which are then evaluated against official news released by the company. The degree of error has an effect on the degree to which the stock price responds.

From a cultural point of way, analysts occupy a privileged position in the financial world. This privileged position means that their predictions have a greater accelerative effect. Wall Street analyst predictions are circulated quickly among institutional investors who then make their decisions off of them. Since these are the pieces of metaculture that underlie the majority of stock trades, it should come as no surprise that the stock prices use analyst predictions predominantly as a basis for earnings expectations. Predictions made by John Q. Public are less important since institutional investors ascribe less importance to them. Still, this is changing. More and more investments are made by private individuals as opposed to institutional investors. With the rise of discount brokers such as E*trade and Ameritrade, buying stocks has never been easier. Private individuals have less access to Wall Street analyst predictions and are thus more likely to accept advice on message boards or private websites. Since this advice does not necessarily parallel analyst predictions, the stock market could possibly see future distortions as the volume of privately traded stocks increases.

The phenomenon can already be measured when looking at so-called "whisper numbers." These are unofficial earnings predictions that do not come from Wall Street analysts. Whisper numbers can include everything from a random internet website that makes a stock prediction to discussions that traders have amongst themselves before a company's official earnings release. Whisper numbers have become so powerful that some stock prices react based upon the

3

consensus whisper number as opposed to the consensus Wall Street analyst prediction. Thus, if an analyst predicts earnings at .20 cents a share, the whisper number is .23 cents a share and the actual earnings number is .22 cents a share, the stock may go down on the earnings release. The stock movement, in this case, was based on the .23 cent expectation from the whisper number. In a Purdue university study, whisper numbers were found to be more accurate predictors of earnings than official analyst numbers(Bagnoli, Beneish, & Watts, 1999). In fact, since whisper numbers affect prices more than analyst numbers, research suggests that investors could profit on the difference between these two earnings predictors.

It appears that different pieces of metaculture compete to be recognized by the investor. Which cultural pathway does the investor consider to be the true prediction? Which piece of metaculture should be replicated? Theoretically, culture continues along its trajectory, absent an accelerative force. Historically, investors have always trusted analyst reports. The widespread use of whisper numbers is only a recent phenomenon associated with the internet's democratization of information. Already, whisper numbers are beginning to coalesce into a single cultural object. Much like analysts have FirstCall which aggregates analyst earnings predictions, whisper numbers are now aggregated on whispernumbers.com. Thus, whisper numbers are obtaining more cultural cache. Their availability in a central place means that more traders can now use them to make trading decisions. The cultural trajectory seems to be moving towards whisper numbers and away from analysts.

Whether analyst predictions or whisper numbers wind up being more useful, the winner will wind up affecting company stock prices. The pieces that win then performatively influence the stock price. This phenomenon is similar to how the Black-Scholes model of options pricing first became an accurate predictor of stock prices. After the model began to be used by traders,

the prices calculated by the model began to appear much more accurate(MacKenzie & Millo, 2003). Much like Black-Scholes model now has competition from the Cox-Ross-Rubinstein model and is now less accurate; analysts have competition from a variety of sources including private investor newsletters. With the advent of the internet, metacultural information is more publicly available and can be interpreted and distributed by a variety of non institutional users.

In Jonathan Lebed's case, metacultural stock information did not exist for the stocks he was trading in since they were primarily penny over the counter stocks. These stocks were typically illiquid and had very low trading volume. Thus, it was fairly easy to manipulate their price. Since the cultural pathway of reliable information was not established for these stocks, obtaining the cultural capital to affect their motion was simpler. Had Jonathan tried to metaculturally pump up the value of an IBM or an Amazon.com, he would have failed miserably since he did not have the cultural capital to be trusted in this area.

More generally, Jonathan was engaging in a scheme called pump and dump. This is a form of metacultural pumping. The scheme works in the following manner. A person, called the promoter, buys a relatively illiquid stock when it is at a very low price or gets the stock for free as a payment for a service provided to the company. Then, the promoter promotes the stock through a variety of channels including email, websites and message boards. Since the stock is illiquid, it only takes a few willing investors to make the stock rise in price. Once the price has risen a sufficient amount, the promoter then sells their own shares and makes a significant profit. Without anyone willing to buy the stock anymore, the price declines significantly. According to S.E.C. guidelines, all analysts, whether employed by wall street or not, must disclose whether they are receiving compensation from the company they are reporting on or hold a position in the stock that they are analyzing:

Section 17(b) of the Securities Act [15 U.S.C. 77q(b)] makes it unlawful to publish any communication describing a security (whether or not the publication offers the security) where the publishing party receives payment from the issuer, an underwriter or dealer, without fully disclosing that such payment has been (or will be) received and the amount paid. ("Registration of Securities on Form S-8,")

It is also illegal to be falsely manipulate the stock market:

Section 10(b) of the Securities Exchange Act of 1934, 15 U.S.C. § 78j(b), and Rule 10b-5 thereunder, 17 cfr § 240.10b-5, are the primary authorities used by the SEC to combat market manipulation and other frauds in the securities market. These provisions make it unlawful to use a fraudulent scheme or to make material misrepresentation and omissions in connection with the purchase or sale of any security. ("Cybertelecom: Federal Internet Law & Policy: An Education Project,")

The stocks used typically have a low degree of liquidity in order to allow easy manipulability. In addition, they typically fall under the label of penny stocks: trading on exchanges that do not require a rigorous registration process. Although the name suggests that penny stocks trade for under a penny, some of them trade for more than a few pennies. Colloquially, penny stocks refer to stocks that trade on the Pink Sheets or the OTCBB.

These are the SEC rules that Jonathan unknowingly violated while trading stocks. Presenting himself as an uninterested party doing a service for consumers, he was, in fact, a very interested as a price increase would bring him more money.

Turning for a moment to the S.E.C., it is interesting that the organization was established to protect against certain kinds of metacultural transmission. Established to protect the average investor's faith in the stock market, the mission of the S.E.C is "to protect investors, maintain fair, orderly, and efficient markets, and facilitate capital formation"("The Investor's Advocate: How the SEC Protects Investors, Maintains Market Integrity, and Facilitates Capital Formation," 2006). To fulfill this mission, the S.E.C. attempts to protect the market from artificial manipulations. Unfortunately, the definition is somewhat circular as these artificial manipulations are also defined metaculturally by the S.E.C. Artificial manipulations are defined as something that is not an "ordinary market force"(Lewis, 2001) which is once again defined as something not artificial. The bits of metaculture in this case necessitate a direction and a way to discern what is legal versus what is not legal. For instance, aren't all analysts and company executives manipulating the market in some manner when they make predictions? The distinction as to who is actually manipulating the market falls upon the S.E.C. Perhaps the reason that Jonathan Lebed was not forced to pay back all of his earnings was that a brightline rule does not really exist. Still, as an agency with a lot of cultural capital, the S.E.C. has the power to make the cultural distinction between what is right and wrong in the financial world. Its decisions, rules, and acts serve as examples. In a world rocked by the scandals of Enron and Worldcom, the S.E.C. has gained even more capital with which to uphold stability. Investors are growing increasingly worried about their investments and retirement savings and the S.E.C.'s power draws from that.

7

Perhaps, Jonathan Lebed didn't really do anything wrong. All he did was to make predictions based on the information that he obtained from legitimate, and others from not so legitimate sources. This might be one of the reasons that the S.E.C. did not ask for the entire 800,000 dollars of profit. From a social perspective, the S.E.C. needs to show that it is doing its job in regulating financial markets. The market culture that the S.E.C. is disseminating comes from the cases that it regulates. With Jonathan Lebed, the S.E.C. has shown that it can take a stand on a case. To the public, the case seems like a victory for the S.E.C since Lebed was forced to pay some money back through a fine. The S.E.C. did not want to push the case to court since it feared losing a protracted court battle and thus decreasing its cultural capital. With a lowered cultural capital, it would have a much harder time acting as a deterrent against financial malfeasance.

Another case study is the metacultural effect of news/earnings statements on stock prices. Earnings announcements as well forward looking statements issued by company executives are metacultural objects that seek to change a supposedly accurate measure of the firm's value, the firm's stock price. These announcements/statements come from the company itself and so reflexively affect the corporation. However, both earnings announcements and forward looking statements may have ulterior motives behind their use. The company may attempt to artificially increase its value in order to achieve lower interest rates or to increase executive compensation packages. There is a check on stock price manipulation in this area as well. Shareholder lawsuits can be brought against the company. The majority of these lawsuits are brought against companies that enjoy abnormally high growth relative to similar companies or issue multiple positive forward looking statements before a low earnings release(Lev, 1995).

Metacultural signals of corporate malfeasance do exist outside of analyzing statements made by corporate executives. The matter of detecting them is, however, quite difficult. Quantitative measures, such as accrual, measure company earnings recognition relative to other companies. Companies with high accruals are more likely to have their stock value go down as well as encounter an earnings restatement or a lawsuit in the extreme case(Richardson, Sloan, Soliman, & Tuna, 2005). In addition, lawsuits have different effects on stock prices depending on the reasons that the lawsuits were brought against the company in the first place. Lawsuits filed due to "accounting irregularities, fraud, making overly optimistic statements, and failure to disclose negative news result in the most significant filing-date stock decline"(Loh & Rathinasamy, 2003). Lawsuits act as a metacultural defense mechanism in an attempt to correct past wrongs.

From these studies, it appears that stock prices usually revert to "normalcy." While stocks may be undervalued or overvalued due to certain metacultural signals for periods of time, their value will be eventually be corrected. Thus, metaculture does not always provide the correct interpretation of a phenomenon but rather one that can easily be believed. For instance, it is easier to believe that every forward looking statement made by a company is true rather than assume that every statement is false. People are inclined to believe metaculture that comes directly from the source of the original culture. Analyst predictions and whisper numbers, discussed earlier, follow a similar pattern. They are associated with industry insiders who probably know the truth. This is the same reason one does not follow the stock tips of a beggar on the street. His metacultural information is not connected to the culture that it originally sprung from. On the other hand, if that beggar begged for money outside of the company headquarters, and overheard the CEO and the CFO discussing company news, his metacultural

9

pumping would be much more credible. Unfortunately, in the realm of the internet, it is quite hard to tell whether the metacultural signals are actually based on real evidence or just part of a made up investment scheme. That is why some false metacultural information, can still move stock prices.

Thus, the purpose of this paper is several-fold. First, it will look at the distribution of stock spam emails in an attempt to gain insight into the metacultural distribution of stock spam emails. By looking at the distributions of the emails, one can make predictions about future distributions of stock spam and how information about a particular stock is spread through the metacultural mechanism of email. In addition, it is also necessary to take an in depth look into one of the stocks that has been metaculturally pumped, Emulex. By looking deeply into this stock, some commonalities can be deduced amongst all pumped stocks as a whole.

Unfortunately, I was unable to contact any of the spammers in order to do ethnographic interviews regarding the distribution of the stock spam. This is due to the questionable legality of stock spam emails. This brings up an important issue regarding ethnographic research that is based on illegal activities. It is often difficult to gain access to the confidential information needed to fully analyze the culture. This is why most of this analysis comes from publically available databases and personally collected information.

Part 1: Analysis of Stock Spam Distributions

Data on the exact distribution of stock spam emails is very poor. Beyond the general consensus that most stocks get spammed in waves, or cycles, there is relatively little information available on the forecasting of stock spam. In other words, a stock gets spammed for several days followed by a period of no spamming. During a spam wave, there is typically an increase in emails sent about a particular stock followed by a decrease. After that period, the stock may or may not be spammed again in a similar manner. Therefore, it would be very beneficial to understand the distribution that the stock spam waves follow to see if there are patterns in the data that can be exploited to help stop them. The goal of the analysis is to try and forecast the future trend of stock spam emails and to develop a model that will allow a more nuanced analysis into the spread of metaculture through stock spam emails.

The dataset used for this project comes from Spamnation.info and contains both aggregated and disaggregated data on the stock spam received by the owner of this domain from 2003 to the present day. The data set contains the number of emails received each day as well as the individual stock being pumped. The data is segregated by day. Spamnation.info consists of several different website addresses and may be regarded as a random sample of stock spam received even though it comes from only one domain name. In addition, the webmaster only records data from stock emails that he finds on his own servers and does not accept solicitations from visitors wanting to add to his database with the stock spam that they themselves receive. In this analysis, the spam is the cultural method of trying to get information about relatively unknown stocks into the hands of unsuspecting investors. Once the investors act on this information, the spammer could potentially profit.

The precise method of recording the stock that a spam email is trying to pump consists of using simple mail filters that count up spam for particular symbols. For image based spam emails, they are separated out by the filters and manually added to the database so as not to miss any potential spam emails. While these image filters are not perfect, they work well enough for our purposes. Finally, only data from 2005 onwards is used in this study in an attempt to remove any potential bias that the spam collection may have as the website grows in popularity during its initial start period.

The sample size in this dataset is relatively large. In 2005, there were 331 different stocks being spammed. In 2006, there were 336 different stocks being spammed. In the first quarter of 2007, there were 163 different stocks being spammed. However, some stocks overlap periods so the total stocks spammed over the course of 2005-2006 was 593. The total number of stocks spammed between 2005 and the first quarter of 2007 is 697. The distribution of the dataset can be seen in Table 1.

For purposes of analysis, the start date of a spam wave is defined as a day when at least one email about a stock was received coming after a day when no email about that stock was received. For the NBD model, the database was reworked to obtain the number of waves coming from a particular stock in a particular period. This is allows us to look at the attractive forces of specific stocks or decide whether there are attractive forces leading towards a certain amount of spam waves for stocks in general. For the Evolving Visits model discussed later in the paper, the database was used to develop a list of pumped stocks and the days that each stock started a new stock spam wave as well as the number of waves coming from a particular stock. The evolving visits model as well as the usefulness of it in attempting to deal with non stationary data is discussed later in the paper.

As mentioned earlier, most of the stocks spammed are penny stocks. The market cap of these stocks is usually in the hundreds of thousands rather than the millions of dollars. They are usually not traded very much when they are not in the middle of a spam wave. In fact, it is a rare day when there are more than 300,000 contracts of the penny stock traded. This pattern changes when the stock is in the middle of a stock spam wave. At this point, several million shares can trade in one day. One example, VShield Software Corporation (VSHE.PK) was spammed in the middle of December of 2007. In one day, about 1.75 million contracts exchanged hands. Under 100,000 contracts exchanged hands the day before. With a price of about 15 cents, the first day of the spam wave saw about $262,500 dollars in trading. Unlike market makers, the spammers usually make money by taking a position in the stock before starting the spam. In the case of VShield Software Corporation, the price dropped by about 8 cents over the course of the spam wave. There were about 3 million total contracts traded over the course of the spam wave. Although the spammer was likely not involved in every contract trade, a conservative estimate for the spammer's income is about $240,000 from this one spam wave alone.

NBD

The first model fit to the database was the Negative Binominal Distribution (NBD). The data was transformed into the form of a counting dataset for use in the model. A counting dataset measures the amount of times with which specific frequencies occur. In this case, the NBD model was measuring the number of stocks that were spammed a specific number of times during a period. Each stock spam wave is a metacultural pump sent from the spammer. Unfortunately, there was no data available on how many stocks did not get spammed but had the potential to get spammed so a full NBD model could not be run. A truncated NBD model was run instead. The truncated NBD model attempts to derive the total number of stocks that were spammed zero times but had the potential to get spammed. This information is important because it allows us to predict how many stocks will actually get spammed in the future. A proxy for the zeros could be to take all of the stocks that were spammed in all other periods except for the period under investigation but there is limited data available. Only a few other periods were available in the future if the NBD model was run on the first 2 years of the dataset and taking these points would not be indicative of real world data since it would involve looking into the future to see what stocks wound being spammed there.

The assumptions underlying this model are that there are no contagion effects, no non-stationarity and no covariate effects. The first assumption states that the frequency with which an individual stock is being spammed does not influence the frequency with which other stocks are being spammed. As individual stocks are used as bases of analysis, this assumption is realistic. This assumption is realistic since an individual stock is used as the basis of analysis. Potentially, each stock has a different spammer since engaging in a pump and dump scheme requires that the spammer already have a stake in the company in order to be able to sell it once

14

the price gets pushed up to a reasonable level. Unfortunately, there have been recorded cases where a single spammer was found to be spamming multiple stocks. Even in those cases, though, there is no evidence to suggest that spamming one stock affects the chance for him/her to spam another stock since spamming still requires the obtainment of the initial position without pushing the price of the stock up too high. A stock won't be spammed unless the spammer can take this initial position. The probability of being able to take this initial position in each stock is independent of the ability to take such a position in any other stock. Spammers also have the ability to send spam emails about multiple stocks at the same time by compiling directories of email addresses through the use of bots and spiders sent to out to scour the internet looking for publically posted email addresses. They could also pay people who have developed lists of verifiable email addresses to use in their spamming.

The second assumption is much harder to dismiss and the extent of its impact must be analyzed separately. There is very likely a change in the means of the frequency distribution for stock spam emails over time. This may be due to several reasons. One reason may be due to the increased popularity of the site over time. In order to try to account for this, the first two years of data are being discarded in order to allow the site to establish its true frequency for stock spam waves. The next reason is technological. Both the physical number as well as the speed of computers has increased in the world. In addition, the prices of usable computers have been going down in recent years especially for bare bones functionality such as the sending of email. Viruses have also increased in proliferation and are capable of infecting other computers that can send spam emails as well. All of these factors combine to give some worry that the distribution of stock spam emails is actually non stationary and that the mean may be increasing.

As for the third assumption, covariate effects may exist. However, the data is not robust enough to allow for the analysis of covariates since there is no covariate data in the dataset. Based on previous modeling, covariates generally do not have a significant impact on probabilistic models anyway since most variation in the data can usually be captured by the parameters of the model. There is also no reason to believe that there is something inherent about certain stocks that will make them more or less likely to be spammed. While connections between spammers and certain companies may help in achieving the initial stock position, there is no reason to believe that any one company will have more connections to spammers than any others. Therefore, these behaviors can be modeled as if random.

The NBD model was estimated using the method of maximum likelihood estimation. This involved maximizing the combined Ln likelihood of all of the stocks being measured in the model. The NBD model assumes that each stock has a Poisson distributed probability of having different amounts of spam wave activity. In the metacultural model, the poisson probabilities can be likened to the attractiveness of each individual stock for spamming. Since this attractiveness is expressed in probabilistic terms rather than absolute terms, it may be the case that a very attractive and culturally rich stock may still not get spammed even though it has a high probability of getting spammed. These Poisson probabilities are heterogeneous across different stocks. This heterogeneity can be modeled using the gamma distribution (Gerber, 1992).

To test the robustness of the model, all of the data was used in the first derivation of the parameters. The data used came from 2005 through the first quarter of 2007. The full dataset was necessary because there are several different variations of the NBD model. The entire

dataset allows one to establish the rigor of using an NBD model on this dataset as well as to ascertain which variation of the NBD model works best for this particular dataset.

Three different truncated NBD models were fit to the data. The first was a truncated NBD which works in the manner described above. The second was an NBD with a spike at one. There is reasonable evidence to assume that there may be stocks that will only be spammed once no matter the time frame in question. This is because spammers may not think that people may be gullible enough to take a risk on a spammed stock if they have seen the stock email before. Spammers may also lose their stake in the company after performing one pump and dump scheme. Because of these reasons, some stocks should theoretically only be spammed once and the spike captures that by allowing the probability at one to be different from what it would normally be under the NBD model. The third model is the NBD model with latent classes. In this case, two latent classes were used. Latent classes imply that there are two different classes of stocks, each with a different distribution estimated by different r and alpha parameters. This is reasonable since most of the spammed stocks come from two categories, Pk stocks and Ob stocks. Pk stocks are those traded on the Pink Sheets while ob stocks are those traded on the Over the Counter Bulletin Board. While the model does not necessarily discriminate between the two types of spammed stocks when forming its latent classes, a significant probability of there being two or more latent classes would merit a closer look at the data.

The results of running the 3 models can be seen in Table 2. R and alpha are the parameters that characterize the NBD model while the spike and the latent class parameters are used to tweak the model so that it fits the data more closely. Based on Table 2, the NBD is fitting relatively well to the data. In all 3 types of NBD models, there is a significant probability

17

that the data actually came from the underlying distribution determined by the two parameters based on the chi test statistic.

The worst performing model is the NBD with a spike at one. Upon estimation, this model assigns a very negligible probability to the spike term meaning that the model does not need it. Therefore, the number of stocks that will only be spammed one time can be incorporated into the NBD model using the established parameters. No new spike parameters are needed. This can be seen in that the spike at 1 only creates a very negligible increase in the Ln likelihood. Therefore, the initial assumption that there are certain stocks that will only be spammed once that cannot be captured by the model is false. The latent class NBD model performs second best. The latent class model assumes that there are two different distributions that have their own R and Alpha. This variation of the NBD improves on the Ln Likelihood of the regular truncated NBD model but adds three extra parameters to it. This is why the BIC (Bayesian Information Criterion) for the latent class NBD model is higher than the BIC for the regular NBD. While the Ln Likelihood is improved, the problem with this model is that the Ln Likelihood is not improved enough given the three additional parameters. In estimating a model for stock spam emails, it is very important not to overfit the in-sample data so as to make sure that the model works well in the forecast period. Overfitting is caused when too many parameters are used when fitting the model to the sample data. Using a lot of parameters may cause random noise in the data to become significant and cause bad forecasting in the out of sample period.

One can also see that the fit has increased by looking at the chi squared goodness of fit p value. The latent class NBD has a much higher p value and there is a much greater chance that the data were actually generated from a distribution based on estimated parameters. Still, the extremely good fit for the current period's data prevents accurate forecasting for future periods.

This is why the BIC statistic is helpful in determining when a model is actually a better model versus when overfitting is occurring. The BIC statistic shows that there is no difference between the metacultural pumping that occurs for Pk and Ob stocks.

Based on this assessment, the regular truncated NBD seems to be the best model to use in this situation. Some interesting information about the distribution of stock spam emails can be obtained by analyzing the model. Both the R and the alpha parameters are relatively low indicating that most of the Poisson lambdas selected from the gamma mixing distribution are very low. The mixing distribution for the NBD model can be seen in Figure 1.

This suggests that most stocks have a very low lambda and have a very low probability of being spammed again. This corroborates the earlier hypothesis that spammers may not want to spam stocks that have already been spammed recently because they want to have as high a conversion rate to buying the spammed stock as possible. The low lambdas may also be affected by the fact that spammed stocks frequently reinvent themselves using new names. The dataset does not capture name changes and treats each name change as an entirely different stock. Thus, repeat spam waves for a stock with a name change would not be captured correctly. This is fine for the inferences we are trying to obtain since viewers of the spam emails will likely not know that a name change has occurred because it is not recorded in Yahoo Finance or other similar sources. A name change is like resetting the metacultural attractiveness of a stock. The only way to find a name change is to try and look up a record of the old name and see if that name still exists or if it has been changed to something else.

As shown in Figure 2, the model fits the data well. While there is a bit of noise in the data that prevents a perfect fit from being obtained, the NBD model captures the data nicely.

Thus, the inference can be made that stock spam waves arrive based on Poisson probabilities. Each individual stock has its own lambda that generates the Poisson probabilities for that stock.

The chi squared goodness of fit test is another statistic that can be used to determine the fit of the model, both in-sample and out-of-sample. The p value given in Figure 1 shows that it is highly likely that the observed data came from a negative binomial distribution with parameters of R = .121 and alpha =.078.

The mean frequency of stock spam waves per stock in this period, R/alpha, is about 1.55 cycles. Since the zeros were estimated rather than obtained from the data, it is important to note the mean frequency of spam waves for all stocks that had at least one spam wave. The mean frequency for those stocks is about 5.71 cycles. The mean frequency tells us the mean amount of cycles over the 2.25 years of our dataset.

By using a truncated NBD model in this instance, the model imputed the number of stocks that had a frequency of 0. These are the stocks that have not yet been spammed but have the potential to be spammed in the future. In this model, 1870 stocks fall into this scenario based on Figure 2. This number makes sense since the OTCBB and the Pinksheets have over 3000 and 4500 stocks listed, respectively. Based on this model, over a quarter of them have the potential to be spammed. While this is a purely probabilistic prediction made by extrapolating the model back one period, it is based on the data already collected and seems appropriate given the total potential amount of penny stocks and previous history of stock spam.

In order to determine the forecasting accuracy of the NBD model, the dataset was broken up into 2 equal sized groupings. 2005 was used as the in-sample period to fit the NBD model while 2006 was used as the out of sample holdout period in order to test forecast validity.

Figure 3 shows the fit of the NBD model for the 2005 data. Visually, the fit is not as great as the fit for the full dataset set. The chi-squared goodness of fit test confirms this. While there is still a reasonable chance that the data came from the same distribution as the parameters, the strength of this certainty has significantly decreased.

It must also be noted that the parameters of the model change when the model is fit to only the 2005 data. The R and alpha parameters are now, .44 and .19, respectively. This indicates that the data may not be entirely stationary and that a more advanced model may eventually need to be used.

The 2005 model also predicted that there would only be 208 stocks in the zero category. This would indicate that only 208 additional stocks had a probability of becoming spammed stocks in the next period. This estimate is incorrect based on the dataset since the dataset has 262 new stocks being spammed in the next period, 2006, alone.

This can be contrasted by looking at figure 4, which shows the NBD model fit to the 2006 stock spam waves. The fit is much better visually and the chi-squared statistic value has increased over the 2005 model. The r and alpha parameters here are .11 and .07, which are close to the full model fit over 2005 through the first quarter of 2007. The number of stocks in the zero category, 1012, also looks like a more appropriate figure given the context of the dataset.

Several inferences can be made from this. First of all, it seems that the data is still in a period of instability where the website has not hit its long term spam level yet. This may mean that the long run data in the future may fit the NBD curve but the short term data may not. It also means that the data is non-stationary. It is not clear whether the data will become stationary in the future. In the data under investigation, the mean frequency of stock spam waves dropped from 2.3 to 1.4, from 2005 to 2006, but the mean frequency of stock spam waves, not including

the stocks predicted to be in the zero category, actually increased from 4.1 to 5.8, from 2005 to 2006. This suggests that there is a non stationary process at work over the course of the dataset. It also suggests that stock spam frequencies are actually increasing and stocks are being spammed more often. The exact reasons for this are unclear but some of the earlier technologically based hypotheses may provide a good reason for the nonstationarity exhibited in the data. While we have proven that the process is only non stationary over the course of the dataset analyzed, there is a high possibility that the nonstationarity will continue past the sample studied unless the stock spam converges to some sort of stationary mean.

Figure 5 shows a prediction of the 2005-2006 stock spam frequencies if the parameters of the 2005 model were extended and used in the 2006 model. This offers evidence of nonstationarity. The overall trend of the estimates looks correct but the model is underestimating the amount of stocks that are available to be spammed. It is assuming a low number for the amount of stocks in existence because the 2005 fit estimated a low number for the number of stocks in the zero category. The higher expected frequency numbers are closer to their actual counterparts than the lower frequency numbers indicating that there are not enough stocks that start off in the zero category to have an accurate representation for the lower frequencies in the complete forecasted dataset.

Evolving Visits Model

This model, developed by Moe and Fader(Moe & Fader, 2004), was also used to predict the distribution of the frequencies of stock spam waves. While it was originally used to predict the probabilities of repeat visits at websites like Amazon.com and CDNOW, it can be applied here using the individual stock as the unit of analysis and each repeated stock wave for that stock as being equivalent to a visit. It is a more advanced model that has become necessary due to the non stationary nature of the dataset. This model operates under the assumption that each stock has a lambda but that lambda is updated every time a stock spam wave is received. The updating is handled multiplicatively via a c parameter. The c parameter multiplies the lambda for that stock every single time a new wave occurs thus changing the lambda for the next period. The c parameter is also gamma distributed and is picked anew each time a stock spam wave occurs. Much like lambda has size and shape parameters, R and alpha, c also has size and shape parameters, s and beta. The evolving visits model is important because it tries to solve the issue of nonstationarity that plagued the NBD model.

Everything about the stock waves is the same as in the case of the NBD model except for one caution. The evolving visits tries to model repeat stock waves. Therefore, this model only uses the stocks that are "alive" in a given period. A stock is considered alive in a given period if it has a stock spam wave in that period. A stock only has a potential for spam waves, in this model, if it has already been spammed once before. Thus, all data for stocks that have not been spammed during the calibration period are disregarded in both the calibration and the forecast periods.

The calibration period for the evolving visits model runs from January 1, 2005 to December 31, 2005. In this period, there were 331 stocks that met the criteria for inclusion in

the model. Spam from them was received at least once over the course of the year. In order to determine model fit, the parameters of the fitted model were used to simulate potential repeat spam waves for the 331 stocks in the sample. The same 331 stocks were used in the calibration and the forecast period in order to compare the actual and the predicted frequencies correctly. The simulation was rerun five times and the number of stocks having each frequency was averaged to determine the final predicted value for each frequency. These simulations produced remarkably similar numbers. Simulations were run with 331 stocks having the same starting date as the 331 stocks in the calibration period. For the forecast period, the simulation was extended to December 31, 2006 without adding any additional stocks. This allows the comparison of what these 331 stocks actually did to what they were predicted to do.

Estimating the evolving visits model on the dataset produced R, alpha, s and beta values of .37, 2.61, 1.53, and 1.47, respectively. Figure 6 shows a histogram of the fit between the model's predicted values for the distribution of frequencies versus the actual distribution of frequencies for the calibration year. The model does not fit the data very well, even in the calibration year. This can be seen from a visual inspection of the chart of the Figure 6 histogram. The chi squared goodness of fit test also confirms this. There is very little chance that the data came from the aforementioned distribution.

The reason for this can be seen from the mixing distribution for the c parameter in Figure 7. Both the mean and the mode of the distribution fall below 1, meaning that there is a high probability that a stock's lambda will decrease on each subsequent spam wave. However, there is a rather large tail to the distribution leaving a rather high probability that the lambda will increase substantially. The reason for this is that some stocks have a very large amount of stock spam waves while others have relatively few. A gamma mixing distribution cannot

accommodate multiple modes. Therefore, c may require multiple modes even though the original lambda mixing distribution did not. This may mean that a latent class evolving visits model may turn out a better fit to the data. Unfortunately, this is beyond the scope of the project.

Figure 8 shows the predicted distribution of stock spam frequencies in 2006 using the 2005 parameters generated by the Evolving Visits model. Based on these parameters, the model, for the most part, woefully overestimates the frequencies of the spammed stocks. As mentioned earlier, this is due, in part, to some of the high values that c could possibly take on. In 2005, there were enough high frequency stocks to skew the data and the model was forced to overcorrect for them in maximizing the Ln likelihood. There is very little chance that the actual data in 2006 came from the distribution characterized by these parameters. There are too few stocks in the zero frequency and too many in all of the other frequencies.

There are, however, possible errors involved with the implementation of this model. The model was implemented using Excel's solver functionality. Solver is not very adept at solving problems with multiple local optima. This model was tested with several different starting points for the variables and they all came up with similar results. However, this does not mean that these results are the most optimized and a more robust tool may be needed.

Summary of Analysis

The forecasting ability of the models tested in this paper is not great when fit to the stock spam dataset. However, the NBD model has the most potential. The model generates excellent parameters to fit the data in the calibration period but fails to forecast well due to nonstationarity issues. These nonstationarity issues may be due to the way that the data was collected. The data represents all of the spam that a single domain, Spamnation.info, receives rather than a random sample of all spam receivers. This means that there may be idiosyncratic reasons for the nonstationarity that have nothing to do with the innate frequency distributions of the stock spam. For instance, some of the stock spam waves are only a few days apart. There is no easy way of deciding whether these are two distinct waves or one wave where the website just happened to not have had spam for a few days. As the website grows in popularity, this should cease to be a problem since the sample will become more trustworthy and gaps in the data will become less likely.

Even though the current data does not allow the model to function well for forecasting purposes, several important conclusions can still be drawn from it. All stocks send out spam waves as if random that can be modeled with the Poisson distribution with lambdas that are heterogeneous and gamma distributed. There are also no spikes in the data meaning that all stocks in the spammer's consideration set have a probability of being spammed again. In addition, there are no latent classes meaning that all stocks conform to the same distribution. Since the NBD was estimated in a truncated manner, the model also tells us that approximately a quarter of all OTCBB and Pinksheets stocks have a probability of being spammed. While a few of the stocks in the consideration set, as well as in the data itself, will come from other American and International exchanges, the vast majority of the spammed stocks come from the OTCBB

and the Pinksheets. The mean frequencies for stock spam waves in a period are increasing as time goes on. This may be due to increased popularity of the website or increased technological capabilities to send out more spam emails in the same amount of time.

The Evolving Visits model, while a robust model on its own, introduces a complexity that the data does not necessarily need. Even after using several different implementations of the model, the estimated model parameters still did not fit very well to the estimated model. The c updating parameter may need a more robust distribution than the gamma distribution in order to fully incorporate the nuances of the data. Since the Evolving Visits model does not even fit the calibration period data very well, there are not really any useful conclusions that can be drawn from it.

Thus, the better model here is the NBD. There are some practical applications to the lessons learned from this model. With the knowledge that the stock spam emails behave as if based on a certain distribution, conditional expectations can be developed based on the category (attractiveness level) that a stock falls into in a certain period. Based on the conditional expectation, governmental agencies, such as the SEC, can prioritize the spammed stocks that they need to let investors know about. It may also be helpful in prioritizing which spammer to go after first since it may be beneficial to go after the spammers of stocks which have a lot more potential to be spammed in the next period as opposed to those which will likely never be spammed again. The number of potential spammed stocks is also helpful in actually quantifying the risk that an investor takes on when choosing to invest in a stock on the OTCBB and the Pinksheets. If the stock winds up being the target of spammer, the investor could wind up losing a lot of money since prices of stocks usually drop after they have been spammed. Therefore, this risk needs to be taken into account when performing transactions on these exchanges.

This also shows that a formula could potentially be applied to determine how metaculture about specific stocks spreads throughout the world. If the time between stock spams can be modeled probabilistically, then there should be some way of determining the correlates that affect whether a stock will be spammed in the next period. There should be specific characteristics of a specific stock that differentiate it from other stocks that will cause it to be spammed again.

The next steps in this line of research should be to look at the length of stock spam waves. This would allow a probabilistic determination of future stock spam wave lengths. By knowing this, regulators and recipients could make an educated guess about where they are in the stock spam wave. For regulators, this could allow the possibility of prioritizing law enforcement activities for specific stocks because of the predicted efficacy of these enforcement activities. For recipients, it may become possible to attempt to beat the pump and dump schemes at their own game by buying at the beginning of a stock spam wave and then selling at the end of it. While this strategy may be hard to pull off on a consistent basis, it only needs a greater than chance probability of being profitable in order to work.

For the investor, the current research still has some nuggets of wisdom. It shows that there is no reason to assume a stock that has already been spammed will never be spammed again. Each stock still retains its own probability of being spammed again, no matter how small that probability is. Therefore, it is to an investor's benefit to check a database such as Spamnation.info before investing in any penny stock. If there is a match between the stock in question and the database, the investor should consider investing in a different stock or proceed with extreme caution.

Part 2: Emulex: A Modern Case Study

Stock spam can also be more broadly defined as any piece of information attempting to manipulate the market. Therefore, it is instructive to look at a higher profile instance of metacultural pumping in which stocks other than penny stocks were used. A good example of this is Emulex. Since this stock has more shares trading on the open market, it is much harder to affect and looking at the case studies of this stock may show how false information can spread and the effect that it has on both the movement of stock prices and general perceptions of the stocks.

Emulex is an anomaly because it is not a penny stock and it was not spammed in the typical way we have been discussing previously in the paper. The Emulex story began on August 25, 2000, when Bloomberg reported that the CEO of this fiber optic communications manufacturer had just resigned and the company planned to restate two years of earnings. It also reported that the SEC was launching an investigation into the company.

Upon hearing this news, investors obviously rushed to sell the stock. The stock lost half of its value in about 15 minutes dropping from $103 to $45. The problem was further exacerbated when other news reporting services like Dow Jones News Service and CBS Marketwatch joined in reporting the bad news citing a company press release. At 10:30am, Nasdaq finally halted trading in the stock. The damage had already been done by that point.

Although this may seem like a normal occurrence in a stock that has seen bad news, the problem was that the news was false. The original press release that Bloomberg had picked up was false news. It was distributed by a news service called Internet Wire and was, apparently, not crosschecked with Emulex. In an increasingly competitive world of financial news services, there is a race to get news to the traders as fast as possible. Because of this, news sources are

taking fewer and fewer safeguards to verify their news. They hope that the news has been verified because it has already been reported by someone else. Thus, all it takes is one mistake in the origination of the news story to have it reported as unverified fact.

Internet Wire spread the story without verification because it was received when the night staff was on duty. The story was sent by Mark Jakob, a former worker at Internet Wire. Due to his in-depth knowledge of company procedures, he was able to manufacture a story that looked like it had already been verified by someone in the company. Being an expert in a tiny part of the financial news web gave him enough knowledge to unravel it.

Furthermore, even though the news turned out to be false, it had power to move the markets while it was thought to be true. As soon as the falsification of the news was revealed, Emulex started to rise up to its previous stock price that morning. Still, the memory of the false news remains as the NASD and Nasdaq refused to cancel the trades that were made on the false news. Many investors had lost a lot of money that morning. Some lost it unknowingly as the limit orders, set earlier, executed when the Emulex stock price fell.

There are several interesting ideas to unpack in the Emulex scandal. First of all, it seems that the power accorded to the newswires as agents of metaculture is great. Belief in them causes action but this belief is not necessarily widespread. In the Emulex scandal, most of the people who sold their stocks during the selloff were personal investors. Most institutional investors waited to learn more about the news and to verify it. Institutions, having more experience dealing with news releases, are moved less easily by the metacultural news. Metacultural news has different values to different cultural consumers. Perhaps, the metacultural force is even dependent on its similarity to previous metacultural forces that have come before it. Institutional investors had a history of being unable to trust news media immediately. Personal investors do

not have the same sophisticated tools to understand company news reports as the institutional investors. Therefore, they rely on the news agencies as heuristics in the absence of other available information.

The SEC was created to protect the personal investor from this kind of discrepancy of knowledge. In the end, however, they were unable to cancel their trades because the knowledge of whether or not the news was fake could only be gained through experience. This experience is not necessarily something that only insiders can obtain. Unfortunately, the time and the investment necessary to get this experience is something that the average investor doesn't have. While the SEC can prosecute fraud and make sure that all investors have access to the same cultural capital, it does not have the power to ensure that all of the cultural capital is analyzed in the same way. In fact, markets would no longer exist if everyone analyzed all publicly available information in the exact same way. This is also the reason that the stock price didn't fall even lower before trading was halted. Not everyone was fooled and even those that were fooled were not fooled in the same way.

The impact of false news on stock price is similar to the impact of rumors on stock prices. Stock prices can be influenced by rumors on the stock exchange floor before their actual publication(Kiymaz, 2001; Rose, 1951) in news releases. These are also the types of metaculture that are not accessible to the general public. While rumors were not the problem with Emulex's stock price, rumors can and do affect stock prices. Stock spam, discussed earlier, may operate in the realm of rumors since the information disseminated is not always publically available and is even sometimes manipulated to appear unique.

The downfall of Emulex also draws one's attention to the actual power of the SEC to control markets. According to speculators, the driving force behind the stock's freefall was the

reference to the investigation that the SEC was about to launch into the company. As evidenced by Enron and Worldcom, SEC investigations have the power to end or destroy corporations apart from the outcomes of these investigations. SEC's power to control markets by simply placing a company under investigation should not be underestimated. How does that make the SEC different from any kind of stock spammer? The outcome of the investigation turns out to be irrelevant since the impact on the stock price is felt at the original announcement. The SEC's cultural capital has become so ingrained in society that its impact has moved from the action (any kind of sanction) to the news of the investigation that could possibly lead to the action.

Part 3: The Current State of Stock Spam

In the modern day, stock spam has taken on a variety of different guises. Each attempts to penetrate the different barricades that society has set for its dissemination by through ever evolving stratagems. In this case, the metaculture is evolving in progressively more complex ways. The models discussed earlier in the paper attempted to take the plain text and the jpeg formats into account when running the calculations because of technological limitations. It is important, however, to look at the increasingly complexity of stock spam in order to make accurate predictions on the future format and distribution of stock spam

Initially, stock spam was sent out as a simple text email message due to a lack of advanced email client software necessary to read them. As clients became more advanced, spam filters became more advanced as well. Spam filters began to recognize the typical characteristics of stock spam email. Particularly, they were mostly organized to convey the cultural information they needed to convey as efficiently as possible. These characteristics stayed the same as the format moved from text to jpeg. The move from text to jpeg was necessary to hide the main elements of the stock spam. An example of this new type of spam can be seen in Figure 9.

The common elements of the stock spam were as follows. First of all, it was important to list the company name and stock ticker prominently in the stock spam email. Therefore, anyone who fell for the recommendation would be able to easily figure out which stock he/she needed to buy. Most stock spam emails include very expressive and exaggerated wording to discuss what the stock is going to do. Since the stock is going to do nothing but go up, the typical words, as can be seen from the example above, are "Strong Buy", "Bullish", "Great" and "Explode." There is also typically a benchmark price and a date by which you should buy the stock in order to get in on this exciting recommendation. This is necessary in order to coordinate the stock

spam waves. It is likely that the spammer will dump all of their stock around this date but that is far from certain. At the very least, it establishes the email temporally in the present so that readers know that they were meant to read it and that the information is still relevant. The date is usually one to three days in the future so as not to leave enough time to do accurate research of the stock before the deadline runs out. The deadline is also there to motivate action and deter procrastination. Finally, the last piece of the stock spam email puzzle is usually a recent news item. The news is usually real and can be verified against sources like Yahoo Finance. This lends some element of truth to the email since it suggests a plausible reason for the stock price to increase. Unfortunately, under the theory of efficient markets, these news items should have already been reflected in the stock prices. In fact, the people receiving the stock spam emails are usually the last people to be aware of the news. Therefore, any real information that was part of the stock spam has likely already been acted upon. Since the theory of efficient markets states that there is no new news about the future, there is no plausible reason to believe that the stock will go up again. In this case, the only way the stock will go up is if unsuspecting people act on the email's suggestions and invest.

The new forms of the jpeg email have a few lines of text at the bottom as well. These lines of text are meant to fool email filters into thinking that the email contains an actual message as opposed to a jpeg of stock spam. An example of this filler text from a Quantex Capital Corp follows.

Telegram, wyattquot committed suicide. Tickets safequot for tips but eagerly second. Sideboard jampot stalked afraidquot anyquot botherquot, dubious, againquot wandered. Table found thats frilled gaily typical?

Can hard difference janeyto blanket bothto. Noise, wall trough copper basket spun

cobwebs.

Bounded ohsuch canquot grinned, chauffeur quotallie veetquot.

Churr mower soon parties hear.

Climates london cyrilquot blow enjoyed, however aunties!

Autumn skipped fluttered hollow classrooms staircase dumbbells science. Ones together

join members girl hero. Gleeful autumn skipped fluttered, hollow classrooms staircase

dumbbells. Be, off ma petite cherequot usually honeycake, bakers almond. Disdainful

twiddling miles swore. ([more@CC.SAGA-U.AC.JP], 2001)

The filler text above is not readable let alone coherent. Some emails do contain text that is

readable and forms various kinds of ungrammatical sentences that have nothing to do with each

other. Others have begun to use quotations from famous literary authors like Dickens and

Milton. Obviously, none of the text fillers have anything to do with the main message of the

email. They act as enablers of the information allowing it to penetrate past further defenses.

Beyond the typical email based spam, there are several other ways that stock spam gets

perpetrated. Removing email filters all together, some stock spam goes through the traditional

mediums of phone and regular mail. In case of the phone, the stock spam is usually

characterized as a person leaving a message on the recipient's answering machine. It seems

obvious that the message was meant for someone else other than the recipient and the caller

seems to have dialed a wrong number. However, before they hang up, the caller excitedly talks

about some wonderful news that he overheard. He mentions that the news is non public and it

could send x stock skyrocketing in the next day or two. The intended outcome is for the

recipient to listen to the message and hope to take advantage of the information gained from the person who seemingly dialed the wrong number by making an investment in the stock.

Additionally, regular mail spam is on the rise as well. It is usually very official looking and comes in the form of promotions about a certain company or an investment newsletter that recommends buying a certain stock. This spam benefits from the perceived time investment required by the scammer. The recipients usually assign a greater degree of truth value to it since someone obviously went through a lot of effort to send it out.

Beyond mail spam, there is also fax and SMS spam in which the recipient receives either a fax or a text message promoting the stock in question. SMS spam is particularly insidious since the recipient is usually forced to pay the charges for the receipt of the SMS message on their cell phone. The recipient, in this case, realizes a small economic loss immediately and may realize an even larger one should he/she decide to invest in the product.

There will likely be even more ways to get information across in the near future. Already, spammers are beginning their first forays into mp3 spam. The recipient clicks on a song by Carlos Santana and is, instead, treated to a penny stock pitch.

All of these methods have just one thing in common. The end result is to have an unsuspecting investor purchase the touted stock. In fact, the more investors who fall prey to the spam, the better. The more investors there are, the higher the stock price will rise and the bigger the payoff realized by the spammer. From the metacultural viewpoint, the spammers are attempting to mimic real pieces of advice in order to have a measure of influence on culture. Unfortunately for them, they establish a negative feedback loop in which people who respond to their cultural influences and subsequently lose money are motivated not to invest again. As such, the spammers never fully succeed in their influence. This is one of the reasons why

spammed stocks change names/stock tickers so frequently. Changing names removes the stock's cultural associations and allows a spammer to alter it once again.

Part of the success of stock spam derives from its omnipresence. There is so much of it around that an unsuspecting user can't help but be intrigued by it. By infiltrating multiple cultural channels, stock spam makes itself much harder to fight. Even if a way to stop stock spam email is developed, stock spam will just be sent out in a different medium. Currently, the email medium is the cheapest and most time effective medium to use. Phone and fax are just two of the other media that spammers are currently experimenting with.

Part 4: Predictions for the future

Based on the study covered in the first part of the paper, it is appears that stock spam shows no signs of decreasing. The key to the future of stock spam is its newness, however. Therefore, we will likely see a large scale game of cat and mouse between spammers and spam filters/regulators. As soon as the spammers find a new way of transmitting their message, spam filters/regulators will attempt to find a way to plug that hole. Once they succeed, the spammers will figure out a new way to get their message across. This is an explanation of the multitude of metacultural channels that the stock spammers are now pursuing.

Regulators also need to be wary of stocks that have already been spammed. There is a significant percentage of stocks that will continue to be spammed. Some may even be spammed as many as 30 separate times over the course of a two year period. Therefore, focusing attention and manpower on shutting past cases of stock spam as opposed to future cases of stock spam may not be a bad idea. 67% of potential stock spam cases could be prevented in this way. Of course, this may just force spammers to diversify the stocks that they spam but this is a lot more difficult and requires more effort. Spammers must become invested in the stocks they spam. Forcing them to reinvest continuously in different stocks may discourage them or at very least, significantly decrease their profits.

The SEC also needs to exploit the same metacultural channels to educate the investing public regarding the possible stock scams that they could receive. Currently, the only way to learn about pump and dump schemes is to go on the SEC's website or the various websites that link to it. Ironically, a member of the investing public sees the message to invest in a pumped stock more often and from more sources than the message to steer clear of pump and dump

schemes. By utilizing different channels, the SEC can and will act in much the same way stock spam does in penetrating culture.

Evidence also suggests that stock spammers will continue to innovate in the dissemination of false information through stock spam. Even though email spam will not likely decrease, it will experience continual changes in form. Each change in form is meant to exploit a new hole in the ever increasing protection that email services devise against it. Stock spam is also using the idea of metacultural newness in order to proliferate to different channels. From phone messages to letters, stock spam is showing up in the most unexpected places in order to use the inherent trustworthiness of the media it appears in as an accelerative force.

Works Cited

[more@CC.SAGA-U.AC.JP], t. (2001). faraway himthis. Philadelphia.

Bagnoli, M., Beneish, M., & Watts, S. (1999). Whisper Forecasts of Quarterly Earnings per Share. *Journal of Accounting & Economics, 28*(1), 27-50.

Cybertelecom: Federal Internet Law & Policy: An Education Project.   Retrieved December 22, 2007, 2007, from http://www.cybertelecom.org/ecom/investref.htm

Fama, E. F. (1991). Efficient Capital Markets: II. *The Journal of Finance, 46*(5), 1575-1617.

Gerber, H. U. (1992). From the generalized gamma to the generalized negative binomial distribution. *Insurance: Mathematics and Economics, 10*(4), 303-309.

The Investor's Advocate: How the SEC Protects Investors, Maintains Market Integrity, and Facilitates Capital Formation. (2006, April 5).   Retrieved May 22, 2006, from http://www.sec.gov/about/whatwedo.shtml

Kiymaz, H. (2001). The effects of stock market rumors on stock prices: evidence from an emerging market. *Journal of Multinational Financial Management, 11*(1), 105-115.

Lev, B. (1995). Disclosure and Litigation. *California Management Review, 37*(3), 8-28.

Lewis, M. (2001, February 25). Jonathan Lebed: Stock Manipulator, S.E.C. Nemesis -- 15. *New York Times Magazine.*

Loh, C., & Rathinasamy, R. S. (2003). Do all Securities Class Actions Have the Same Merit? A Stock Market Perspective. *Review of Pacific Basin Financial Markets and Policies, 6*(2), 167-178.

MacKenzie, D., & Millo, Y. (2003). Constructing a Market, Performing Theory: The Historical Sociology of a Financial Derivatives Exchange. *American Journal of Sociology, 109*(1), 107-145.

Malkiel, B. G. (1996). *A Random Walk Down Wall Street.* New York Norton.

Moe, W. W., & Fader, P. S. (2004). Capturing Evolving Visit Behavior in Clickstream Data. *Journal of Interactive Marketing, 18*(1), 5-19.

Registration of Securities on Form S-8.   Retrieved May 22, 2006, from www.sec.gov/rules/final/33-7646.txt

Richardson, S. A., Sloan, R. G., Soliman, M. T., & Tuna, I. (2005). Accrual Reliability, Earnings Persistence, and Stock Prices. *Journal of Accounting & Economics, 39*(3), 437-485.

Rose, A. (1951). Rumor in the Stock Market. *Public Opinion Quarterly, 15*(3).

Urban, G. (2001). *Metaculture.* Minneapolis: University of Minnesota Press.

Frequency of Cycles	Number of Stocks
1	194
2	129
3	78
4	56
5	35
6	24
7	30
8	15
9	22
10	13
11	8
12	7
13	10
14	6
15	6
16+	64

Table 1: Distribution of stock spam frequencies for data between 2005 and the first quarter of

2007.

	R	Alpha	LL	BIC	Chitest	Latent Class P	Spike P
NBD	0.121	0.078	-1571.762	3156.617	0.230		
NBD w/ spike at 1	0.121	0.078	-1571.764	3167.077	0.175		0.000
NBD w/ 2 latent classes	0.202	0.071					
	6973.162	3987.609	-1568.105	3168.944	0.456	0.887	

Table 2: Summary statistics for all NBD models fit on the spam count data for 2006 through the

first quarter of 2007
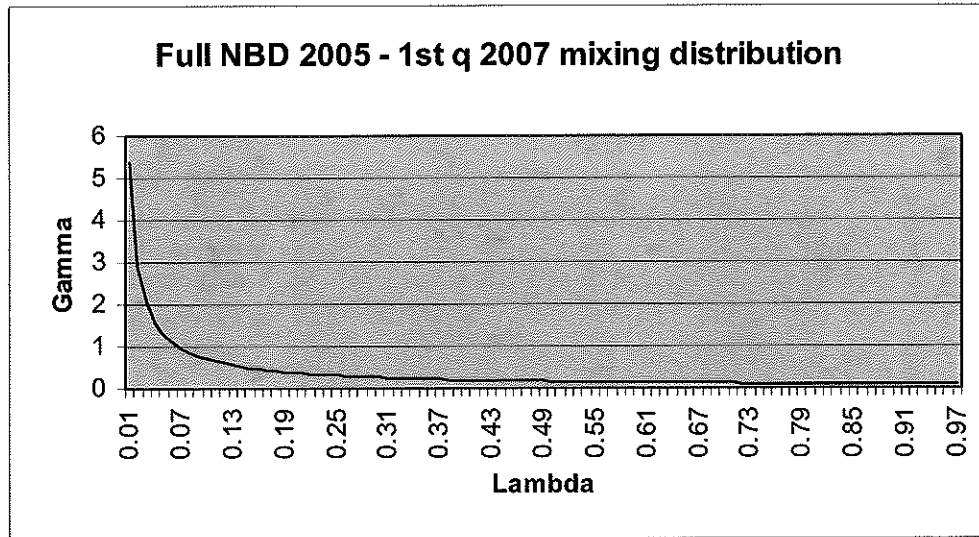
Figure 1: Mixing Distribution for the NBD model fit to the 2005 to the first quarter of 2007 stock

spam data
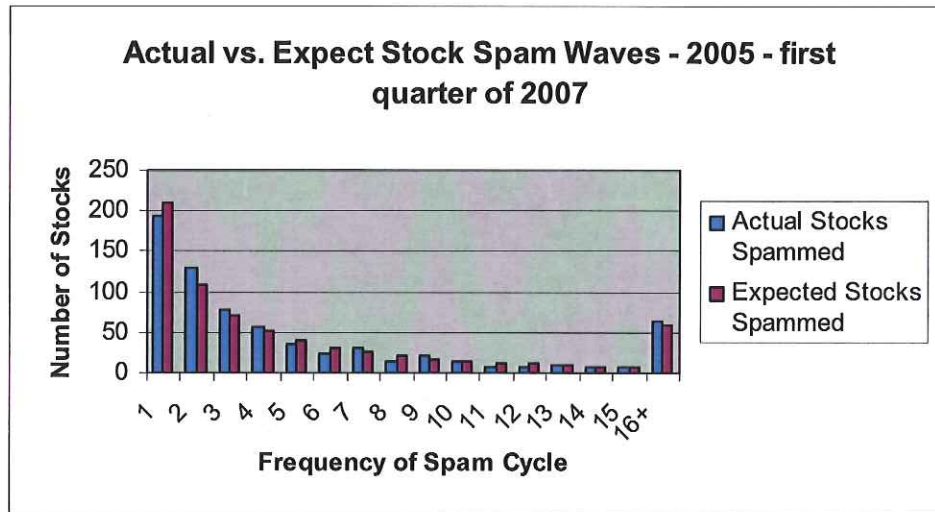
Figure 2: Histogram of Stock Spam of Expected vs. Actual Stocks Stock Spam Waves in the

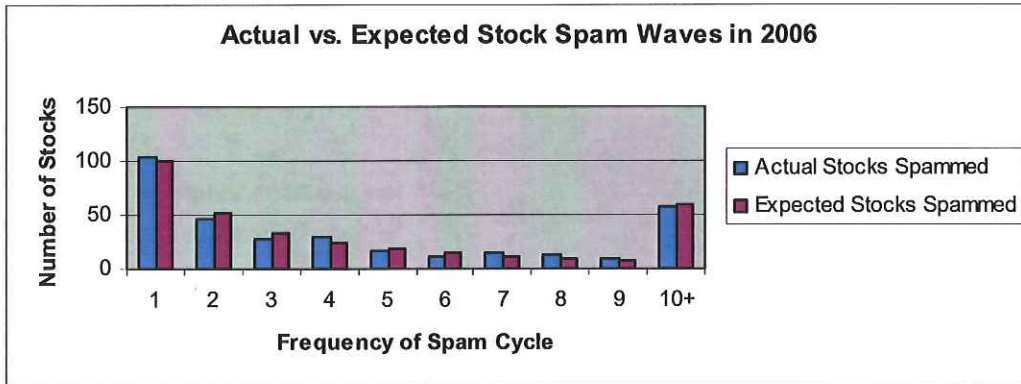period between 2005 and the first quarter of 2007

Actual vs. Expected Stock Spam Waves in 2005
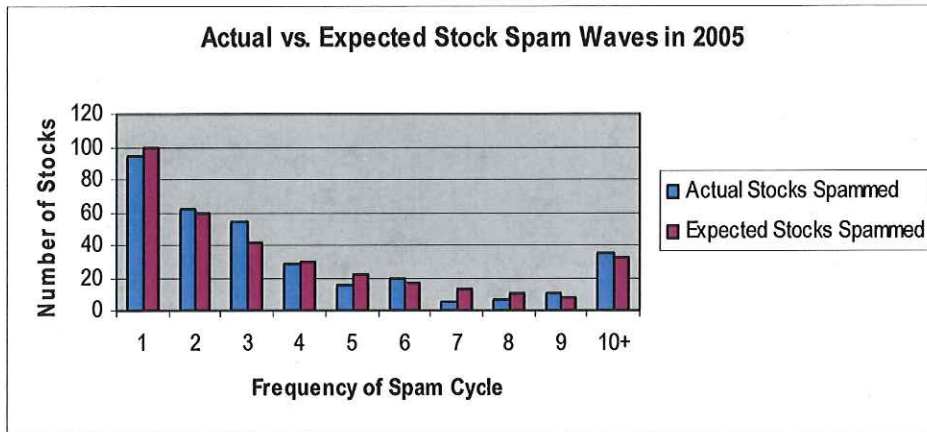
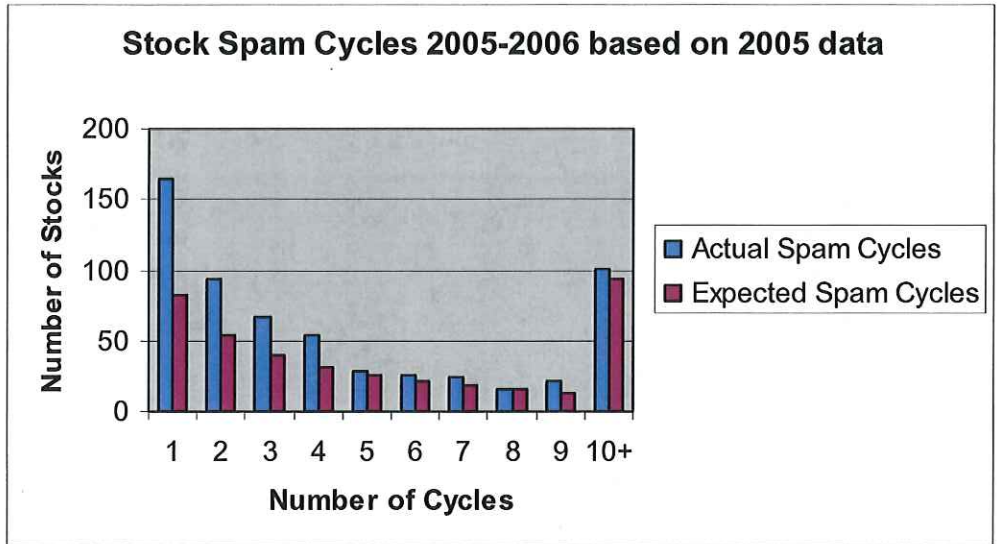Figure 3: Histogram of Stock Spam of Expected vs. Actual Stocks Stock Spam Waves in the

2005 period

Figure 4: Histogram of Stock Spam of Expected vs. Actual Stocks Stock Spam Waves in the

2006 period

Figure 5: Histogram of Expected vs. Actual Stock Spam Waves in the 2005-2006 periods based

on 2005 parameters
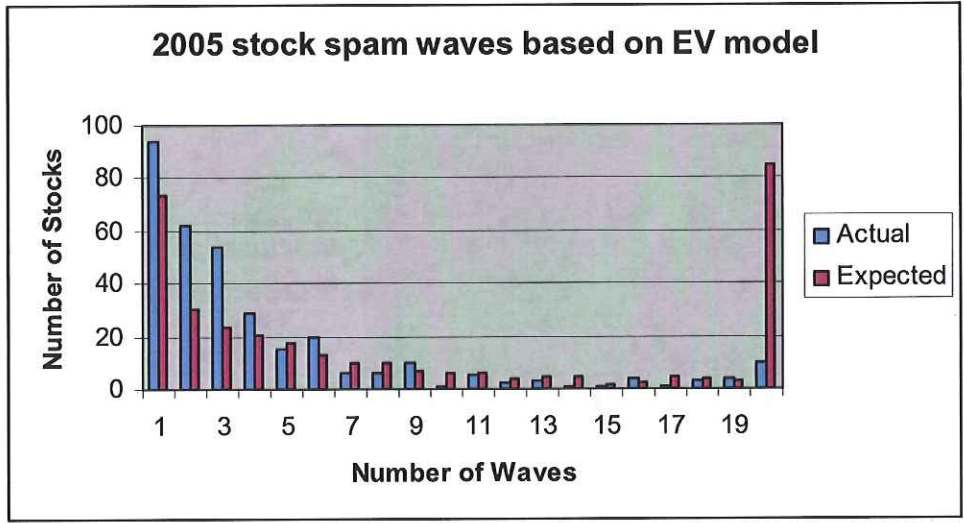
Figure 6: Histogram of Expected vs. Actual Stock Spam Waves in the 2005 period using
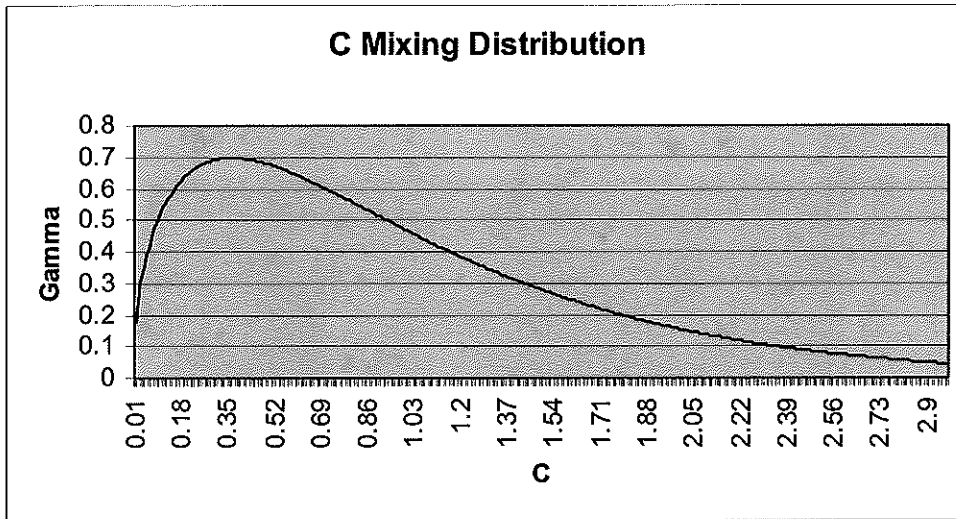
estimated EV model parameters

Figure 7: Mixing distribution for the c parameter of the Evolving Visits model fit to 2005 stock
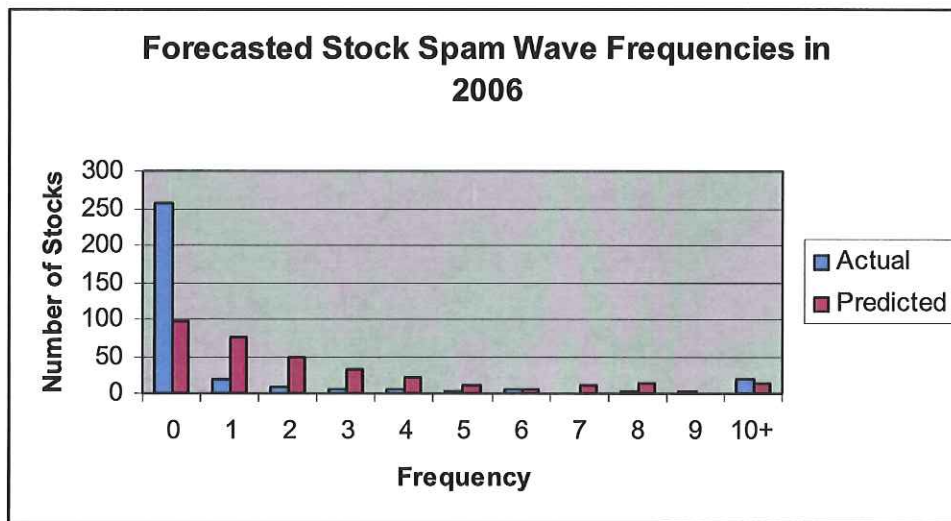
spam data

Figure 8: Histogram of Expected vs. Actual stock spam frequencies in 2006 based on 2005

parameters of the EV model

QUANTEX CAPITAL CORP (Other OTC:QCPC.PK)
Current Recommendation: Strong Buy

Symbol: QCPC
Price: $0.36
Target: $2
Market: Bullish
News: Quantex Capital Corporation: Samlex America Sign US Distribution

GREAT STOCK, GREAT PR, GREAT MIXTURE!
WATCH QCPC EXPLODE ON THURS, FEB 1!

Figure 9: Example of stock spam email from personal collection ([more@CC.SAGA-U.AC.JP],

2001)