# Forecasting Motion Picture Box-Office Returns and Analysis of the Hollywood Stock Exchange

By

James Cameron Davidson

Advisor: Rene Carmona

Submitted in partial fulfillment
of the requirements for the degree of
Bachelor of Science in Engineering
Department of Operations Research and Financial Engineering
Princeton University
April 17, 2006

I hereby declare that I am the sole author of this thesis.

I authorize Princeton University to lend this thesis to other institutions or individuals for the purpose of scholarly research.

James Cameron Davidson

I further authorize Princeton University to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

James Cameron Davidson

**Acknowledgements:**

I would like to thank my advisor Professor Rene Carmona for his understanding and help with a constantly evolving thesis.

I would like to thank my loving family – without their support I would not be where I am today.  Thank you, Mom, for the hours of thoughtful editing.  Thank you, Dad, for your insight and help.

Next, I would like to thank Laura Haas who kept me sane throughout this endeavor.  Your support helped me more than you know.

TABLE OF CONTENTS

Introduction: Forecasting, Gambling, and the Movie Industry…………..……... p. 1

Chapter 1: The Challenge of Forecasting Box-Office Returns……………………. p. 3

1.1: Uncertainty in Motion Pictures……………………………………..p. 3
1.2: Literature Review…………………………………………………...p. 5

Chapter 2: An Introduction to Prediction Markets and the HSX…………………...p. 10

2.1: Examples of Prediction Markets……………………………………p. 10
2.2: Efficient Markets…………………………………………………... p. 14
2.3: The Power of Prediction…………………………………………… p. 16
2.4: An Introduction to the Hollywood Stock Exchange………………..p. 17
2.5: The Hollywood Stock Exchange Market…………………………...p. 19

Chapter 3: Forecasting Motion Picture Returns…………………………………….p. 22

3.1: Motion Pictures as an Industry…………………………………….. p. 22
3.2: A Discussion of Relevant Data……………………………………..p. 25
3.3: Hypotheses…………………………………………………………p. 33
3.4: Method and Analysis………………………………………………. p. 36
3.5: Conclusion………………………………………………………… p. 57

Chapter 4: Analysis of the Hollywood Stock Exchange as a Forecast Vehicle…….p. 60

4.1: A Discussion of Hollywood Stock Exchange Data………………... p. 60
4.2: An Analysis of the Predictive Power of the HSX…………………..p. 61
4.3: Incorporation of Information within the HSX…………………….. p. 66

Chapter 5: Concluding Remarks……………………………………………………p. 75

5.1: Summary of Analyses………………………………………………p. 75
5.2: Potential Avenues for Future Applications…………………………p. 77

Appendix A: Frequency of Binary Variables……………………………………… p. 80
Appendix B: Example of Film Data that Trained Box-Office Model……………... p. 81
Appendix C: Examples of Prediction from the Trained Box-Office Model………. p. 83
Appendix D: Examples of HSX MovieStocks Trading Data……………………… p. 84

Works Cited……………………………………………………………………….. p. 86
End Notes…………………………………………………………………………  p. 89

# Introduction:

Forecasting, Gambling, and the Movie Industry

Wherever there has been uncertainty in human life, there has followed speculation, forecasting, and almost inevitably, gambling on that uncertainty. Whether it is the weather, the roll of a die, or the outcome of a football game, humanity deals with uncertainty in the future by making predictions. As computer models become more prevalent and powerful, society now relies on tools they build to produce their forecast and predictions for them. For the weather, one merely has to turn on a television or search the internet for a viable prediction. There are tools online that will predict the likelihood of nearly any outcome during a game, given all history of a football team.[1] In an effort to mitigate risk that is inherent in uncertainty, we seek to understand it and better forecast the future.

Our world is centered on forecasts, whether it is the weather or financial outlooks. Humans base nearly all of their decisions upon their views of the future – ranging from what clothes to wear based on the weather forecast to the makeup of their stock portfolios. A meteorologist's career is based on forecasting the weather. An investment manager's job is to forecast the future of companies. The problem with these forecasts is that they are all too often wrong.

As a result of people's attempts at forecasting, some believe they can more accurately predict the future than others. They attempt to benefit from what they perceive as their advantage and as a result gambling emerges. Gambling is no small activity as this industry had estimated gross revenues of $73 billion in 2003.[2] With the invention and adoption of the internet, gambling has gone online as well. Individuals now have access to casino games and gambling sites twenty-four hours a day from their own homes. Internet gambling was valued at $5.7 billion in 2003 and is projected to triple in value by 2009.[3] Man's fascination with uncertainty has led to his obsessive love of gambling and forecasting.

Uncertainty is also called risk just as often and gambling can in some ways serve to mitigate the risk of uncertain futures. As uncertain as any other human pursuit, the movie industry can prove just as dangerous. In truth, producing a movie can seem as much of a gamble as betting on the line of a football game. Ways for mitigating the risk inherent within movie production and distribution become necessary to allow for further development of the film industry.

# Chapter One:

## The Challenge of Forecasting Box-Office Returns

### 1.1: Uncertainty in Motion Pictures

To the casual moviegoer, movies are on par with a walk in the park or relaxing on the couch. However, creators and distributors know the true nature of the business. With such huge blockbusters like *Titanic* (grossing $1.8 billion worldwide), *Lord of the Rings: The Return of the King* (grossing $1.1 billion), and *Harry Potter and the Sorcerer's Stone* (grossing $968 million), it is possible to come to see motion pictures more as a business and less as art.[4] The movie industry seeks to sell us our entertainment and we are a fickle people.

Uncertainty is the name of the game in the motion picture business; where thousands of screenplays are in development, only about 450 or 500 get produced every year, and of these 450 productions, roughly half will receive a theatrical release.[5] In fact only three or four movies out of ten will even turn a profit.[6] That being said, the movie industry still can be very lucrative. When those three movies do hit commercially, they can be billion dollar-making blockbusters. Domestic box-office returns on motion pictures represented a $9.5 billion industry in 2004. Returns also have continued to grow dramatically, and are up 25% since 2000.[7]

A movie represents a huge investment on a studio's part, with production costs, actors' incomes, and advertising and marketing costs, which can account for nearly half of box-office revenues.[8] As a result, a way to forecast the returns for movies and mitigate the risk of failure would be critical to this industry. A reliable method to determine whether a movie would prove profitable would be incredibly valuable.

A large part of the uncertainty in the box-office returns of motion pictures is due to their status as a consumer product. While consumers can access reviews and receive information about the quality of a movie through word of mouth, the only true way to assess the quality of a film is to actually see it. This consumption is complicated by the very nature of marketing a movie. Studios often market their products aggressively with elaborate trailers, teasers, and powerful casts in an effort to create buzz and convince an individual to consume. Movies are a hedonistic, experiential good where the reason for consumption is pleasure, thrills, and entertainment, and not a physiological reward. So while movies can be perfectly marketed with top stars, it is ultimately the audience that determines the movie's fate. In the end, the movie represents itself as its own entity, the sum of all its parts, which lives or dies depending on the audience's response.

Movies are a complicated and difficult-to-understand product. Often movies are a one-use product, whose shelf life is only a few weeks long. After years of production and development, motion pictures have a small window in the marketplace and battle against a rotating bill of competitors. The infrastructure of release is currently experiencing dramatic changes due to the development of new technology. This coming change though makes it only even more important to have an accurate method to forecast a film's revenues in analysis of different media and timing of release.

Adam Smith's free market economic theory seems to be holding true for a new, developing type of market: Prediction Markets (also known as information markets, decision markets, idea futures, and virtual markets), which speculative markets trading assets, whose final values are tied to a future event. Researchers are finding these prediction markets to be incredibly accurate at forecasting the future. It seems that the aggregation of many semi-informed individuals with proper incentives is incredibly accurate.

One of the most famous prediction markets is the Hollywood Stock Exchange (HSX). With over 1.3 million traders, the HSX has changed from a play market to an important research tool for industry executives.[9]

## 1.2: Literature Review of Forecasting Box-Office Returns

Predicting the success of a motion picture has often been considered a "wild guess,"[10] even by famous researchers like B.R. Litman. This is due primarily to the difficulty found in forecasting the demand for the films. Predicting box-office returns has been done using two main models: **quantitative analysis** of factors that would influence box-office returns, and **behavioral models** of an individual's decision-making process with regards to patronizing the theaters.

One of the first, and arguably most influential, studies in predicting the financial success of motion pictures was Litman's first paper in 1983. His quantitative model used three groups of factors:

1) creative characteristics of the movie, including star power, director's prestige, genre, and MPAA rating;
2) scheduling and release factors, such as release time, competition, number of theaters, and distributor's strength; and
3) marketing effects, including advertising budgets, reviews, and awards.

Using a multivariate regression analysis, Litman makes a "revenue equation." Removing variables with no statistical significance, he found the remaining variables have an $R^2$ value of nearly .5, demonstrating that they explained half of the variance of the dependent variable (the box-office returns). In his later studies he continued to use his linear regression modeling to develop a more comprehensive set of explanatory variables.[11] Based on the results of his regression models, he goes on to explain that the movie industry might not be as uncertain as previously believed.[12]

Sharda and Delen's quantitative model used a similar, but not identical, set of explanatory variables. Instead of using a regression model, Sharda and Delen used multi-layered neural network architecture. As their dependent variable, instead of predicting an actual dollar amount, they placed box-office returns into nine distinct return classifications between "flop" (<$1M) and "blockbuster" (>$200M). They trained their network with a data set of approximately 600 movies. Their network correctly predicted the box-office returns to the correct "bucket" 30% of the time and within one adjacent "bucket" 72% of the time.[13]

Sawhney and Eliashberg developed a model for forecasting box-office returns based on beginning box-office return data. They modeled the customer's decision to patronize a movie in two steps: (a) the "time to decide" variable, an indicator of a consumer's decision of whether it would be valuable to patronize the movie (whether he would want to), and (b) the "time to act" variable, an indicator of the consumer's decision to actually go to a film. Their model used similar variables as previous studies, including genre, special effects, sexual content, rating, star presence, and reviews. Using 101 movies to build their regression model, they tested their model on 10 movies. Without taking into account early box-office returns, their model has a $R^2$ value of .42 for their final set of significant movie attributes. Predictably and perhaps without merit, their model becomes dramatically more accurate when it considers more of the early box-office data. Their model produced a mean standard error of 71.1%, 51.6%, 13.2% and 7.2%, for no data, one, two, and three weeks of data, respectively. Their model of a "time to act" indicator and a "time to decide" indicator are only weakly predictive with $R^2$ values of only .2 and .1, respectively.[14]

De Vany and Walls (1999) expanded on their 1996 paper which found that motion picture box-office returns converge on a Pareto distribution. Their work centers on the effect of stars on the expected box-office returns. They found that only 19 stars are positively correlated to the likelihood of a hit movie and that despite this correlation, each star still carries significant risk. Using a sample of over 2,000 movies, they find that the mean return of the box-office is dominated by increasingly rare blockbuster movies. They conclude that "anything can happen," meaning that one cannot accurately predict revenue and "nobody knows anything," indicating that since movies are involved in such a complex exchange of information among customers, it is impossible to attribute the success of a movie to casual factors.[15]

These studies break the casual factors down into three distinct categories: Creative Aspects, Studio Actions, and Non-Studio Factors. **Creative Aspects** include casting, genre, rating, and whether the movie is a sequel. Popular knowledge suggests that the biggest stars draw the biggest crowds and as a result would yield the highest box-office returns. Contrary to this belief, De Vany and Walls (1999) have raised questions with this conventional line of thinking. Genres such as action and comedy have been shown to draw larger box-office revenues. This may be because they have stars attached or because they have larger production costs. Other studies have also shown that genre actually has little to no predictive ability. Ratings of R or NC-17 that restrict a customer's ability to view the movie can be detrimental to box-office returns, shown by both Litman (1983) and Sawhney and Eliashberg (1996).

**Studio Actions** include production budget, advertising budget, release date, and number of screens. Litman (1983) reported a positive correlation between production budget and box-office revenues, but a question of causality remains. Advertising budget also has been shown to be a key predictor by Elberse and Eliashberg (2003). The relationship between box-office returns and advertising has not been fully explored. Intuitively, studios will likely market and advertise for movies they think will be more successful. Despite this economic truism, it could be that the advertising does in fact drive box-office returns by encouraging consumers to patronize the movie. Release date has been shown by Litman (1983) to be very important, as films released during the summer or during the Christmas holidays have significantly higher returns. Logically, the number of screens a film is released on will highly correlate with box-office returns, as shown by De Vany and Walls (2002). Yet if the intent is to mitigate risk before production, the value of this indicator is questionable.

Finally, the previous work considers **Non-Studio Factors**, which include reviews, perceived quality, and at times, initial box-office information. Both Litman and Ahn (1998) and Sawney and Eliashberg (1996) show that reviews correlate with box-office results. It is interesting to note though that causality is again in question. Are the reviewers' expert and accurate portrayals adept at forecasting a movie's success or do reviewers affect consumers with their biases? A number of studies, including that of De Vany and Walls (2002), use an indicator of the public's perception of the movie and the nature of word-of-mouth communication of this information.

# Chapter 2:

## An Introduction to Prediction Markets and the Hollywood Stock Exchange

### 2.1: Examples of Prediction Markets

With the invention and adoption of the internet, online trading markets have also begun to grow. One type of online market that has developed is the prediction market, also known as information market, decision market, idea future, and virtual market. Prediction markets are speculative markets where assets, whose final values are tied to a future event, are traded. Prices of these assets can then be analyzed based on the market's prediction about the likelihood of the future event.

Prediction markets can be used to mine a collective or market estimate of the expected value or probability of a random variable or future event. The price an asset trades at reflects information dispersed across the entire population of traders. Nonetheless, the market prediction is not just an average of individual opinions. Instead, it is a complex integration reflecting the competition of traders as they obtain and use information and act in response to other individual traders obtaining and using their own information. This game theoretic interplay allows for a complex method that integrates a vast amount of relevant information. The actions of traders enable prediction markets to aggregate information from individuals and incorporate it and other sources of information within the market. By their trades, the participants weigh this information by reaching agreement on a purchase and divestiture price. The price of an event reflects the consensus probability that the event will occur.

Perhaps the most famous of these prediction markets and the one that sparked the relatively new fascination with their use would be the Iowa Electronic Market (IEM).[16] The Iowa Electronic Market was created in 1988 at the University of Iowa's Henry Tippie College of Business. The market was created as an educational and research tool where participants could invest small amounts of money ($5 - $500) on futures that paid off according to future political events. The initial logic behind its creation was that individual voters had no incentive to answer truthfully when being polled. This betting exchange could provide incentive to traders and capitalize on their participants' incentive, which might allow for a means to aggregate the market's information and correctly predict the outcome of elections. Prediction markets reached the public eye when the Iowa Electronic Market correctly predicted the outcome of the 2004 presidential election. The Iowa Market continuously predicted the results of U.S. presidential elections with greater accuracy than that of the leading polling companies.[17] The Iowa Market allows real gambling on future events like the outcome of elections. Berg, Forsythe, Nelson and Reitz (2000) summarize the results of the IEM and demonstrate that it not only yields very accurate predictions, but on the whole outperforms major polling organizations.[18]

Pricing within another event futures market, Tradesports.com, is similar to that of the Iowa Electronic Market. For example, in the last California gubernatorial election, there was an "Arnold Schwarzenegger victory" contract that would pay out $1 dollar if he won. The price reflected the market's view of his likelihood of victory. If the Arnold contract trades at 50 cents, it would imply that the market believes his chance of winning is 50%, while a price of 70 cents would reflect the belief that he has a 70% chance to win.

Tradesports.com is likely the most famous and often used prediction market that currently exists.[19] Owned by the Trade Exchange Network, Tradesports.com (like the IEM) allows real gambling on future events, although it is predominantly focused on sporting events. Tradesports.com and similar websites take the lead in defining the scope of a contract, but then allow users to post offers to sell and accept offers to buy from other users. Researchers like Telock (2004) have used Tradesports.com to trade data that generalize an individual's perception of utility, specifically with the demonstration of a reverse favorite/ long-shot bias.[20] He cautions, however, that because one set of prediction markets may demonstrate inefficiencies does not mean that it will generalize and apply to all information markets. He shows that while in fact Tradesports.com's sports betting demonstrates a reverse long-shot bias, on the same prediction market the financial betting is in fact much more efficient, despite even less liquidity and trading volume.[21]

Trading in only play money, another series of prediction markets operate as online games. Yahoo! Research teamed up with O'Reilly and together built the Tech Buzz Game. They describe it as "a fantasy prediction market for high-tech products, concepts, and trends."[22] Playing for pride or fun, these traders bet on what rising technologies will be searched for on Yahoo! the most frequently. While it may seem intuitive that play markets will in fact be much less accurate than their real money counterparts, Servan-Schreiber (2004) of NewsFutures, Inc. found that in fact pride is a good motivator and provides enough incentive for accuracy.[23] In fact, because the only way to build up more capital in fantasy markets is through skill, the participants that control and can influence the market the most are proven to be more accurate in the past and will likely increase the accuracy of the overall market.

Another fantasy prediction market, mentioned before and expanded upon later, is the Hollywood Stock Exchange (HSX). It is a play online market game in which participants can buy and sell shares predicting the future for movies, actors, directors, and other film industry-related options. Like the Iowa Electronic Market, the HSX has also been remarkably accurate. In a famous prediction, the Market correctly predicted 35 of the 40 major Oscar nominees in 2002 and all of the nominees in 2004. [24] The accuracy of these markets' predictions has provoked a significant amount of research and speculation about the power of a large number of only semi-informed individuals. James Surowiecki's 2004 book, The Wisdom of Crowds, has touted the power of these markets to aggregate information.

Many companies, including Google, have begun tests to determine the feasibility of implementing a modified market as a useful decision support tool. Google has built a number of markets to forecast product launch dates, estimate office opening expense, and other factors of strategic importance.[25] It has been argued that the huge success of the Google search engine has in fact benefited from a prediction market-esque property. It uses the breadth of the internet and the interconnection created by individuals, their "wisdom from the crowd," to vote for and value individual websites.[26]

The U.S. government, via a project within the Defense Advanced Research Projects Agency (DARPA), also speculated that prediction markets may be useful as an indicator of future terrorist attacks. They proposed to establish a "Policy Analysis

Market" that would allow for trading in forms of geopolitical risk. Because of the controversial nature of the subject, it was criticized as a "terrorism futures market" and subsequently the Pentagon canceled funding for the program. It was feared that such a market would in fact create an immoral incentive for terrorists.[27]

With all of these prediction markets, the question of efficiency and accuracy remains. Economists have toiled over financial markets for years, but since these markets are primarily research-oriented, the opportunity to examine them and apply them to a commercial setting (namely, the motion picture industry) is still open.

## 2.2: Efficient Markets

There is an old financial joke that demonstrates what is known as the efficient market hypothesis. An economist is walking down Wall Street and sees a twenty dollar bill on the sidewalk, but decides not to pick it up. He decides this not because he is too rich to be bothered, but instead because he had concluded (without needing to verify) that the bill is a fake. His reasoning goes along the lines that if it were a real twenty dollar bill, then someone else would have picked it up already. A Wall Street trader's take on the joke depicts the economist walking past the bill, while a trader following him picks up the twenty and leverages it in a complex play to take the economist's money.

Roughly speaking, the efficient market theory says there are no free lunches in financial markets. The price of a security reflects a rational assessment of its true value after taking account all available information. Under the economic theory of rational expectations, information incorporation is accounted for in markets. The rational expectations theory suggests that prices reflect the sum of all information available to all market participants.[28] Even if participants with access to exclusive inside information exist, the price in the market will equilibrate to exactly the point as if all participants had this information. The rational explanation is that the price offered and bid upon reveals such private information to all ignorant participants.

The Efficient Market Hypothesis was developed in 1960 by Eugene Farma and is generally represented by three forms. The first and weakest is that all historical trading information is priced into the security. This means that an analysis of the past prices, volatilities, volume, etc. will provide no further advantage. The second form, known as the semi-strong form, is that all publicly known information (quantitative and qualitative) is priced into the value of the security. The final and strongest form states that all public and private information is priced into the security. This would mean that even in the instance of insider trading, no advantage is available.

Some research has shown that prediction markets exhibit the characteristics of an efficient market and the information aggregation properties of rational expectations. Pennock, Debnath, Glover and Giles (2002) have demonstrated that online exchanges, and specifically the HSX, demonstrate signs of efficiency, notably price coherence and forecast accuracy.[29] As a result, the prices have been shown to be an accurate incorporation of all the information on the security.

## 2.3: The Power of Prediction

Prediction markets and betting markets have recently been studied as a means of insight into human nature and as a means of information aggregation. Research into prediction markets has been similar to that of financial markets and one can expect to see similar results as well.[30] Hanson, Oprea, and Porter (2004) examined the effect of manipulators on predictive markets. They found that in fact manipulators will not affect the accuracy of prices. Participants that are not manipulators set a different threshold where they are willing to accept trades and as a result protect the market from bias.[31] In an earlier paper Hanson and Oprea (2004) showed that manipulators and insiders increase the informational content represented within a market.[32]

Pennock et al. (2002) developed a model of information incorporation and demonstrated that relevant events are reflected immediately and correspond to market price swings.[33] Using metrics like average logarithmic score and entropy of a security, Pennock determines the informational content incorporated into a market. The expected entropy loss is synonymous with the expected information gain, first demonstrated by Shannon.[34]

Prediction markets can suffer the same inaccuracy as other more standard markets. Liquidity and other factors that skew the risk profiles of the market can work to distort implied market probabilities. It is important to take these potential pitfalls into account to make sure that individual securities are, for example, liquid enough not to suffer from the mis-pricing by a few participants. Prediction markets, like any financial market, are also slaves to their traders. Any irrationality or inefficiency demonstrated will be a result of the perceived utility, irrationality, or ignorance of its traders.

There have been numerous studies on horse track betting [like Jullien and Salanie (2000)] that have demonstrated the "favorite / long shot bias." [35] The bias causes the expected returns of betting on a horse to increase with the probability of the horse winning. It has been speculated that expected utility gamblers understand that winning big with a long shot outweighs the more likely win on a favorite. In the event of winning it big, the reward is great and gamblers can pay off all debts. If they lose, they only risk a small amount with little expectation to win. Woodland and Woodland (2001) demonstrated reverse "favorite / long shot bias" in betting on the National Hockey League,[36] where people place too high a premium on the favored team and overestimate their chance of winning.

## 2.4: An Introduction to the Hollywood Stock Exchange

The Hollywood Stock Exchange was founded in 1996 by Max Keiser, a former stockbroker, and Michael Burns, a former investment banker. They later sold their market to Cantor, a financial services company. It is a fantasy prediction market that allows traders to invest in MovieStocks, StarBonds, and AwardOptions.

The success of the Hollywood Stock Exchange stems from the core of interested and informed traders. As Servan-Schreiber, Wolfers, Pennock, and Galebach (2004) demonstrated, in order for a fantasy market to be accurate, it needs motivated and knowledgeable traders. Because of the sheer size and breadth of the movie industry, motivated and knowledgeable traders are exactly what the HSX has. Over half of the country's population attends the movies.[37] Many people are passionate about movies and the HSX offers them a venue to exercise that passion.

Upon joining the Hollywood Stock Exchange, an individual is given 2 million Hollywood Dollars ® (H$). The Hollywood Stock Exchange allows trades in MovieStocks (which represent the future box-office success of releases), StarBonds (which represent the future popularity of celebrities), and Award Options (which pay off in the event an actor receives an award), all in a fictional money market game. The MovieStocks are interesting in terms of potential research because they are a prediction of a very quantifiable result. In pricing the MovieStocks, the market participants take into account a large amount of information that they then translate into a price at which they would be willing to buy or sell. This market has been demonstrated to efficiently aggregate available information into an accurate estimation of box-office returns.[38] The success of prediction and the accuracy of these MovieStocks have begged the question: what specifically contributes to and causes its precision?[39]

The ability of this market to aggregate information is a large part of what has made it so interesting to researchers. There is a significant amount of information involved in the process of pricing the future returns of a movie, including individual taste, public response to advertising dollars spent by the producers, number of theaters in which the movie opens, reviewers' outlooks on the movie, and more. Looking at the market's ability to aggregate this information, as well as its ongoing reappraisal due to new information, provides insight into the prediction market's ability to aggregate information.  This is the underlying reason why it has been so successful at predicting the future.

The HSX value lies in its capacity for supplementing research materials.  Because of the demonstrated accuracy shown by the HSX in the past, the HSX has become an important source of information to the movie industry.  In just a few years the HSX has gone from an experimental online game to having a prominent seat at the Oscars.

## 2.5: The Hollywood Stock Exchange Market

While the price of a MovieStock is subject to the whims of the market's traders, the price is ultimately grounded to the box-office returns because the stock is delisted after the movie is in theatrical release for four weeks.  Shareholders receive H$1 per share for every US$1 million the movie grossed in the US domestic market.  Traders buy (short) stocks they believe underestimate (or overestimate) the movie's future performance.  The price of a MovieStock is the market's forecast of the movie's future four weeks box-office gross following its release.

New MovieStocks become listed through the HSX Initial Public Offering system. During a movie's lifetime, if successful, it will go through five stages: concept, development, production, wrap, and release.  A MovieStock is listed long before it is a sure thing to be released and often there are just concepts and rumors flying, without any true plans for development.  As a result, uncertainty is priced into a MovieStock from its Initial Public Offering.

The price of a MovieStock adjusts after its first weekend in wide national release (at least 650 theaters). On the Friday before opening weekend, all trading is halted; on the following Sunday, the price readjusts to H$2.8 times the movie's actual weekend box-office gross (in millions). This halt is due to the fact that some individuals with greater access to box-office information can unfairly capitalize on this asymmetry of information. This 2.8 factor is used to predict the movie's four-week gross based only on its opening weekend. As a result, the market price before and up to the halt and the subsequent readjust reflect the traders' prediction of 2.8 times the opening weekend's gross.

In an effort to mimic realistic components of other financial markets, on each trade a 1% trading commission is deducted from the total amount of the trade. Although on Saturdays, this fee is waived. There are also rules governing the maximum number of shares (50,000) an individual can own of an individual MovieStock in order to limit the amount of influence a single trader has over a particular security. Because this market is merely an online game, insider trading is allowed and almost encouraged. It has been speculated that this is another reason that the market is so accurate. It seeks to incorporate all possible information into MovieStock prices and, as a result, may be a more realistic reflection of reality.

Key to maintaining the market is the HSX's Virtual Specialist, which acts as a market maker on all the market's securities. It provides liquidity and maintains order in the market. Unlike other prediction markets where buyer and seller match up and find a price for a transaction, the Virtual Specialist is always on the other side of every trade. The Virtual Specialist maintains its own portfolio and can create shares of a particular security on the fly, but it does not always do so and will sometimes match players up. When buyers and sellers act, the Virtual Specialist uses a mathematical system to control the changes in price to reflect the supply and demand for that security. In doing so, the Virtual Specialist controls the volatility of the securities. The Virtual Specialist is a perfect solution to the potential lack of liquidity in the market. As a result, the HSX avoids a lot of the pitfalls that befall other markets with a lack of liquidity. When securities are illiquid, they can be grossly mis-priced. Liquidity in a security increases competition and narrows the spread between bid and offer prices. Because of the increased competition, the price will be a more accurate reflection of the underlying asset. This Virtual Specialist works by matching up two or more traders seeking to purchase and divest a number of shares. Subsequently, the Virtual Specialist credits their accounts or employs its own portfolio and acts as a market maker. The Virtual Specialist also has the potential to create new MovieStock shares on the fly due to demand.

# Chapter Three:

## Forecasting Motion Picture Returns

### 3.1: Motion Pictures as an Industry

The film industry is first and foremost an information industry. A movie is essentially the production of an informational good and is sustained by the exchange of information. A movie lives and dies on the exchange of information within its marketplace, in the form of the reputations of its actors and film creators, and the standing of the creative content of the individual movie. Creative elements within a film are all important when they are hired only for short periods of time and each can provide critical input to the final product. As a result, their reputations for success are crucial to marketing themselves within the industry.

The motion picture industry is primarily comprised of three distinct steps: production, distribution, and exhibition. As their title implies, producers deal with the production of the movie and any aspects that falls within its creation. Distributors serve as the means to move the complete produced movie through a nationwide distribution network. Exhibitors take the final product and display it in theaters they own in which the end consumer will view the film.

Movies typically spend several years in production. Producers assemble groups of individual actors and craftsmen on a temporary basis, as these groups last only for the duration of the project (be it filming, editing, music scoring, etc.). There are many steps to the process, with each less and less likely to succeed. The production costs are sunk at the completion of the movie and are then unrecoverable. Even a finished movie is not guaranteed to be seen on the big screen. The production and marketing of a feature film represents a huge investment and assumption of enormous risk. According to the MPAA, the average cost of a feature motion picture as of 2001 was $47.7 million and had increased to $64 million by 2003.[40]

The production of movies is very costly and risky, and unfortunately the vast majority of studios never recoup these costs. Desai, Loeb and Veblen (2002) describe the three main risks that a motion picture faces as completion risk, performance risk, and financial risk. Studios face completion risk due to the huge burdens of investment required and the malleable relationships between producers, creative talent, and financiers. Films face performance risk due to uncertainties regarding the public's and critics' views of stars, creative content, competition, and other factors that make it incredibly difficult to accurately forecast box-office revenues and profits. Unfortunately, it seems that every movie is unique in terms of characteristics in production and the environment it is released within. Because a film is such a difficult financial instrument to predict, investors face significant financial risk as well. As one can see from the MPAA released numbers, film investors are now required to invest more and more money to compete with increasing production and marketing costs. As a result of this increasing financial risk, investors are requiring larger returns in order to make it a profitable investment. Investors are also seeking ways to begin to mitigate these powerful risks. One example of this is the decision by studios to finance and produce established intellectual content in the form of franchises.

After the initial production, the next step in the process toward the consumer is the film's distribution. The main issues that distributors face involve the release date, the scope and location of the release, the development and implementation of a national advertising campaign, and finally individual contracts with exhibitors. The decision regarding timing of the release date considers the seasonality of demand for entertainment and the issue of competition throughout the movie's run. Release date timing is especially important since the performance during the first week of release for an individual film accounts for nearly 40% of total domestic revenues.[41]

> *Fueled by popular media outlets fascination for reporting on "the number one movie" in any given week, securing a successful opening weekend has become almost a necessary requirement for long-term success across all windows. A movie that fails to open strongly almost always loses the attention of the media, audiences, and exhibitors. Timing the opening carefully therefore is crucial.* [42]

Scope of release refers to the number of theaters in which a film will initially open to the public. There are three main types of release: wide, limited, and platform. Wide release indicates that the movie will begin screening on several thousand screens with an extensive advertising campaign. Limited release is where the film is released only in a few cities (primarily New York and Los Angeles), often without strong expectations of a national wide release. This can serve to build up what is called "buzz," or "word-of-mouth," that can be essential for movies without a clear national appeal. Word-of-mouth communication is typically viewed as critical to the formation of demand for films as products. Logically, word-of-mouth advertising is a particularly important driver of the success of entertainment goods because such goods are often consumed collectively and serve as important topics in daily conversation. Platform release is when a picture is released in a smaller number of theaters (commonly in bigger cities) with a smaller local advertising campaign.

### 3.2: A Discussion of Relevant Data

Based on a rational analysis of the mechanics of the motion picture box-office market and an examination of research done in forecasting motion picture box-office returns, the following variables will be used to build a forecast model. The factors that are believed to influence a film's performance are genre, MPAA rating, director's financial success, actors' financial success, franchise, opening date, an estimate of production budget, an estimate of advertising budget, number of theaters in which the film opens, and the competition the movie will face upon opening.

Because motion picture production companies are primarily private corporations with strong incentives not to release certain sensitive information due to competition, data is often difficult to uncover. As a result, the data for this analysis was collected from three websites, The Numbers (www.the-numbers.com), Box-Office Mojo (www.boxofficemojo.com), and the Internet Movie Database (www.imdb.com), which represent the most respected authorities on box-office and film information. The Internet Movie Database categorizes films and contains MPAA rating information, full cast lists, genre, and other information. The Numbers focuses on financial information and contains box-office information, theater counts, budget estimates, and so forth. Box-Office Mojo contains a combination of both forms of information and was used when one or the other source did not contain a piece of needed information. Four hundred data points were collected, representing the 100 top grossing movies of 2005, 2004, 2003 and 2002. Again, because the production companies are private companies, information for smaller movies was unavailable all too often and, as a result, the data set was limited to the top 100 grossing films in a given year.

The dependent variable in this study will be domestic box-office revenues; this does not include auxiliary distribution media, such as video rentals, international markets, or derivative products like soundtracks. I used many of the same explanatory variables as found in the studies discussed above; however, I consider most of them in a different light. A description of these variables follows, along with an explanation of why several variables considered above are not relevant here. The variables are broken into two primary classifications: **Creative Aspects**, dealing with the aspects of movie that are more art than business, and **Studio Estimates**, dealing with the aspects of most concern to a studio shareholder, the financial numbers.

Within **Creative Aspects** I consider five explanatory variables: genre, MPAA rating, director's previous financial success, actors' previous financial success, and franchise. The primary interest is the target audience of the film. The model I am developing captures this factor using two indicators, genre and MPAA rating. These provide a good approximation of the general audience that will be interested in a given movie and targeted by advertising.

*Genre* has been shown in the past to have little to no statistical significance [Litman (1983), Litman and Kohl (1989)]; however, various studies have demonstrated that in fact certain genres are at times profitable and significant. A question of causality is implicit within an analysis of genre because certain genres (such as Science Fiction or Animated) may require a more powerful and well-known cast, a greater production budget, and so on, and therefore may receive higher box-office returns. Nevertheless, testing genre should reveal whether it has a significant impact independent of budget.

Specifically, each movie was classified according to IMDB's categorization in any combination of fifteen genres: Action, Adventure, Animated, Biography, Comedy, Crime, Drama, Family, Fantasy, History, Romance, Science Fiction, Sport, Thriller, and War.[43] I have represented each of these genres by a binary variable. It is important to consider genre as an influential creative aspect of box-office receipts, based on the past research and an understanding of the mechanism that the motion picture industry employs to attract consumers.

*MPAA Rating* is another commonly used explanatory variable when predicting the success of a film. The Motion Picture Association of America assigns one of five rating categories to movies: G, PG, PG-13, R, and NC-17.[44] These ratings reflect the level of sexual content, violence, and adult language and situations contained in movies before their theatrical release. While earlier studies have shown that MPAA rating has no significance in predicting box-office returns [Litman and Kohl 1989], it makes rational sense that the non-restrictively rated movies appeal to larger audiences when all other influencing factors are kept equal. As a result, the higher the rating and thus restriction on a movie, the lower the returns expected at the box office. Despite this line of thought, a condition of ceteris paribus can rarely be assumed in a creative industry like film. While a rating maybe restrictive, it is important to consider the film in its entirety. In order to have a broad look at the factors that can potentially influence a film, it is important to consider a mechanism that restricts consumption on such a broad scale. I represent each of these ratings gathered from the IMDB as a binary variable.

Regarding *Director's previous financial success*, to my knowledge, no study has demonstrated an impact of a director's star power on the financial success of a movie. Because a director exerts the most creative controls over the final product of a film, it follows that he can exert the greatest impact when it comes to the public's acceptance of the film as a product. It makes intuitive sense that if a director is a proven race horse, i.e., has been financially successful in the past, he should be more likely to do the same in the future. One of the issues with previous studies was their lack of a quantifiable variable to indicate a director's star power. Most relied on references such as *People* magazine and other subjective approaches to estimate the public's approval of a given director.[45] I will be using the average box-office return for a director prior to the release of the motion picture. This data will be drawn from a data set comprised of 422 directors drawn from The Numbers (the-numbers.com) and Box Office Mojo (boxofficemojo.com).

*Actors' previous financial success* refers to how often stars have been shown to have an impact on movie success. However, it is still a greatly contested issue [De Vany and Walls (1999)]. Again, an issue that has faced researchers in the past is the lack of a quantifiable means for judging a star's box-office power. As with director's power, I am going to use previous financial success as a metric to determine whether tried and true stars are still likely to a have positive effect on their next movies. For actors, however, the variable for star power will be an average of each major actor's average box-office returns prior to the release of the movie. Actors performing in cameos and bit parts are not considered to be participants in a movie. Major actors were determined by Box-Office Mojo's definition of "Players" in a movie and the data for average box-office return is drawn from a list of 271 “Player” actors are also drawn from The Numbers (the-numbers.com) and Box Office Mojo (boxofficemojo.com).

*Franchise* suggests that whether or not the movie is within a series can make a dramatic difference on box-office returns [Ravid (1999)]. Sequels and movies in a series essentially have a ready-made audience that does not require nearly as much advertising to reach.  The first movies in a series in fact serve as advertising for later films. Intuitively, it makes sense that a piece of intellectual property within a film, concept, storyline, or characters that has already proven successful will be more likely to be successful.  This is represented by a simple binary variable indicating whether or not the film was produced as a part of a franchise.

Several factors that have been used in the past that I will not be using include technical effects and awards.  Technical effects, while possibly a creative element, are considered directly within the production budget and therefore are an unnecessary factor. There are also significant issues with quantifying the level of technical effects without considering it as the costs already included in the production budget.  Subjective measures lack the ability to translate to a universal method, which is what studios require after all if they are seeking to mitigate their risk.  Awards also are problematic because they are only given out after production and release.  Thus, even if they are an important factor, they cannot help studios predict how well a movie will do before production begins.  As a result, awards and even public perception of the final product are not useful for a model which serves to mitigate the risk inherent within studio production.

The second category of variables examined will be the **Studio Estimates** variables, including Opening Date, Production Budget, Advertising Budget, Number of

Opening Theaters, and Competition at Release. These variables on the whole represent estimates of practical business aspects of the motion picture industry. If Creative Aspects represent a motion picture as art (albeit quantified and classified art) then Studio Estimates represent a motion picture as a product.

*Opening Date* has been shown repeatedly to be an important indicator in box-office returns. The specific month in which a film is released has been analyzed previously and in a number of studies has been shown to be significant in predicting box-office returns [Krider and Weingberg (1998) and Sharda and Delen (2002)]. For instance, movie consumers have been shown to be more likely to patronize a theater during vacation times, such as Thanksgiving, Christmas holidays, and the summer [Litman (1983)]. As a result, the date on which a film is released can have a huge effect upon its box-office reception. Using binary encoding, the data was classified for a given month of release.[1]

*Production Budget*, or an expert and industry estimate of the production budget such as those of Box-Office Mojo and The Numbers, can potentially be a good indicator of box-office returns. Because a studio's information remains private, the actual production budget numbers often are unavailable. Thus, the industry participants are forced to make estimates based on other studios' budgets. Studios presumably possess the same expertise as that found at Box-Office Mojo and The Numbers. As a result, they should be able to make production estimates of the same caliber before shooting begins. It should be noted however, that the motion picture industry is notorious for going over

[1] This means if the film was released in December, then it would be coded a 1 for the December variable and 0 for the other months.

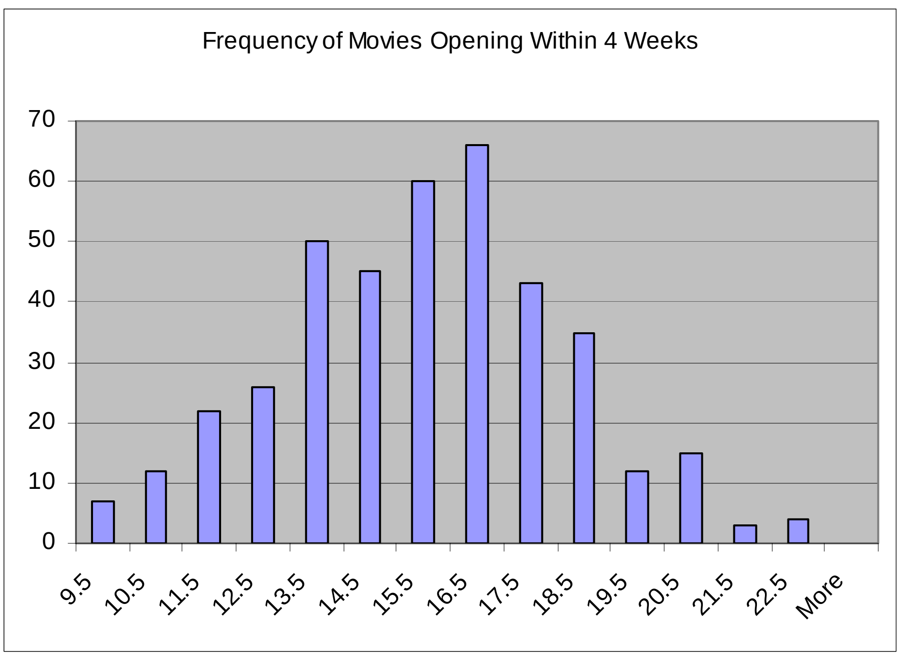
budget and having issues with their estimates. That being said, the estimates still should be a good indicator for box-office returns, as they include accurate information about actors' salaries, the creative team's salaries, an indication of technical effects, and so on. These production budgets essentially provide a financial indication of the quality of the final film. You get what you pay for. An issue with causality will inevitably remain, as studios are likely to spend more money on movies that they believe will make more.

*Advertising Budget* based on expert estimates can be a good indication of box-office return. Intuitively, since studios spend multimillions to grab customers' attention in an effort to create buzz about a certain movie, advertising is clearly thought to cause and influence box-office returns. Information about the level of a film's advertising budget thus would logically be indicative of the box-office returns. Like production budgets, causality is an issue with advertising budgets because studios are likely to spend more money attracting customers when they believe that the film is likely to make more.

*Number of Opening Theaters* represents the total opportunity for consumers to view a film and box-office returns should be highly correlated to this number. Considering motion picture box-office returns as a practical business, the more individual theaters in which a film is released, the greater the increase in possibility for sales. Due to the mechanics of the film industry, some films are released in a small sample of theaters in an attempt to build up buzz and word-of-mouth advertising before they progress to wide release. As a result, the data set contains a disparity between films released in 8 theaters and films released in 3,000 theaters. This indicates the strength of the data set, which includes a broad range of films.

*Competition at Release* can make or break a film. Movies are not released in a vacuum and often face stiff competition at the box office. Each individual movie competes for a portion of the same amalgamation of entertainment funds and competes against other movies released within the same viewing window. A movie's box-office returns can therefore be expected to be negatively correlated with the competition existing in the marketplace of is released into. These forces have been shown to negatively influence success in the past [Litman and Kohl (1989)]. Some studies have treated this almost as a subjective measure, but I have established a quantifiable manner of classification. Looking at *Figure 3.2.1*, a histogram of Competitive Release, one sees that the majority of movies are released within four weeks of the releases of thirteen – eighteen other movies, non-inclusive. Accordingly, I have decided to use a binary variable to represent movies released in a heavily competitive environment.

*Figure 3.2.1* Histogram of Competitive Release



The factors I am excluding have been primarily classified in the past as **Non-Studio Factors,** which are logically factors over which the studios have no direct influence. These include reviews, perceived quality, and initial box-office information. My primary reason for excluding them from my model of box-office return is that these are factors are only knowable after production and public release of a film. As a result, they are not useful to a model seeking to mitigate investment risk of a studio. These factors could be useful in building an optimal advertising budget or tailoring release to optimize profit, on the other hand, provided one knows the motion picture studio's initial investment in production and release.

There are also issues of causality with reviews specifically. Might a reviewer be affected by the public's love for a movie and therefore write positive reviews for movies that will already do well? Or might the public base their beliefs on a reviewer's opinion and thus the reviewer can cause a movie to become successful? The subjective nature and difficulty in properly quantifying perceived quality make it a much less valuable candidate for inclusion in a model forecasting box-office returns.

### 3.3: Hypotheses

Several intuitive beliefs are apparent within the mechanisms of the motion picture industry and I would like to test some of these intuitive beliefs. The previous research and literature has found that stars have a positive impact on the financial returns of a film at the box-office. I will seek to demonstrate this conclusion again and perhaps address some of the questions that remain open for some studies. Brand-name actors are paid an amazingly large sum up front just for production of a movie and they often have the potential for profit sharing stipulated in their contract on the back end. Clearly, the major studios feel that these stars are worth the investment. Despite their demonstrated economic beliefs, work has been done to show that only the top twenty actors and actresses in Hollywood serve to mitigate risk and guarantee box-office returns. The hypothesis to be tested is that stars with proven financial success (the metric being used to define star power in a different way than other researchers) will be positively correlated (beneficial) and a statistically significant explanatory variable of domestic box-office returns.

Hollywood is often criticized for its lack of creativity and reliance on proven scripts or franchises in production. However, it follows logically and has been demonstrated that if a film worked the first time, it is likely that it will work and be profitable a second time. I will hypothesize that being a part of a franchise causes a film to have a large audience already waiting upon its release and thus should be beneficial to its box-office returns.

Also, on the practical side of the motion picture industry, the number of theaters in which a film opens directly determines and can limit the number of tickets a consumer can purchase. As a result, this factor should be beneficial, significant, and highly predictive of box-office returns.

Ratings serve as a measure to restrict possible viewers and as a result an R rating should perform worse than a PG-13 or PG movie that can appeal to, and be consumed by a larger audience of both adults and children. I will test whether in fact ratings are significant in limiting box-office returns, and thus indicate their ability to limit consumption of motion pictures. The hypothesis that I will test is that the "R" classification is in fact detrimental.

Competition intuitively should be detrimental to a film's release, but looking at the histogram of competitive release (*Figure 3.2.1*), most films are released within a similar competitive window. As a result, I doubt that competition is truly significant in predicting box-office returns and will seek to demonstrate this hypothesis.

The date of release for a film has been shown in the past to be very influential and significant in predicting box-office returns and I would like to test whether this is in fact the case. The previous literature has demonstrated that a summer release is beneficial and a significant predictor of box-office returns. My analysis will attempt to replicate these previous results.

Logically, the film's production budget is comprised of three main factors: cost of the creative cast, cost of development, and cost of various rights, such as filming licenses. As a result, a film's star and director power will be positively influential on its production budget. This power will be represented by these creative elements' previous financial successes in the model developed. Certain genres also necessitate larger production budgets. For example, Animated films require a large amount of technical effects and technology costs and therefore the cost should be significantly higher to produce. Financially successful creative actors should represent a greater amount of talent, at least in producing box-office returns, and thus the cost of the creative element should lead to a greater increase in the budget.

Advertising expenses should be positively influenced by production budget, star and director power, and should be negatively impacted by whether or not the film is part of a franchise. The final factor important to advertising expenses should be the amount of competition; however, it is unclear to me whether the greater competition would cause an increase in advertising as an attempt to capture a greater portion of the public's interest. Or the greater competition could reflect that the distributor choose this weekend specifically, either acknowledging the increased competition and in essence letting those customer dollars go elsewhere, or believing that they will capture that weekend with less advertising due to the nature and appeal of each movie. Regardless of the effect of competition, it logically would be a significant issue in the decision of creating and allocating advertising budget.

## 3.4: Method and Analysis

### 3.4.1: Summary Statistics

Using the data set compiled from the three websites of the top 100 grossing films for 2002, 2003, 2004, and 2005, I am first attempting to replicate some of the findings reported by previous studies, as well as a number of other predictions enumerated in the hypothesis sections. I will begin with analysis of the summary statistics for the data set.

*Table 3.4.1.1*

**Summary Statistics for Non-Binary Variables**

| | Min: | 25% Quintile | 50% Quintile | 75% Quintile | Max: | Mean: | Std Dev. |
|---|---|---|---|---|---|---|---|
| Total Gross | 22918387 | 36969082 | 56119030 | 93948746 | 441226247 | 80201874 | 69071793 |
| Opening Theaters | 2 | 2442 | 2827 | 3208 | 4163 | 2595 | 989 |
| Advertising | 5000000 | 20000000 | 25000000 | 35000000 | 75000000 | 28934263 | 11863457 |
| Director | 0 | 0 | 558873 | 51270894 | 355251351 | 31823786 | 49570789 |
| Actor | 0 | 21268600 | 35022441 | 50144823 | 306471167 | 37033610 | 30389911 |
| Budget | 400000 | 25000000 | 45000000 | 75000000 | 207000000 | 54322500 | 37339205 |

Note: The discrepancy between the median and mean of the Total Gross as well as the same disparity with the Director financial values indicates that the data is skew right. Also, note the skew left Opening Theaters variable

Table 3.4.1.1 contains summary statistics for the Non-Binary Variables. A preliminary examination of the data set demonstrates that the data set is robust, ranging from incredibly limited opening release (two theaters) to a giant national release (over 4,000 theaters). Despite this indication of a good sample set, this data set will inevitably contain some biases. The first hint at bias comes from the median and the mean of the opening theaters which both indicate that the average film is one experiencing wide release. Despite this high average, there is some significant variability with a standard deviation of roughly 1,000. Also, the Opening Theaters are clearly skew left, indicating that despite a few huge opening releases, the smaller, more limited or platform releases are more frequent.

Looking next to the Total Gross, again one sees a huge variability with a standard deviation of nearly $70 million and an average of $80 million. This is a reflection and a result of the fact that the data set by necessity had to be of the most successful movies in a given year. One can also already see that the larger returns may be dramatically larger than the lower values with the discrepancy between the median of $56 million and a mean of $80 million. This discrepancy makes intuitive sense when considering the nature of the industry. Motion Pictures are characterized by the tremendous blockbusters that make it worthwhile to be involved in the industry in the first place. This will be something that will have to be taken into account when one considers the results of the model built to predict box-office returns.

Looking at the Director data, a similar skew right bias is demonstrated by the discrepancy between the median and mean. It is important to understand though that a large portion of the Director variables are zero, which is largely due to the fact that a large portion (over 25%) of directors are unproven financially prior to the production and release of a film. The skewed nature of the data is largely attributable to a small number of dramatically successful directors. It would seem logical to expect similar characteristics within the Actors variable, however, the same skew right is not demonstrated and there are many fewer zeros in the data set. This could be an indication that studios do not rely on unproven actors to star in their films as frequently. Rather than using unproven but potentially huge undiscovered stars, studios tend to rely on a small set of demonstratively strong actors when building their casts. Also, note that the average of the cast's previous financial success (represented by the Actor variable) is roughly $7 million more than the average of the Director data.

It is also interesting to note that while a film's budget ranges from $.4 million to $207 million, the advertising budget minimum is dramatically larger. It seems that in order to reach the national market, a film requires a set amount of advertising. This may be a result of the nature of this data set being based on the top 100 films of a given year. In order to achieve success, a certain benchmark may be required when it comes to advertising expenditures. Also, note that the Advertising median is much smaller than the Budget median and the standard deviation is much smaller, indicating a much tighter data set around this smaller median. It should be noted that due to the private, competitive nature of the industry, many data points for the Advertising budget are missing, yielding a data set of only 251 movies. The films' budgets (while still considered precious and hoarded) are much more available and the data set missed only 20 films' production budget values.

*Table 3.4.1.2*

**Box-Office Breakdown ($MM): Release, Genre, Rating, Competition and Franchise**

| | Min: | 25% Quintile | 50% Quintile | 75% Quintile | Max: | Mean: | Std Dev. |
|---|---|---|---|---|---|---|---|
| *Month of Release* | | | | | | | |
| December | 23.5 | 39.7 | 54.0 | 103.7 | 377.0 | 87.9 | 79.8 |
| November | 27.0 | 36.0 | 60.1 | 133.5 | 286.4 | 91.0 | 70.8 |
| October | 25.2 | 32.8 | 53.8 | 79.3 | 160.9 | 59.3 | 32.6 |
| September | 25.5 | 35.7 | 46.2 | 54.1 | 127.2 | 50.9 | 23.0 |
| August | 26.2 | 37.9 | 58.2 | 95.2 | 228.0 | 71.5 | 43.8 |
| July | 25.5 | 35.8 | 65.7 | 138.6 | 305.4 | 92.6 | 69.2 |
| June | 24.0 | 44.9 | 76.4 | 128.4 | 373.6 | 101.2 | 74.8 |
| May | 24.5 | 45.1 | 104.3 | 204.3 | 441.2 | 143.4 | 123.8 |
| April | 22.9 | 34.3 | 51.6 | 67.7 | 241.4 | 60.0 | 40.9 |
| March | 23.6 | 34.5 | 43.7 | 75.9 | 176.4 | 58.1 | 34.3 |
| February | 25.2 | 34.7 | 48.4 | 75.6 | 370.3 | 73.0 | 71.4 |
| January | 25.5 | 40.9 | 52.0 | 60.2 | 88.1 | 53.2 | 17.2 |
| *MPAA Rating* | | | | | | | |
| G | 25.5 | 52.3 | 75.6 | 114.4 | 339.7 | 96.8 | 78.2 |
| PG | 22.9 | 38.8 | 52.3 | 107.3 | 441.2 | 87.1 | 78.0 |
| PG-13 | 23.5 | 38.1 | 59.5 | 106.8 | 403.7 | 86.0 | 72.6 |
| R | 24.0 | 33.3 | 47.4 | 74.5 | 370.3 | 62.4 | 49.0 |
| *Genre* | | | | | | | |
| Action | 23.6 | 39.2 | 62.6 | 120.6 | 403.7 | 94.1 | 82.0 |
| Adventure | 22.9 | 37.9 | 65.3 | 128.7 | 441.2 | 103.2 | 93.7 |
| Animated | 25.5 | 42.2 | 79.3 | 157.1 | 441.2 | 114.5 | 103.2 |
| Biography | 25.9 | 33.4 | 34.5 | 68.5 | 107.0 | 53.0 | 29.8 |
| Comedy | 22.9 | 37.7 | 56.7 | 94.9 | 441.2 | 78.6 | 63.1 |
| Crime | 24.5 | 34.5 | 52.2 | 80.0 | 205.3 | 63.6 | 38.1 |
| Drama | 23.5 | 34.8 | 50.2 | 74.9 | 370.3 | 65.2 | 48.7 |
| Family | 22.9 | 40.8 | 60.9 | 102.1 | 441.2 | 89.1 | 78.8 |
| Fantasy | 22.9 | 51.4 | 84.0 | 194.2 | 441.2 | 131.7 | 111.2 |
| History | 23.5 | 36.7 | 49.6 | 61.9 | 78.1 | 49.9 | 19.5 |
| Mystery | 23.6 | 40.2 | 59.7 | 88.6 | 286.4 | 78.6 | 58.4 |
| Romance | 22.9 | 37.2 | 52.0 | 88.1 | 380.3 | 72.1 | 58.8 |
| Science Fiction | 25.5 | 35.6 | 57.2 | 129.2 | 403.7 | 97.2 | 98.5 |
| Sport | 32.5 | 40.4 | 52.3 | 64.4 | 81.2 | 52.9 | 15.4 |
| Thriller | 23.5 | 35.3 | 51.6 | 77.5 | 373.6 | 71.8 | 58.2 |
| War | 23.5 | 42.5 | 64.3 | 95.2 | 133.4 | 70.8 | 35.1 |
| High Competition | 22.9 | 34.4 | 52.6 | 85.4 | 441.2 | 80.2 | 76.8 |
| Franchise | 25.5 | 48.5 | 85.8 | 160.9 | 441.2 | 126.0 | 104.5 |

*Table 3.4.1.2* demonstrates the box-office revenues break down statistics for the binary variables (in 2002 – 2005 constant millions of dollars). In examining this table, first looking at month of release, where the first striking statistic is the May box-office total which at every quintile is higher than any other month. May's average (both mean and median) is also dramatically larger than any other month. Interestingly though the standard deviation for May is also dramatically larger, indicating a more spread out distribution of box-office returns for films released in May. Movies released within September and March seem to fare worse than others with the two lowest averages (again both mean and median). It may be that May's peak and September and March's trough are reflections of the film industry's oft touted seasonality.

Appendix A shows the frequency that the binary variables are seen within the data set. Referencing the frequencies of occurrence within the data set may also support the possibility of seasonality. For nearly every variable classification there was a discrepancy between the mean and median that indicated the box-office returns were skew right. This is merely a reflection that the distribution of box-office returns, regardless of classification, is skew right.

Next looking at MPAA ratings, it is interesting to note that G-rated films do have dramatically higher returns and R-rated films lower returns, which follows my hypothesis. Looking at Appendix A, G-rated films account for less than 4% of all films within the data set and so the summary statistics are biased by a few blockbuster releases.

In an examination of Genre, it is evident that Animated and Fantasy films seem to perform better than their counterparts. But both of their means are dramatically larger than their medians, indicating a large skew right in the data set. Also, it is interesting to note that Animated is incredibly uncommon (less than 1%) and Fantasy is relatively uncommon (roughly 10% of all movies See Appendix A). Comedy, Action, Adventure, and Thriller are much more common, with Comedy representing nearly 50% of all films. When examining genres, it is important to remember that an individual movie can be classified as being part of multiple genres (such as *King Kong* which is classified as Action, Adventure, Drama, Fantasy, Science Fiction, and Thriller).

3.4.2: Box Office Linear Regression Model

EQ. 3.1:
$$\begin{aligned} &BORETURN_i = \beta_0 + \beta_1 \cdot MAY_i + \beta_2 \cdot FRANCHISE_i + \\ &\beta_3 \cdot ADVERTISING_i + \beta_4 \cdot ANIMATED_i + \beta_5 \cdot FAMILY_i + \\ &\beta_6 \cdot FANTASY_i + \beta_7 \cdot THRILLER_i + \beta_8 \cdot DIRECTOR_i + \beta_9 \cdot ACTOR_i + \varepsilon_i \end{aligned}$$

This regression model and equation is a result of removing all statistically insignificant variables that were originally considered in an effort to make a more accurate model. EQ. 3.1 refers to the multivariate linear regression model developed and used to forecast box-office returns of motion pictures. Where $B_0$ is the regression constant, MAY is a variable indicating if a film is released in May, FRANCHISE indicates the film is following in a series of other related films, ADVERTISING represents the films advertising budget, ANIMATED, FAMILY, FANTASY, and THRILLER demonstrate that the film’s genre classification includes but is not limited to Animated, Family, Fantasy and Thriller genres, and finally DIRECTOR and ACTOR, which are a representation of the previous financial successes of the creative elements (cast and director) creating the film. Specifically, DIRECTOR is calculated as the average gross of all of the director’s previous films. ACTOR is the average of the leading actors' average gross of all previous films they starred in. The parameter $\varepsilon_i$ is assumed to be a normally distributed variable indicating noise or error within the data.

*Table 3.4.2.1*

**Total Domestic Gross**
**Regression Coefficients**

| | Value | Std. Error | t value | Pr(>\|t\|) |
|---|---|---|---|---|
| (Intercept) | -6820106.02 | 9643033 | -0.7073 | 0.4801 |
| May | 52724368.7 | 12398088 | 4.2526 | 0 |
| Franchise | 30776209.7 | 8980972 | 3.4268 | 0.0007 |
| Advertising | 2.5159 | 0.3391 | 7.4188 | 0 |
| Animated | 29689921.2* | 15940495 | 1.8625 | 0.0637 |
| Family | -31681670.9 | 11019391 | -2.8751 | 0.0044 |
| Fantasy | 29454820.1 | 10237459 | 2.8772 | 0.0044 |
| Thriller | -18776127.4 | 7402966 | -2.5363 | 0.0118 |
| Director | 0.2864 | 0.0764 | 3.7496 | 0.0002 |
| Actor | 0.2826 | 0.1121 | 2.5205 | 0.0124 |

R-Squared - 0.5354

Note: A * indicates that the coefficient is only significant at a 90% confidence level. That while the intercept may be negative, it is due to the fact that films are never released without $5 million in Advertising and the majority have a large indicator of director or actor financial success. Also, note that this intercept is not significantly nonzero, indicating that there is no base value for box-office returns and that the variables used in the model combine to explain any systemic baseline values that may be present within the data set.

A common standard to judge the quality of a regression is the $R^2$ statistic. The $R^2$ statistic gives the proportion of the variance of the dependent variable explained by the regression of the explanatory variables. $R^2$ is mathematically arrived at as follows:

$$R^2 = 1 - \frac{SSE}{SST}$$

Where $SSE = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$ and $SST = \sum_{i=1}^{n} (y_i - \bar{y})^2$

This model's $R^2$ indicates that the regression using those nine variables only explains 53.54% of the variance in the dependent variable or domestic box-office returns for films. While this looks like a poor fit, in fact by looking at the previous literature, it is apparent that this performs as well as nearly every other model using only pre-release information. The Adjusted $R^2$ for this model is .5247 and while this is a relatively low value, this is due to the unpredictable nature of the motion picture industry. Unfortunately, it remains very hard to quantify qualities that lead one movie to become a blockbuster and another to flop. This makes it supremely hard to model box-office returns. As a result, and as an examination of the previous literature demonstrates, this model, while seemingly not a good fit, is as good as can be expected of box-office returns.

Most interesting and surprising about the model is the huge coefficient that a film released in May can expect ($52 million). This means that a film is expected to earn $52 million more if it is released in May. There are several possible explanations for this result. It may be that this is truly a reflection of the seasonality that is believed to exist in the motion picture industry. May brings the start of summer and May films accordingly receive higher revenues as the public's interest in entertainment peaks. Looking back through the data, some of the largest grossing films in the data set were in fact released during Memorial Day weekend which would logically be prime placement for a new summer blockbuster. Combining the holiday release and the start of a summer run, it seems to make sense that a film released in May could expect to have a much better showing. It is unclear if a May release is the cause of a larger release or if studios expecting a good release for a film seek to dominate the summer market and release it in May. This would have the effect of artificially inflating the returns for films released in May and could be another cause for this very significant (with a P value of essentially 0) coefficient.

Being part of a Franchise is also very beneficial for a film. This model shows that the franchise is worth nearly $31 million. This is the hypothesized result and is logical since a franchise builds upon the advertising of another film. Within Genre only four classifications were significant: Animated and Fantasy, which were both beneficial, and Family and Thriller, which were both detrimental to box-office returns. Animated was only significant on a 90% confidence interval. It is interesting to see though the interplay between Animated, Fantasy and Family genres, because often an Animated film is also classified as a Family or Fantasy film. As a result, a film classified both as Family and Animated has genres that essentially cancel each other out in terms of box-office returns. In addition, Thriller as a separate genre did not perform as well. This could easily be explained by the cyclical nature of the public's tastes in film. Horror and Thriller typically do much better in down economic markets and Comedy and lighter movies do better in bull markets.[46]

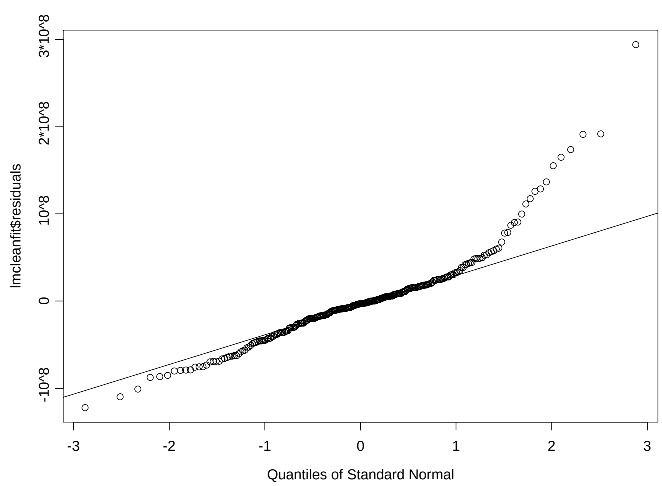
Advertising budget is very significant in predicting box-office returns and is an interesting statistic to examine. For every dollar spent on advertising, the studios will get a 2.5 multiple back at the box office. Logically, this would mean studios would spend an infinite sum on advertising. Clearly, there must be some point of diminishing marginal returns in this practical world. In an attempt to demonstrate diminishing marginal returns of advertising dollars, a regression of box-office returns against a logarithm of the advertising term yielded unsatisfactory results. It confirmed the squared term to be insignificant. This is likely due to the limited amount of advertising information. The 2.5 multiplier for advertising budget is likely to be only applicable to the range of $20-$40 million in which the majority of the advertising budgets within the data set lay.

Finally, analyzing the Director and Actor variables, one finds that a financially successful creative cast and crew are beneficial to box-office returns of films. This indicates that star directors and actors are more likely to perform better in the box office, if their past films have also done so. The coefficients for the two are nearly identical which would mean that the previous financial success of the director is as important as that of the cast he is directing. For every additional average dollar of success that a cast or director brings to a film only 28 cents more is generated at the box office. It begs the question of whether actors and directors are worth the exorbitant sums that studios pay them if they expect to be compensated for every dollar of financial success brought to the film. The complete analysis of the return on investment for actors is beyond the scope of this thesis, although I am sure it would prove very interesting to motion picture studios. A summary analysis based on rudimentary simplifications will follow later.

Because of the nature of the data set for actors' and directors' financial data (containing so many zeros representing unproven casts or directors), an analysis to determine if unproven directors and actors are truly worth nothing was interesting. A regression of a new variable created to represent "New" actors and director (i.e. Actor or Directors with no average box-office return for movies previously created and released) against box-office return within my box-office model demonstrated that these new manufactured variables were statistically insignificant. This demonstrates that "New" casts or directors are expected to be truly worth nothing at the box office.

Several variables that were hypothesized to be significant, were subsequently demonstrated as immaterial. The production budget in the past has been found to be a significant predictor of box-office returns and was not one here. This is interesting but actually an understandable result. In subsequent analysis, a film's production budget is shown to be an important predictor of advertising budget. In performing various regressions production budget was significant when advertising budget was not included. This indicates that all information contained within the production budget is also contained within the advertising budget variable, thereby making the production budget insignificant in our final regression. Another variable that was hypothesized to be significant is also explained by the advertising budget. Recall in the discussion of distributors, the primary issue they faced was determining the scope of release. This encompasses two factors: the first being the number of theaters showing an opening film and the second the advertising budget allotted to the marketing of the film. As a result, these two decisions are closely tied and highly correlated and again we find that the information contained within the variable for opening theaters and thus opening theaters are insignificant.

*Figure 3.4.2.1*




Another method for determining how good a fit the model represents is to look at the residuals of the Regression Model as compared with the quantiles of the normal distribution. One sees that the right tail is heavier and this indicates the regression is not as accurate for films with higher gross returns at the box office. This reflects what was demonstrated on the cursory analysis of the summary statistics, that box-office returns are skew right. This also fits with an understanding of the nature of the motion picture industry, as it is characterized by blockbusters that make dramatically more money than their competititors.

The final analysis with regards to the strength of a linear regression box-office model preformed was prediction. Breaking the dataset into 350 training movies and 50 films tested on the box-office model, predictions resulted in an average absolute error of nearly $30 million and a mean percent error of 60%. The actual total box-office gross fell within the 95% confidence interval predicted by the box-office model only 33% of the time. This model is demonstrably poor. Appendix C contains examples of the prediction as well as summary statistics.

3.4.3: Production Budget Regression

EQ 3.2:
$$\begin{aligned} PRODBUDGET_i &= \beta_0 + \beta_1 \cdot FRANCHISE_i + \beta_2 \cdot DIRECTOR_i + \beta_3 \cdot ACTOR_i + \\ &\beta_4 \cdot ADVENTURE_i + \beta_5 \cdot ANIMATED_i + \beta_6 \cdot FANTASY_i + \beta_7 \cdot SCIFI_i + \beta_8 \cdot WAR_i + \varepsilon_i \end{aligned}$$

Again, this regression model and equation is a result of removing all statistically insignificant variables that were originally considered in an effort to make a more accurate model. EQ. 3.2 refers to the multivariate linear regression model developed and used to forecast the production budget of motion pictures. Understanding what factors go into this forecast will allow a better understanding about what studios are thinking in terms of their development of films. Within EQ. 3.2 $B_0$ is the regression coefficient, FRANCHISE indicates the film is following in a series of other related films, ADVENTURE, ANIMATED, FANTASY, SCIFI, and WAR demonstrate that the film's genre classification includes, but is not limited to, Adventure, Animated, Fantasy, Science Fiction, and War, and finally DIRECTOR and ACTOR, which are representations of the previous financial success of the creative elements (cast and director) creating the film. Remember, DIRECTOR is calculated as the average gross of all of the director's previous films. ACTOR is the average of the leading actors' average gross of all previous films in which they starred. Again, the parameter $\varepsilon_i$ is assumed to be a normally distributed variable indicating noise or error within the data.

*Table 3.4.3.1*

**Production Budget**
**Regression Coefficients**

| | Value | Std. Error | t value | Pr(>\|t\|) |
|---|---|---|---|---|
| (Intercept) | 22124125.4 | 2644674 | 8.3655 | 0 |
| Franchise | 8216238.97 | 3992884 | 2.0577 | 0.0403 |
| Director | 0.1856 | 0.0309 | 6.0058 | 0 |
| Actor | 0.208 | 0.0493 | 4.2168 | 0 |
| Action | 24272105.1 | 3331947 | 7.2847 | 0 |
| Adventure | 11021256.5 | 3583164 | 3.0758 | 0.0023 |
| Animated | 14422633.2 | 6376942 | 2.2617 | 0.0243 |
| Fantasy | 17692120.6 | 4364243 | 4.0539 | 0.0001 |
| Science Fiction | 12196880.1 | 4869752 | 2.5046 | 0.0127 |
| War | 37476697.9 | 8047932 | 4.6567 | 0 |

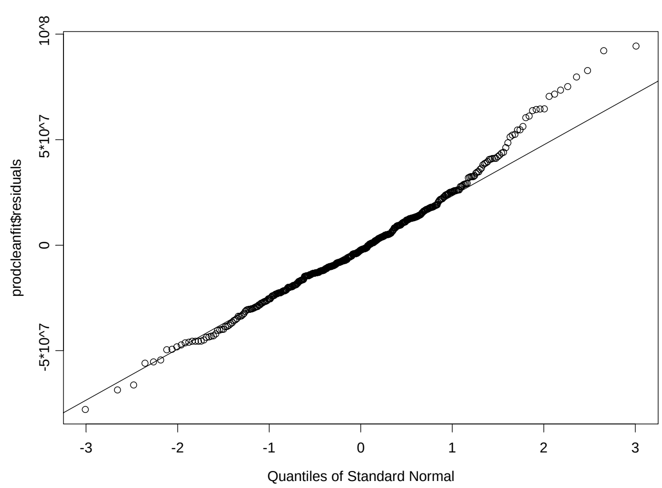
R-Squared - 0.5558

First, we can see that the intercept is significant, indicating that films from the data set are starting from a baseline production budget of roughly $22 million. This is likely due to the fact that the data set is comprised of the top 100 grossing films per year. It is a reflection that in order to be included in this data set, the film required a certain amount of budget which then presumably translates into a certain level of quality and marketability. Looking at the creative crews' influence on the budget indicates an interesting finding. For every average dollar the film's cast or director has made in the past, the film's budget will cost $.18 or $.2 (respectively) more. The analysis of box-office return demonstrated that for every average dollar the film could expect $.28 more return at the box office. Thus, depending on how the gross box-office revenue is shared, it becomes clear whether or not the more expensive talent is marginally beneficial.

Based on a summary analysis, the typical metric for profit excluding ancillary revenues is ½ * Box-office Returns – Budget. This has been shown to be accurate and relevant in a number of studies and will be the metric I use to estimate the profit a film generates for its production studio.[47] Thus, increasing the average gross of a cast by one more dollar of previous financial success will yield ½ * $.28 - $.19 (average of director and cast) = -$.05 in marginal profit. This would indicate that rational studios would not ever upgrade a cast if all other considerations were equal. Understanding that this is a very basic summary analysis, it still reflects the hypothesis that perhaps high-end talent is not worth the price paid.

Examining the effects of genre on production budget, it makes intuitive sense that genres which require more technical effects or large scale computer graphics would greatly increase the cost of production for a film. The results one finds are fairly intuitive. All six of the genres found to increase a film's production budget (Action, Adventure, Animated, Fantasy, Science Fiction, and War) would be logically expected to require a greater level of technical effects due to explosions and large fight scenes in Action, Adventure, and War films or large CGI-generated imaginary worlds in Animated, Fantasy, and Science Fiction films. It is interesting to note that my previous study demonstrated that Animated films could expect roughly $30 million more at the box office and this budgetary analysis indicates that they will cost roughly $14 million more to make.  Using the same summary analysis of profit ½ * $30 M - $14 M = $1 M of profit indicates that the budgetary increase would not be prohibitive to a profitable return for the release of an Animated film.  In contrast, the box-office model indicated that a Fantasy film should generate again roughly $30 million, but this budgetary consideration reflects an increase in $18 million.  Using our metric for profit generated by the change to a Fantasy release (½ * $30 M – $18 M = -$3 M), one finds that it would not be a profitable venture, all other things being equal.

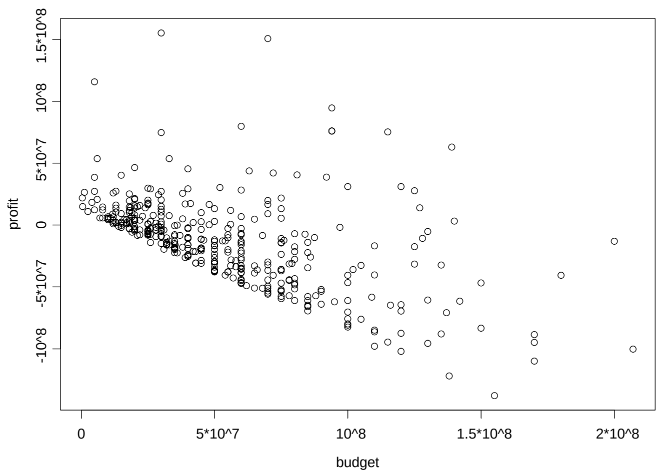
*Figure 3.4.3.1*




A similar analysis of the residuals created by this production budget model indicates the same results as was seen with box-office returns. For large budgets one sees the model does not have as strong a fit.  There is a heavy right tail in the distribution of the residuals of the model.  Once again, this reflects what was hinted at in the analysis of the summary statistics, that the production budget is skew right.  This fits with a rational understanding of the motion picture industry, which while characterized by huge blockbusters, recognize that nearly two-thirds of released films are unprofitable.  This can largely be attributed to the enormous production budgets that go into creating some of these films.

**Budget vs. Profit**

*Figure 3.4.3.2*




Again using the metric Profit= ½ Domestic Box-Office Gross – Production Budget a standard that has been shown to be a good approximation for profit in the past. Looking at *Figure 3.4.3.2* one can see a plot of profit versus budget in which the movies have predominantly negative profits. By examining its applicability, one finds that roughly two-thirds of the films within this data set are not profitable with break-even films being in the 65.5% quantile. This is an accepted figure and a standard number quoted often by industry executives and MPAA studies and confirms the metric used to approximate profit. The median profit was -$10.15 million and the mean -$13.28. This is a further demonstration that this is risky business with production budget that represent an enormous investment and a huge assumption of risk.

### 3.4.4: Advertising regression

EQ 3.3: $ADBUDGET_i = \beta_0 + \beta_1 \cdot BUDGET_i + \beta_2 \cdot FRANCHISE_i + \beta_3 \cdot ACTOR_i + \beta_4 \cdot ANIMATED_i + \varepsilon_i$

Again attempting to understand industry decisions, I examined the relationship between various factors that went into the production of the movie and their result on the studios' allocations of advertising budget. Using similar variables as the models above and parsing the model down to only variables that were significant in predicting advertising budget, the regression model became EQ 3.3.

*Table 3.4.4.1*

**Advertising Budget**
**Regression Coefficients**

| | Value | Std. Error | t value | Pr(>\|t\|) |
|---|---|---|---|---|
| (Intercept) | 14899965.3 | 919243.6 | 16.2089 | 0 |
| Budget | 0.2057 | 0.0134 | 15.3466 | 0 |
| Franchise | 3250318.43 | 1258610 | 2.5825 | 0.0104 |
| Actor | 0.0377 | 0.0156 | 2.4172 | 0.0164 |
| Animated | 5961590.07 | 1922422 | 3.1011 | 0.0022 |

R-Squared - 0.5966

Note: Both the budget and the intercept are strongly significant. The intercept is likely due to the tightness of the data set as well as the inherent advertising requirements necessary to reach a national market. The $R^2$ also indicates the best fit we have modeled yet, explaining nearly 60% of the variance of the advertising budget.

It is interesting to see that the specific month of release and amount of competition had no significance in predicting a film's advertising budget. Also, only the actor's previous financial success, but not the director's, was significant in predicting a studio's advertising budget. Interestingly, the intercept is strongly significant. This indicates some systemic threshold that a film must reach in advertising in order to fall within the top 100 box-office grosses in a year, as those were included within the data set on which the model trained.

The larger the film's production budget the more likely a studio will be to spend additional money advertising and promoting the film. It makes sense that the studio, will pour more money into a venture that they believe will be dramatically successful. While it may be a bit counterintuitive, if a film is part of a franchise, the fact will increase the amount of advertising spent on the film. This is likely a result of the very nature of franchise movies. The intellectual property has already proven financially successful. As a result, the studio is more inclined to spend advertising dollars because they believe people will come out again and see the new film.

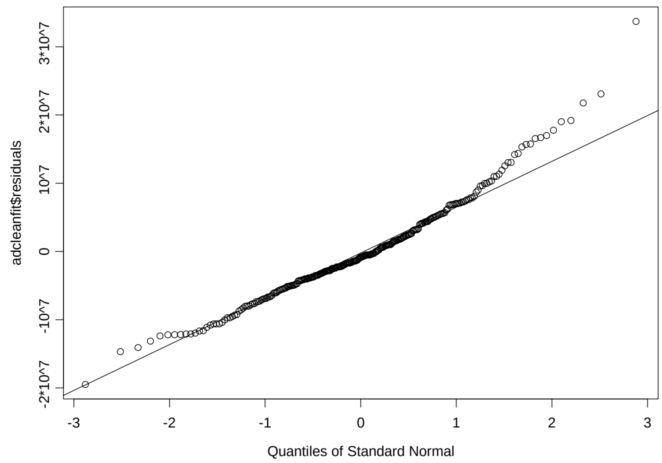
If the cast has been demonstrably financially successful in the past, it will increase the advertising budget, albeit slightly. It makes sense that if a studio is spending money for a proven cast, they want to ensure that the public knows their stars are in a film. Thus, it is rational for the studio to spend more on advertising for bigger (more financially successful) stars.

Animated is the only genre that was significant in predicting advertising budget. A film that is classified as Animated can be expected to spend $60 million more on advertising than a non-animated film. While I am unsure of the exact reasoning behind the increased advertising budget, I suspect it may be largely due to the fact that Animated films are predominantly advertising to a young demographic (children). Due to the nature of marketing films to children (who are known to be repeat consumers by nature and who bring parents along with them time and time again)[2], advertising dollars for children's movies could easily pay for themselves. Despite this argument, Family as a

[2] It is this youth repeated viewing that cause studios like Disney to have the greatest generated profit from an individual piece of intellectual property.

specific genre was not significant in this regression, so this explanation may just be a rationalization.

*Figure 3.4.4.1*

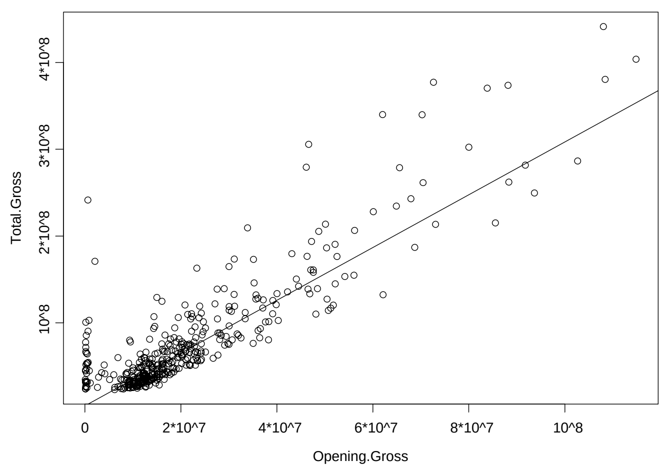



Looking at the plot of the residuals of this regression versus the quantiles of the standard normal distribution, this regression again demonstrates the strongest application for the data set. The residuals of the regression are the closest to normal if they deviate slightly when the advertising budget is higher.

### 3.4.5 Opening Gross

The final thing that will be considered will be predictions using opening weekend gross. This will be considered in an effort to demonstrate why certain previous studies demonstrate a much more accurate model using early box-office release information. While the model created in this paper only explained 53.54% of the variance and in prediction had a mean absolute error of $29 million and a mean percent error of 60%, this is largely due to the inexplicable nature of cinema audiences and the difficulty researchers have in forecasting returns from these audiences. A very simple linear regression minimizing absolute deviations yields the results found in *Figure 3.4.5.1*.

*Figure 3.4.5.1*




Simply looking at the regression indicates a relatively tight fit. Examining Table 3.4.5.1, it is clear that the $R^2$ value is dramatically higher than our other box-office model that employed 9 distinct and significant factors. The fit of the regression shows that while a large number of films open small and then go on to generate much larger returns in the long run, the fit still captures what appears to be the trend in the data.

*Table 3.4.5.1*

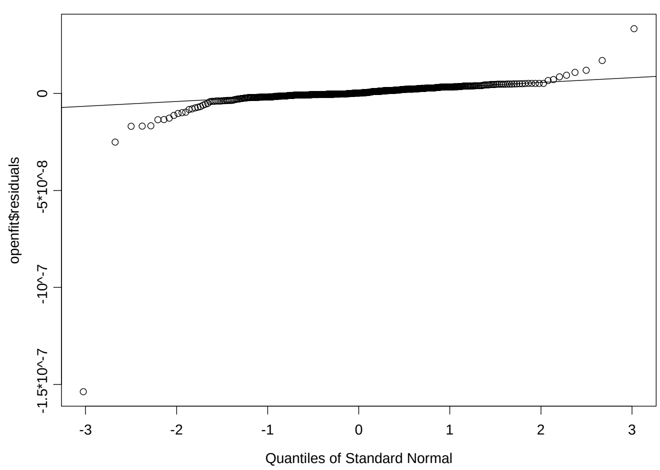
**Total Domestic Gross**
**Regression Coefficients**

| | Value | Std. Error | t value | Pr(>\|t\|) |
|---|---|---|---|---|
| (Intercept) | 10740805.6 | 2513486 | 4.2733 | 0 |
| Opening.Gross | 3.1795 | 0.0867 | 36.6741 | 0 |

R-Squared - 0.7717

According to this model a film should be expected to make approximately $10.7 million and 3.18 times its first weekend's box-office gross receipts. While there are clearly films that make less than $10 million during their entire release, this model was trained on a data set of the top 100 grossing films and is therefore biased and will predict larger returns.

*Figure 3.4.5.2*




From an analysis of the residuals plotted against the quantiles of the standard normal, it is clear that the regression residuals are nearly normally distributed. This is only a further indication of the strength of the regression and model using only the opening weekend box-office gross receipts.

### 3.5: Conclusion

The limitations of previous studies provided the reason for this study and the impetus that started this examination. Previous studies were primarily academic pursuits with no practical application. Using post-release factors misses fundamental causal relationships and the analysis yields no useful information for future film returns. While these studies largely demonstrate a more accurate classification and prediction of films already released, they offer no fundamental benefit to studios seeking a seemingly oracular forecast in an effort to mitigate the giant risk that producing a film now represents. A studio has the power to alter a few key characteristics, and my model sheds some light on the relationship between these aspects and their impact on domestic box-office returns.

So what can studios do to lessen their risk and further understand the relationship between their decisions and the public's response in the form of movie attendance and box-office return? First, the intuitive belief in established intellectual properties (in the form of franchise) being a safer path was confirmed. Studios tend to spend more on these established films and this extra spending is again demonstrably justified. Second, considering the creative cast that is all so important in the creation of the film, established and proven financially successful stars and directors are valuable assets in a film. Director's success was shown to be worth the same as the financial success of the cast. A further exploration of production budget and profitability of film demonstrated that the cost of financial success of actors was prohibitively expensive and strongly recommended that Hollywood rethink its high valuation of its iconic stars. It is also interesting to note that while directors are still excessively expensive for studios, they are not as prohibitive as actors. Directors should take notice that if the major studios can be convinced of actors' false box-office generating power, there may be more wiggle room in directors' compensation.

With regard to time of release, it is clear and demonstrated without a doubt that May is clearly an important month to win for opening release date. May presents large potential for generating box-office returns. Aside from May, no other month was important with regards to box office return. Studios and distributors need to think less about when to release and more about the manner in which they advertise. Competition did not affect box-office returns, counter-intuitively. This is likely due to the fact that the market is already saturated and competition will exist in some form or another regardless of the date of release. Number of theaters in which a film is released does not matter because implicit in the decision of the scope of release is the advertising campaign. Advertising is very beneficial, paying off $2.5 per dollar spent. But this must be placed into perspective and the data set that trained this model must be taken into account. Within the films that contained advertising information, the disparity was not very great and it represented a tight set. Thus, while the model predicts an essentially infinite return on advertising investments, the $2.5 figure is only applicable on the range contained within the dataset ($20-30 million) and diminishing marginal returns from advertising are achieved with larger advertising budgets.

A large number of films lack information on production budget and advertising budget and the values which are available are also largely suspect. Motion picture studios are notoriously secretive and their reports can be deceptively low due to intricacies in contracts, such as director and cast waving their typical salaries and opting for revenue sharing. It is often said in the industry that the "Majors (large studios) lie down and the Indies (independent studios) lie up." As a result, though, a large number of the films in my data set did not contain advertising budget information. The fundamental nature of secrecy within the motion picture business was the initial impetus for comprising the data set of the top 100 grossing films for the past four years. This composition serves to encompass the blockbusters that glorify Hollywood, but fails to incorporate the box-office bombs that are all too often the outcome for the majority of films. As a result, these forecasts are going to overemphasize greater box-office returns.

Analysis with regards to understanding targeted and implied audiences based on genre demonstrated a few interesting results and leaves some questions for studios. Animated films generate larger gross receipts and are profitable, but Fantasy films, while generating larger gross receipts, have increased their production budgets to the point where they have become unprofitable. Thrillers are not beneficial but this may be due to society's cyclical interests.

The motivation for the cursory analysis of the relationship between opening box office returns and total domestic box-office gross was an attempt to understand the weakness of the box-office model and demonstrating why other studies produced dramatically more accurate box-office prediction models. It remains true though that the work that has been done with similar restrictions has been largely in the same class of accuracy. What if there was a mechanism that, without relying on post-production information, could still generate more accurate predictions about box-office returns? If there exists such a Delphic Oracle, the major motion picture studios should take notice.

# Chapter Four:

## Analysis of the Hollywood Stock Exchange as a Forecast Vehicle

### 4.1: A Discussion of Hollywood Stock Exchange Data

As the discussion of the mechanics of the Hollywood Stock Exchange explained above, a MovieStock begins its trading life with an Initial Public Offering (IPO). Then it is the magic of the Virtual Specialist, the mechanism that governs the MovieStock price (the "Special Sauce" of the HSX), interacting with the market participants that determine the price of a MovieStock. Before the first weekend of a film's wide release the price of the MovieStock is halted to be readjusted after its performance during the weekend. The readjust price equals the box-office receipts multiplied by a 2.8 multiplier (this multiplier is designed to replicate the box-office revenues over the four weeks of trading while the movie is in release).[3] Before the halt and readjust, the trading and the price of a given MovieStock reflect an estimate of 2.8 times the box-office return of the film's first weekend in release. The MovieStock then resumes trading for the next four weeks, after which time a MovieStock is delisted and any market participant holding a MovieStock receives the amount the film grossed in its first four weeks of release, in exchange for every share held. Thus, after the readjust, the trading price reflects an estimate of the

[3] Extended weekends and holiday releases have slightly different multipliers to better represent the films likely performance during the four weeks of release.

box-office return for a film in its first four weeks of release. All box-office returns are only for domestic sales and do not consider international or ancillary revenues.

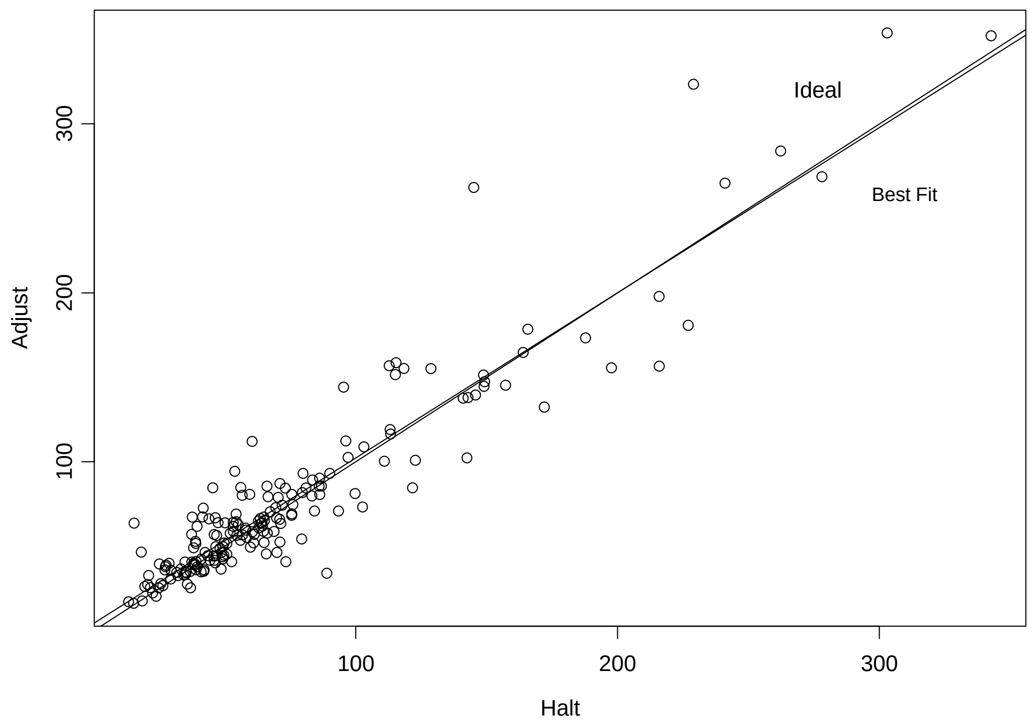
In order to facilitate an analysis of the Hollywood Stock Exchange's forecasting ability; the price at these specific times where the exchange dictates and alters a MovieStock price is a necessary piece of information for analysis. Thus, the information for the top 100 grossing films of 2004 and 2005 was purchased from the HSX. This information includes the Film Name, the IPO price and date, the Halt price, the Adjust price, the Delist price, daily close prices and timestamps for each of these actions.

A MovieStock is listed long before it is a sure thing to be released and often it is merely a concept, without any true plans for development. As a result, a film may be listed for months and years before it is eventually put into wide release. Accordingly, films that are heavily anticipated and take a long time to finalize all the various factors together (production, financing, and creative support) can be listed for a very long time on the HSX. The longest film list within our data set is the *Fantastic Four*, which was listed for roughly seven years before it was released and then delisted four weeks afterwards.

### 4.2: An Analysis of the Predictive Power of the Hollywood Stock Exchange

Comparison of the MovieStock price with the actual amount a film grossed at the box office in the US yields a demonstration of the accuracy and forecast ability of the HSX as a market. This is demonstrable in two fashions due to the nature of the HSX market. One method of analysis is a comparison of the amount of a movie's gross receipts during its first weekend times its multiplier (2.8) and the trading value at the price halt before this first weekend of release. Remember that one of the rules of this exchange is that a MovieStock price is halted before the first weekend of its release and is adjusted to 2.8 times its first weekend's gross receipts after the first weekend. The second method is analysis of trading after release and the accuracy in predicting the box-office returns of a film during its four weeks of release.

*Figure 4.2.1*
**Pre-Release Halt Price vs. HSX Adj. Revenues ($MM)**




Note: The HSX estimates for MovieStock has a slight bias to over price the worse performing movies and under price the best performing movies. This may be an indication of the risk-seeking nature of the market participants.

*Table 4.2.1*

**HSX Halt Price**
**Regression Coefficients**

| | Value | Std. Error | t value | Pr(>\|t\|) |
|---|---|---|---|---|
| (Intercept) | 4.3663 | 2.4062 | 1.8146 | 0.0711 |
| Halt | 0.978 | 0.0258 | 37.9712 | 0 |

R-Squared 0.8793

Note: the intercept is only significant at a 90% confidence level

Analysis of this regression reflects the power of the HSX. With a very high $R^2$ value, this regression explains nearly 90% of the variation in the box-office returns the first weekend of a film's release (multiplied by a constant). The ideal prediction would be along the line y=x indicating a perfect prediction for a film's box-office return. Looking at *Figure 4.2.1*, it is evident that the best fit line very nearly coincides with the ideal. The correlation between the two, the halt price and the readjust price, is 93.8%.

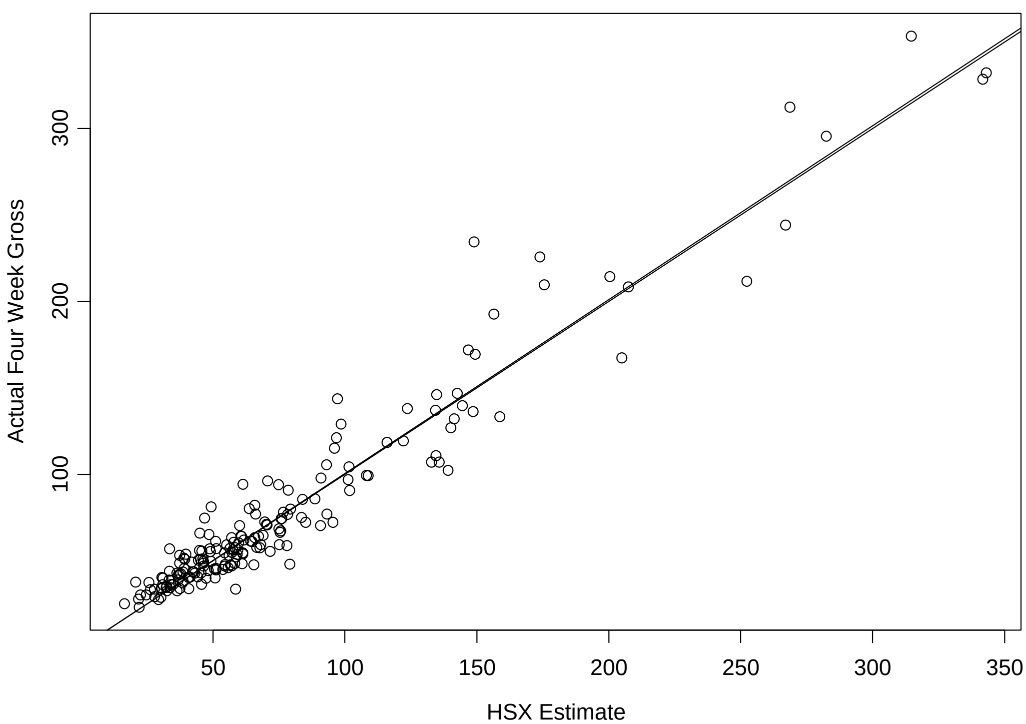
This regression hints at biases that can explain market participants' actions and preferences. Individuals prefer to bet on potential sleepers with a small chance of a very large payoff, rather than betting on movies believed to do well, with a high probability of only a moderate payoff. Because these payoffs are in fantasy H$ and not real money, the incentive may lie in an effort to make it to the leader board as fast as possible, or to sell off on EBay. If this is the case, the incentives would cause the risk-seeking behavior demonstrated.

Because this market is so remarkably accurate, it is a likely indicator that the market represents a very good accumulation of information. Discussed above, this is one of the tenets of efficient markets. One can claim reasonably that if the market accumulates all available information, it will be perfectly accurate. The HSX market participants are acting within a fantasy market with enjoyment as the primary incentive.

Therefore they are very likely to actively pursue all available information, including watching the films in which they are investing. But these individuals represent a select few members of the audience. The select few have been demonstrated to be accurate at predicting the more general sentiment of the United States as a potential audience.

Moving to an examination of the second method for demonstrating the accuracy of this prediction market, the comparison between a Hollywood Stock Exchange estimate of the four week box-office gross is compared to the true gross. This estimate is the trading value one week after release with three more weeks remaining.

*Figure 4.2.2*
**HSX Estimate vs. Actual Four Week Box-Office Gross ($MM)**




Note: The best fit line is just slightly above the ideal with an insignificant intercept and a slope of 1.0053, just .0053 off the ideal of y=x.

*Table 4.2.2*

**HSX Estimate**
**Regression Coefficients**

| | Value | Std. Error | t value | Pr(>\|t\|) |
|---|---|---|---|---|
| (Intercept) | 1.7299 | 1.9945 | 0.8674 | 0.3868 |
| HSX Estimate | 1.0053 | 0.0207 | 48.497 | 0 |

R-Squared 0.9224

Note: The intercept is not significant and the $R^2$ is very high, indicating a good fit for the regression.

The second method yields the same qualitative results and slightly better quantitative conclusions. Intuitively, the market should be more accurate after the incorporation of information regarding the first weekend's box-office returns. The best fit line is very nearly on the ideal and the regression has a very high $R^2$ which demonstrates that the HSX estimate explains very nearly all the variance within the four week gross for a film. The correlation between the HSX estimate and the actual four week gross was 96%.

In 1999, Brandon Gray founded Box Office Mojo, which offers box-office predictions and has become a staple publication for movie buffs. With over a million unique viewers per month, Box Office Mojo has grown to be one of the largest box-office informational sites online. Gray is a well-respected reviewer and forecaster and as a result, his box-office estimates have been analyzed similarly to the HSX. Pennock, et al. (2001) demonstrated that the Box Office Mojo predictions were in fact 4% better in terms of mean percentage error than that of the HSX. Since the Box Office Mojo publication was released after the HSX was established, it is possible that Brandon is influenced by the HSX forecast as a simple accurate foundation for his estimates. It is also very likely that the HSX traders are receivers of Brandon Gray's publication and are influenced by his predictions.

It is important to note that while the HSX is highly accurate directly before a film's release, as explained in the development of a box-office prediction model above, these forecasts and predictions are of limited use to studios. Their investment is largely sunk at the time of release and there exist few (if any) ways to alter the costs associated with the film's production at that late of time.

This analysis has demonstrated that the HSX as a market can arrive at an accurate prediction of a film's opening gross receipts directly before its release. Further, it will be interesting and informative to examine the mechanism that the HSX employs to arrive at this accurate prediction. The manner in which information is accumulated, aggregated and contributed to the exchange will directly be demonstrated by the accuracy of the prediction days in advance.

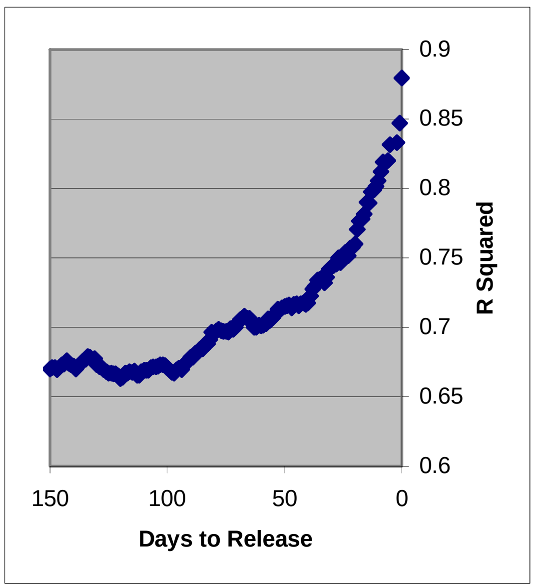
## 4.3: Incorporation of Information within the Hollywood Stock Exchange

### 4.3.1: Analysis of the Progression of R - Squared

Using two metrics for information aggregation, the manner in which the HSX accumulates information through the accuracy of its predictions will be examined. The first method requires an analysis of the progression of $R^2$ value of the regression of prices days before the film is released and its adjust price. This method was first used by Levy in 2005.[48] $R^2$ is a measure of the strength of the fit of a regression and not an indication of prediction accuracy. Despite this, $R^2$ represents a convenient summary of relationships between predictions and actual outcomes, and can be used as a tool to understand the amount of information within a market at a given time by their relative differences. The second method is a derivative of Logarithmic Score as used by Debnath (2003) called Average Logarithmic Normalized Difference.[49]

The HSX is accurate in predicting returns when a film is about to be released but it would be much more useful for studios to accurately predict box-office revenues well in advance (i.e. before production as a means to mitigate risk). In order to understand how the HSX becomes accurate and examine the potential for its use to achieve these goals, one can examine the $R^2$ value of regressions of the price of MovieStock a number of days before their film's release. Analyzing the evolution of $R^2$ and days to release gives a sense of how information is incorporated into the HSX.

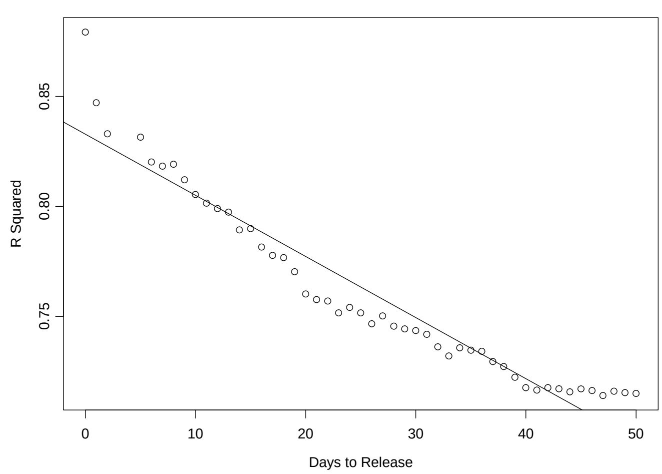
*Figure 4.3.1*
**Evolution of R-Squared and Aggregation of Information**




Examining *Figure 4.3.1* it is appears that the information and the prices of films over 100 days before their release are significantly less accurate than immediately before release. Even these less accurate predictions still are more than 15% better than the pre-release box-office model developed here or any other model used in the literature mentioned above. From a cursory analysis, it appears that the fit becomes rapidly better in the last 50 days prior to release and that the increase in $R^2$ appears approximately linear. Prior to 50 days before release, the $R^2$ value varies slightly. Indeed, it appears to remain relatively constant between values of .65 and .7. Using $R^2$ as an indication of information aggregation of the market this phenomenon makes rational sense. A small amount of new information about a film is available 50 days before the release. The national advertising campaign builds buzz roughly two months before a film's release. As a result, the information accumulated to make the predictions more accurate would be the perceived quality of the film, the film's advertising campaign and an estimation of the public's response to the perceived quality.

While for an individual film a big news break or information being presented to the public would likely have a dramatic effect, individual films have no direct impact on the market. As a result, no block or dramatic individual jumps are evident. Examining the 50 days directly preceding a film's release produces the Close Up on *Figure 4.3.1*.

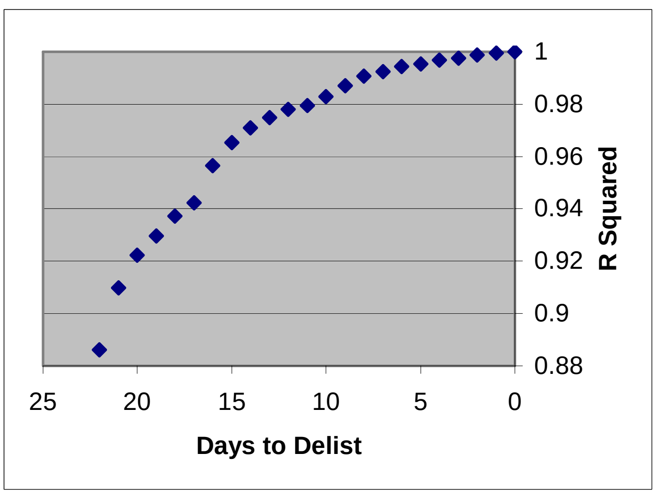
Close Up on *Figure 4.3.1* and a Linear Fit




Note: The X-axis is reversed from *Figure 4.3.1* to demonstrate the decrease of total information aggregated as measured by fit within the HSX as Days to Release increases.

The best fit line is *R Squared =.8328 - .0028 (Days to Release)* with an $R^2$ of .9228. This indicates the progression of the $R^2$ of the market estimates is linear for the 50 days prior to the release of a film. This fit indicates that information is accumulated in a linear fashion from 50 days prior to release until the release. *Figure 4.3.2* shows that after the release and the incorporation of information from the opening weekend, the incorporation of new information slows as the market approaches its accurate prediction. Intuitively, this makes sense as the potential variability after the opening weekend is dramatically decreased. As a result, new information about the film will not change the markets view of box-office potential. Also, the total amount of information about the film is essentially available to the public once reviews have been written, the advertising campaign has run its course, and an individual has actually seen the film. Box-office information is added each day that a film is in release tempered by the recognition that the opening weekend is such a large portion of the box-office revenue and indicative of the total gross that a film will receive. The information available only after the release has less of an affect on the market and less of an affect on the total box-office returns.

*Figure 4.3.2*
**Evolution of R-Squared from Release to Delist**



Naturally, as the market is saturated with all relevant information, it approaches certainty with respect to its prediction. Information is no longer incorporated in a linear manner post-release. *Figure 4.3.2* demonstrates a decrease in the rate of information accumulated represented by a decrease in the accuracy.

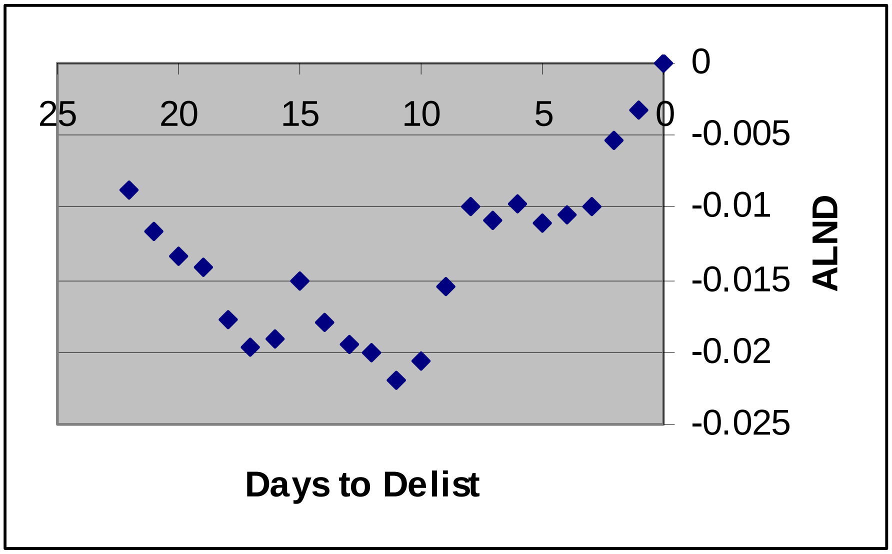
4.3.2: An Analysis of Average Logarithmic Normalized Difference

Average Logarithmic Normalized Difference is defined as:

EQ. 4.1: $$ALND = 1/N \cdot \sum_{i=1}^{N} \log \left| {(p_{delist} - p)}\big/{p_{delist}} \right|$$

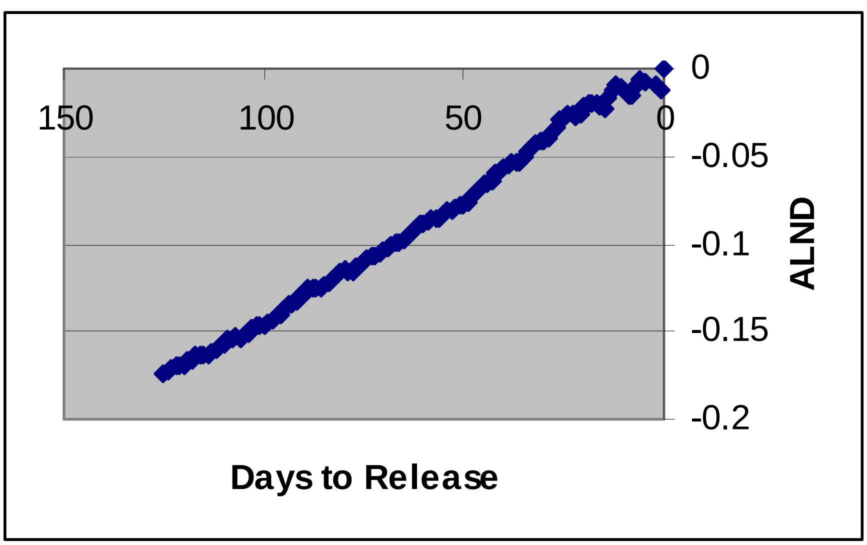
It essentially represents how close the MovieStock's estimate is to the actual box-office gross, either in the opening weekend or within the first four weeks of release, a certain number of days before release or being delisted. Looking at the progression of the ALND relative to other values, one can gather sense of information integrated within the market. An examination of the Average Log Normalized Difference after a film's initial release, found in *Figure 4.3.3*, demonstrates the market actually has as accurate a sense of a film's success directly after its release as immediately preceding its delist. As the MovieStock is traded in release, the certainty and accuracy of market take a small dip. But as the four weeks continue and each daily gross contributes as a smaller and smaller percentage of the total domestic gross, the certainty and information available is completely assimilated into the market. This may be evidence of a systematic error of the HSX traders.

*Figure 4.3.3*
**Evolution of ALND from Release to Delist**

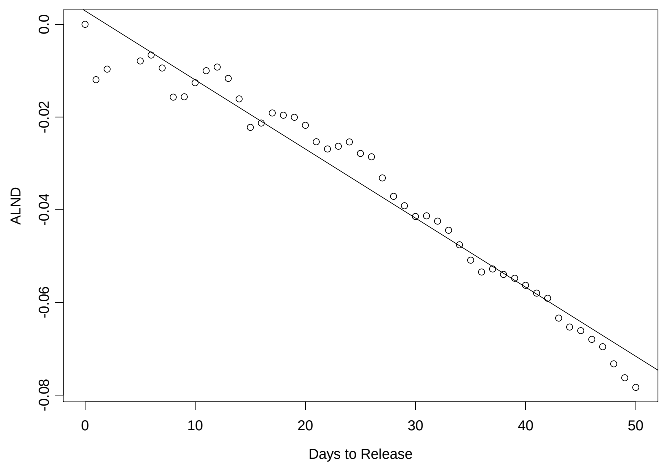



An examination of the Average Log Naturalized Difference before release as seen in *Figure 4.3.4*, demonstrates a linear relationship of ALND and days to release. Unlike the analysis of the progression of $R^2$, it appears there is a linear relationship between ALND and Days to Release even prior to 50 days until release.

*Figure 4.3.4*
**Evolution of ALND until Release**



Close Up on *Figure 4.3.4* and a Linear Fit




Note: Similar to with the preceding close up, the X-axis is reversed from *Figure 4.3.4* to demonstrate the decrease of the total information as measured by log normalized difference aggregated within the HSX as Days to Release increases.

The best fit line in the Close Up on *Figure 4.3.4* illustrates linear accumulation of information and is represented by the equation: ALND=.0029 - .0015 (Days to Release). The regression has an $R^2$ of .9613 indicating a good fit, and confirming that the ALND progresses linearly with each day closer to the release of the film. As with $R^2$, this indication of a linear accumulation of information makes rational sense within the motion picture industry. It further confirms the result and the explanation that information is accumulated linearly.

## 4.4: What are the Ramifications for Studios?

Consider the information that is available to the Hollywood Stock Exchange marketplace of participants in the first 150 days prior to release. Likely, studios and public relation officers for the director and cast have tied cast lists and director to a specific project. If the film is a piece of established intellectual property (a franchise or a transfer of intellectual property from another medium), the marketplace will likely incorporate their expectations based upon a rational assessment of what qualities can be transfered from the seminal work to the new derivative. From this limited amount of information the marketplace is able to ascertain with greater accuracy than any model developed the amount a film will gross in the box-office (although for all they know the movie is merely a chance that a concept with a creative cast linked to it will be created).

The linear increase of information directly up to the release of the film improves the accuracy of the marketplace prediction dramatically. Now consider what information it is likely being contributed to the marketplace. Film trailers have been released for larger films, often as many as two or three different trailer or teasers released in concert with the beginning of the advertising campaign. As a part of the film's advertising campaign, early releases to members of the industry can result in early reviews and buzz created. Finally, the competition against which a film is being released against has been solidified and their quality is being clarified via their own advertising campaigns. By assimilating this new information, the market is able to take a dramatic jump in accuracy.

Within other industries, prediction markets are being molded into decision market, where executives use the public collective beliefs to aid and in some cases determine the outcome of important issues facing the company. Motion Picture Industry executives have long relied on diverse focus groups from which they believe they can ascertain the public's desires and interest. It has become clear that industry executives are not the accurate barometers of public interest they desire to be. In fact, according to this research they should leave prediction of public interest to a self-selected public group with proper incentives in place. If industry executives provided a group similar to the Hollywood Stock Exchange a modicum of information (as they will likely ferret out most of it themselves) industry executives could get accurate predictions of domestic box-office returns (as demonstrated) and likely other ancillary window revenue streams too. The information required by this market does not necessitate the creation of an entire movie. In fact, this information is requires to be little more than a concept and the face of the creative cast. This can dramatically limit the sunk costs a studio faces prior to release in production. Seeking to create a model that can help mitigate investment risk on behalf of the studios, the analysis demonstrated that the only suitable known method for forecast is in fact the HSX.

# Chapter Five:

## Concluding Remarks

### 5.1: Summary of Analyses

#### 5.1.1: Forecasting Box-Office Returns

It is clear from the analysis of box-office determinants that there exist factors which studios should consider while producing motion pictures. Light has been shed on certain factors like May release, advertising budget and genre choice. While some evidence has been presented for the beneficial affects of advertising budget, no analysis was performed on an optimization of advertising via various media nor further analysis of the diminishing effects of advertising budgets. Because of the secretive nature of the motion picture industry, this further analysis would probably require sponsorship of a major studio and access to their proprietary information.

In a summary analysis of the effects of a financially successful director and cast, the analysis suggests that actors may not be worth what they cost studios in increased production budgets. Again, because studios do not release a large amount of information including actor salaries, further analysis will likely require studio sponsorship. In fact, because studios so often manipulate budget numbers direct analysis of individual actor's contracts will likely be required to analyze the profitability of Hollywood's faces.

The analysis of box-office returns also further demonstrates the validity of the intuition that Litman, one of the leading researchers of box-office determents, vocalized about film box-office prediction, namely, they are a "wild guess". The model developed here explains only 54% of the variance. While it provokes helpful questions about studio decisions, it falls short of the original intention of mitigating investment risk. The model only correctly predicted the success of a third of the films on which it was tested and its weakness is readily apparent. This weak fit sparked analysis of the Hollywood Stock Exchange in seeking to find a reliable method for studios to potentially forecast box-office returns and use this forecast to mitigate production risks.

5.1.2: HSX Accuracy and Informational Content

Analysis of the Hollywood Stock Exchange demonstrated the strength of the prediction markets ability to forecast box-office returns immediately before release. The research exhibits the accuracy of the market price as a prediction of a films opening weekend box-office gross and as a prediction of the four-week total gross.

Examination of the incorporation of information into the market demonstrated that up to 50 days before release, both the information contained within the market and accuracy of the prediction increase linearly. A regression of the market price and estimation of the films opening weekend gross 150 days before release is dramatically more accurate than any model developed previously. This indicates that with less information than contained in many models, and information that is available to studios prior to production of a film, the market arrives at a more accurate prediction than these better-informed models.

What are the implications of these results? First, secondary parties can mine online information aggregators like the Hollywood Stock Exchange and enjoy some assurances as to the accuracy of the information they can obtained. Building a model that incorporates HSX information has the potential to help studios predict how audiences will respond to a potential film. This serves to potentially allow studios to mitigate their investment risk dramatically earlier than models contained within the literature reviewed.

## 5.2: Potential Avenues for Future Applications

The obvious conclusion from the research is that studios need to build a prediction market or harness the Hollywood Stock Exchange and develop a model surrounding this market. This model would have the potential to mitigate the huge investment required for production of a major motion picture and would be worth millions to studios. Future research into the mechanism of information incorporation, and the optimal level of information a studio should provide their prediction market, would be beneficial to the practical application of their model. The prediction market has demonstrated the ability to more accurately assimilate information, resulting in a very accurate prediction. This ability can be harnessed and these predictions can transform risk that exists in the motion picture industry currently.

Converting the HSX or the creation of a new a real money market would give theaters and film producers the potential to hedge their investment and lessen their risks. Already the potential for a real money market is hinted at by the sale of HSX accounts on EBay.com. Pennock, Lawrence, Giles and Nielson (2001) cited a trading value of H$1 million = $1 US based on previous sales on EBay. A real money market would need to trade many millions of dollars in order to provide a venue that will stabilize Hollywood returns since production budgets can be huge.

It has become clear, though, that improvements are necessary. The famous director George Lucas has declared. "I predict that by 2025 the average movie will cost only $15 million."[50] Studios are spending too much on production budgets and not receiving sufficient return on investment to justify the risk assumed. Thus, some method of mitigation needs to be developed.

One limitation of my analysis was the simple metric for the estimation of profit. This metric does not take into account any ancillary revenues, so it is not a full indication of economic incentive for the studios. This being said, forecasting box-office returns has already been demonstrated to be incredibly difficult, at best. The simplification was made in an effort to exhibit the strength of the Hollywood Stock Exchange relative to a linear regression model and the models presented in the relevant literature reviewed. The model also demonstrates that a further analysis of actor's box-office value would prove lucrative to studios.

Another limitation of the analysis was the variable indication of franchise. I did not take into account whether adaptations of intellectual property from other media counted within the franchise classification. As franchise was shown to be a strong determinant of box-office returns, further research into what qualifies as franchise would likely prove beneficial.

The power of prediction markets lies in their ability to aggregate information from the many participants. The complex interplay between traders allows for the accumulation of all the data each participant is acting upon and as a result the market price is a reflection of this data. This power has the potential to be leveraged. In this case, motion picture studies have a huge opportunity to control the risk inherent in their business, if they choose to embrace it.

Appendix A: Frequency of Binary Variables

| | % of All Movies |
|---|---|
| *Month of Release* | |
| December | 0.13 |
| November | 0.0925 |
| October | 0.095 |
| September | 0.06 |
| August | 0.0725 |
| July | 0.0925 |
| June | 0.09 |
| May | 0.0775 |
| April | 0.08 |
| March | 0.0975 |
| February | 0.0625 |
| January | 0.05 |
| *MPAA Rating* | |
| G | 0.0375 |
| PG | 0.2075 |
| PG.13 | 0.4825 |
| R | 0.2725 |
| *Genre* | |
| Action | 0.32 |
| Adventure | 0.27 |
| Animated | 0.065 |
| Biography | 0.0175 |
| Comedy | 0.475 |
| Crime | 0.16 |
| Drama | 0.455 |
| Family | 0.185 |
| Fantasy | 0.16 |
| History | 0.025 |
| Mystery | 0.0875 |
| Romance | 0.1825 |
| Science.Fiction | 0.1 |
| Sport | 0.0425 |
| Thriller | 0.365 |
| War | 0.035 |
| High Competition | 0.595 |
| Franchise | 0.1625 |

Appendix B: Example of Film Data that Trained Box-Office Model

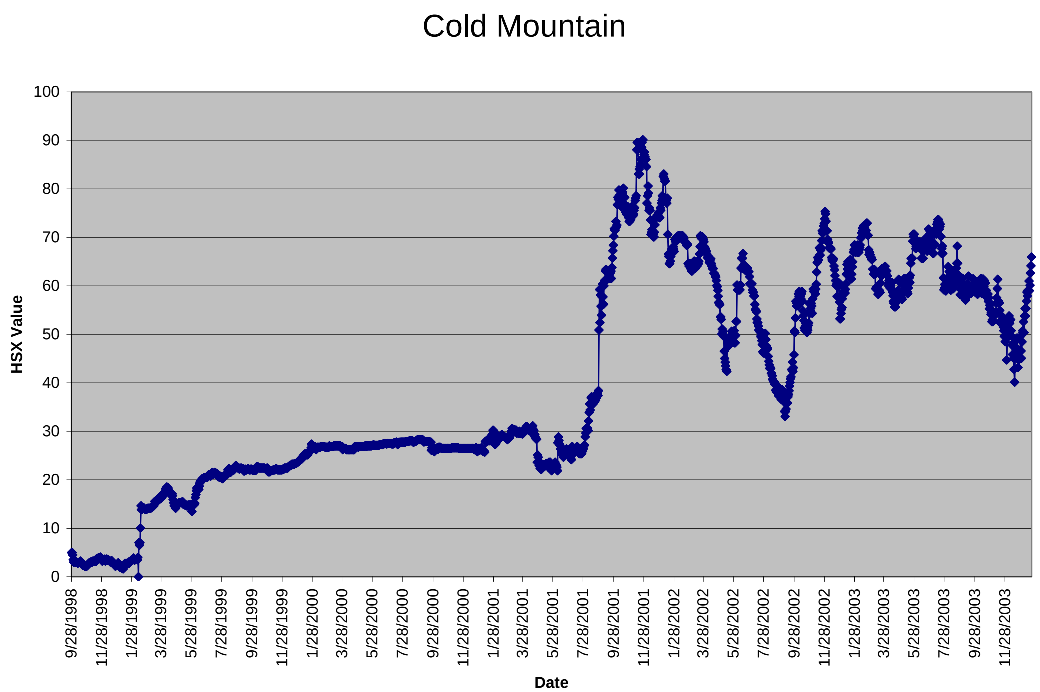
| Title | Total Gross | Opening Theaters | Open Date | Month | Dec | Nov | Oct | Sep | Aug | Jul | Jun | May | Apr | Mar | Feb | Jan | Comp | Fran | G | PG | PG-13 | R |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Garden State | $26,782,316 | 9 | 7/28/2004 | 7 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| All About the Benjamins | $25,916,319 | 1,505 | 3/8/2002 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| Brown Sugar | $27,363,891 | 1,372 | 10/11/2002 | 10 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 |
| 28 Days Later | $45,064,915 | 1,260 | 6/27/2003 | 6 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| Calendar Girls | $31,041,759 | 24 | 12/19/2003 | 12 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| The Hours | $41,675,994 | 11 | 12/27/2002 | 12 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| Johnson Family Vacation | $31,203,964 | 1,317 | 4/7/2004 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 |
| One Hour Photo | $31,597,131 | 7 | 8/21/2002 | 8 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| Barbershop | $75,782,105 | 1,605 | 9/13/2002 | 9 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 |
| Frida | $25,885,000 | 5 | 10/25/2002 | 10 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| The Pianist | $32,572,577 | 6 | 12/25/2002 | 12 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| The Lizzie McGuire Movie | $42,734,455 | 2,825 | 5/2/2003 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| Monster | $34,469,210 | 4 | 12/24/2003 | 12 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| Eternal Sunshine of the Spotl | $34,400,301 | 1,353 | 3/19/2004 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| Napoleon Dynamite | $44,540,956 | 6 | 6/11/2004 | 6 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 |
| Fahrenheit 9/11 | $119,194,771 | 868 | 6/23/2004 | 6 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| Return to Never Land | $48,430,258 | 2,605 | 2/15/2002 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 |
| Lost in Translation | $44,585,453 | 23 | 9/12/2003 | 9 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| The Count of Monte Cristo | $54,234,062 | 2,007 | 1/25/2002 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 |
| A Walk to Remember | $41,281,092 | 2,411 | 1/25/2002 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 0 |
| Big Fat Liar | $48,360,547 | 2,531 | 2/8/2002 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 0 |
| Clockstoppers | $36,989,956 | 2,540 | 3/29/2002 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 |
| Jonah: A VeggieTales Movie | $25,581,229 | 940 | 10/4/2002 | 10 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 |
| Jackass: The Movie | $64,255,312 | 2,509 | 10/25/2002 | 10 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| About Schmidt | $65,016,287 | 6 | 12/13/2002 | 12 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| Drumline | $56,399,184 | 1,836 | 12/13/2002 | 12 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 |
| Rugrats Go Wild | $39,402,572 | 3,041 | 6/13/2003 | 6 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| The Fighting Temptations | $30,250,745 | 2,026 | 9/19/2003 | 9 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 |
| Mystic River | $90,135,191 | 13 | 10/8/2003 | 10 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| Honey | $30,308,417 | 1,942 | 12/5/2003 | 12 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |

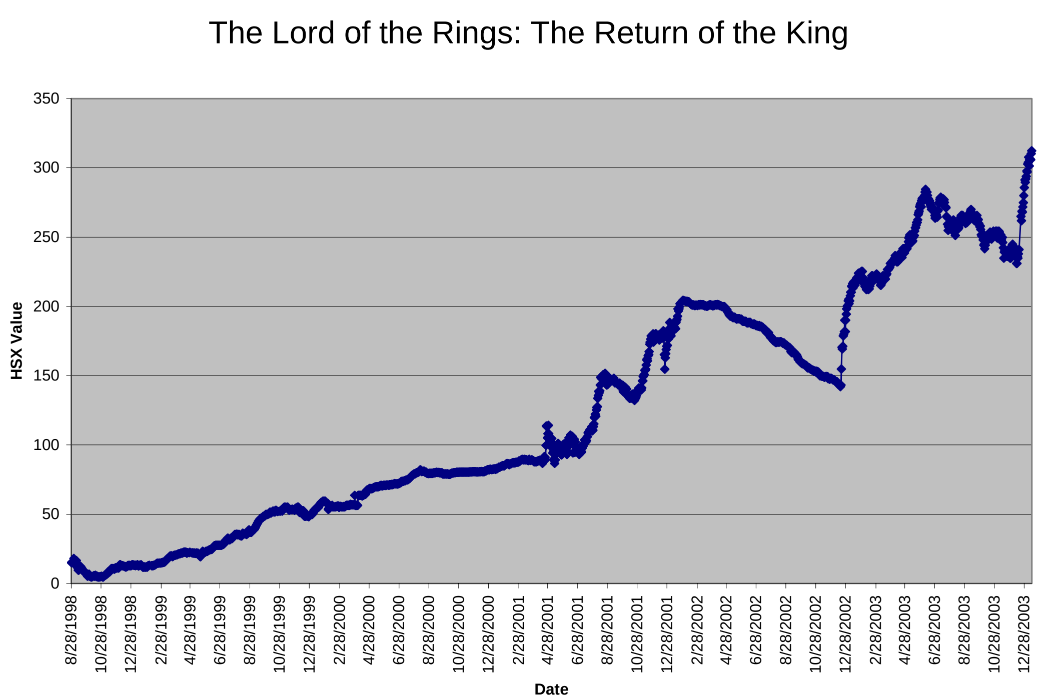
| Title | Budget | Advertising | Act | Adv | Ani | Bio | Com | Cri | Dra | Fam | Fan | His | Mys | Rom | Sci Fi | Spo | Thr | War | Director | Actor |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Garden State | $2,500,000 | $5,000,000 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 48496350 |
| All About the Benjamins | $15,000,000 | $7,000,000 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 31833642 |
| Brown Sugar | $8,000,000 | $7,000,000 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 28275450 |
| 28 Days Later | $8,000,000 | $7,000,000 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 13996713 | 55307020 |
| Calendar Girls | $10,000,000 | $7,000,000 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 12189009 |
| The Hours | $25,000,000 | $8,000,000 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 21995263 | 29258892 |
| Johnson Family Vacation | $12,000,000 | $8,000,000 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| One Hour Photo | $12,000,000 | $10,000,000 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 62406545 |
| Barbershop | $12,000,000 | $10,000,000 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 31378300 |
| Frida | $12,000,000 | $10,000,000 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| The Pianist | $35,000,000 | $10,000,000 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 12143069 | 0 |
| The Lizzie McGuire Movie | $17,000,000 | $10,000,000 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 58625100 |
| Monster | $5,000,000 | $10,000,000 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 26548989 |
| Eternal Sunshine of the Spotl | $20,000,000 | $10,000,000 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 27204470 |
| Napoleon Dynamite | $400,000 | $10,000,000 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Fahrenheit 9/11 | $6,000,000 | $12,000,000 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 60969777 |
| Return to Never Land | $20,000,000 | $13,000,000 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Lost in Translation | $4,000,000 | $13,000,000 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4859475 | 30846563 |
| The Count of Monte Cristo | $35,000,000 | $15,000,000 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 45739580 | 18721000 |
| A Walk to Remember | $11,800,000 | $15,000,000 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 37122018 |
| Big Fat Liar | $15,000,000 | $15,000,000 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 51135530 | 68164414 |
| Clockstoppers | $26,000,000 | $15,000,000 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 23017152 |
| Jonah: A VeggieTales Movie | $14,000,000 | $15,000,000 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Jackass: The Movie | $5,000,000 | $15,000,000 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 28505000 |
| About Schmidt | $30,000,000 | $15,000,000 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 13650200 |
| Drumline | $20,000,000 | $15,000,000 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 |
| Rugrats Go Wild | $25,000,000 | $15,000,000 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 31118301 | 79828962 |
| The Fighting Temptations | $30,000,000 | $15,000,000 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 27870728 |
| Mystic River | $30,000,000 | $15,000,000 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 34854611 | 24864962 |
| Honey | $18,000,000 | $15,000,000 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 59197200 |

Appendix C: Examples of Prediction from the Trained Box-Office Model

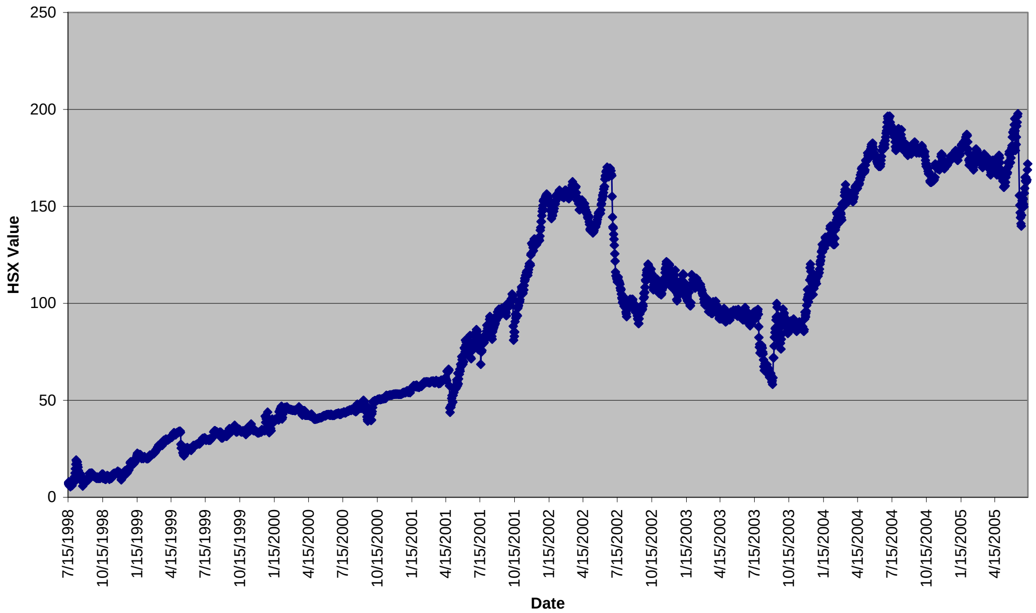
| Title | Predicted Fit | Standard Error | Absolute Error | Percent Error | Lower Bound | Upper Bound | Actual Total Gross | Within Bounds |
|---|---|---|---|---|---|---|---|---|
| Rumor Has It | 72,361,630 | 6,115,110 | 29,958,717 | 71% | 60,315,742 | 84,407,517 | 42,402,913 | 0 |
| Munich | 108,079,221 | 13,671,954 | 66,965,831 | 163% | 81,147,436 | 135,011,005 | 41,113,390 | 0 |
| The Ringer | 37,123,196 | 6,271,470 | 2,481,247 | 7% | 24,769,301 | 49,477,090 | 34,641,949 | 1 |
| Fun with Dick and Jane | 56,987,015 | 11,497,981 | 49,857,290 | 47% | 34,337,647 | 79,636,384 | 106,844,305 | 0 |
| Cheaper by the Dozen 2 | 95,309,364 | 11,676,868 | 15,476,098 | 19% | 72,307,613 | 118,311,116 | 79,833,266 | 1 |
| The Family Stone | 51,868,028 | 6,475,419 | 7,432,119 | 13% | 39,112,383 | 64,623,673 | 59,300,147 | 1 |
| Hoodwinked | 28,927,264 | 14,617,424 | 9,615,143 | 25% | 133,041 | 57,721,487 | 38,542,407 | 1 |
| King Kong | 180,078,821 | 14,656,814 | 33,477,879 | 16% | 151,207,004 | 208,950,637 | 213,556,700 | 0 |
| The Chronicles of Narnia: The Lion, the Witch and the Wardrobe | 243,644,266 | 26,407,866 | 34,969,551 | 13% | 191,624,568 | 295,663,965 | 278,613,817 | 1 |
| Memoirs of a Geisha | 129,396,942 | 12,858,729 | 75,105,575 | 138% | 104,067,094 | 154,726,790 | 54,291,367 | 0 |
| | **Average:** | 9,852,109 | 29,587,173 | 60% | | **Total Count:** | | 16 |

Appendix D: Examples of HSX MovieStocks Trading Data

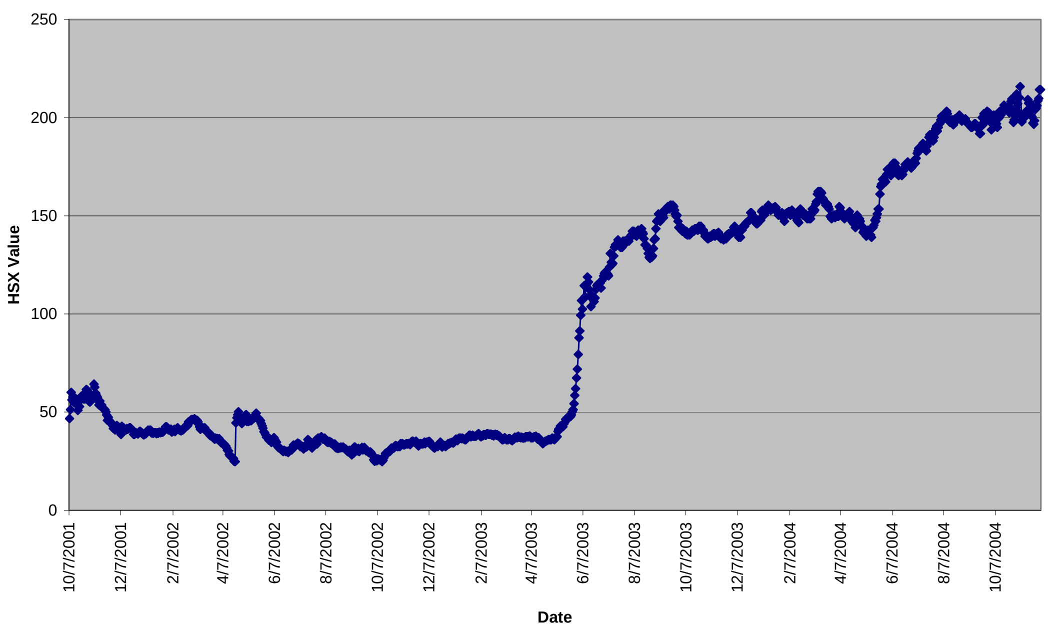


## Batman Begins




## The Incredibles



WORKS CITED:

American Gaming Association Industry Fact Sheet (http://www.americangaming.org/Industry/factsheets/statistics_detail.cfv?id=7)

Berg. J., Forsythe, R., Nelson, F., and Rietz, T. (2000) "Results from a Dozen Years of Election Futures Markets Research," University of Iowa

Cowgill, Bo. (9/21/2005) "Putting crowd wisdom to work" (http://googleblog.blogspot.com/2005/09/putting-crowd-wisdom-to-work.html)

De Vany, A. and Walls, W. D. (1999) "Uncertainty in the Movie Industry: Does Star Power Reduce the Terror of the Box Office?" presented at the annual meeting of the American Economic Association, New York, January 1999

Debnath, S., Pennock D., Giles, C. L., Lawrence, S. (2003) "Information Incorporation in Online In-Game Sports Betting Markets," *Fourth ACM Conference on Electronic Commerce*, pp. 258-259

Desai, M.A., G.J. Loeb, and M.F. Veblen (2002) "The Strategy and Sources of Motion Picture Finance," Harvard Business School Teaching Case, No. 9-2-3-007

Dortch, Shannon. (1996) "Going to the movies – demographic aspects of movie attendance" *American Demographics*, Dec. 1996 <http://www.findarticles.com/p/articles/mi_m4021/is_n12_v18/ai_18894246>

Einav, Liran. (2002) "Seasonality and Competition in Time: An Empirical Analysis of Release Date Decisions in the U.S. Motion Picture Industry," based upon a Harvard University dissertation 2002 <http://www.stanford.edu/~leinav/Release_Dates.pdf>

Eliashberg. J, Elberse, A. Leenders, M. (2005) "The Motion Picture Industry: Critical Issues in Practice, Current Research & New Research Directions" <http://www.hbs.edu/research/pdf/05-059.pdf>

Grossman, S. (1981) "An introduction to the theory of rational expectations under asymmetric information," *Review of Economic Studies*, 48(4) 541-559

Grove, Lloyd. (2006) "Lucas: Big pics are doomed" *Daily News*, March 5, 2006 <http://www.nydailynews.com/front/story/397167p-336664c.html>

Hanson. R., Oprea, R., and Porter, D. (2004) "Information Aggregation and Manipulation in an Experimental Market," Interdisciplinary Center for Economic Science George Mason University 9/21/2004

Hanson. R. and Oprea, R. (2004) "Manipulators Increase Information Market Accuracy," George Mason University July 2004

Jaeger, James. *The Movie Industry* <http://www.mecfilms.com/moviepubs/memos/moviein.htm>

Jullien, B. and Salanie, B. (2000) “Estimating Preferences under Risk: The Case of Racetrack Bettors,” *Journal of Political Economy*, 108, 503-30

Krider, R. and Weinberg, C. (1998) “Competitive dynamics and the introduction of new products: The motion picture timing game,” *J. Marketing Res.* 35(1) 1–15

Levy, Todd. (2005) *Buying and Selling Movies: An Analysis of the Hollywood Stock Exchange,* an ORFE Princeton University Senior Thesis

Litman, B.R. and Ahn, H. (1998). “Predicting Financial Success of Motion Pictures” appearing in *The Motion Picture Mega-Industry*. By B.R Litman. Allyn Bacon Publishing Inc. Boston, MA

Litman, B. R., and Kohl, L. S. (1989) “Predicting Financial Success of Motion Pictures: The '80s Experience,” *Journal of Media Economics*, 2, 35-50

Litman, B.R. (1983). “Predicting Success of Theatrical Movies: An Empirical Study” *Journal of Popular Culture,* Vol. 16, No. 9, pp.159-175.

Motion Picture Association Worldwide Market Research, *US Entertainment Industry 2004 MPA Market Statistics*

Pennock, D., Lawrence, S., Giles, C. L., and Nielsen, F. A. (2001) “The Power of Play: Efficiency and Forecast Accuracy in Web Market Games,” *Science 291: 987-988* 2/9/2001

Pennock, D., Debnath, S., Glover, E., and Giles, C. L. (2002) “Modeling Information Incorporation in Markets, with Application to Detecting and Explaining Events” E-business Research Center Working Paper. September 2002

Prechter, R. (1999) *The Wave Principle of Human Social Behavior*, New Classics Library. 1999

“Prediction market.” *Wikipedia, The Free Encyclopedia*. 15 Jan 2006, <http://en.wikipedia.org/w/index.php?title=Prediction_market&oldid=35290123>

Ravid, S.A. (1998) “Information, Blockbusters and Stars: A Study of the Film Industry”. Working paper, New York University.

Robertson, David. “Making Monster Money,” InsideEdge Magazine <http://www.hsx.com/about/insideedge_p1.htm>

Sharda, R. and Delen D. (2002) "Forecasting Box-Office Receipts of Motion Pictures Using Neural Networks,"
<http://orcs.bus.okstate.edu/sharda/research/Forecasting%20Box-Office%20Receipts100102.pdf>

Servan-Schreiber, E., Wolfers, J., Pennock, D. and Galebach, B. (2004) "Prediction Markets: Does Money Matter?" originally published in *Electronic Markets*, 14-3, September 2004.

Suriweski, James (2004) *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations.* Random House.

Sawhney, M. and Eliashberg, J. (1996) "A Parsimonious Model for Forecasting Gross Box-Office Revenues of Motion Pictures,"
<http://knowledge.wharton.upenn.edu/papers/785.pdf>

Shannon, C. and Weaver, W. (1963) *Mathematical Theory of Communication.* University of Illinois Press, Urban Ill.

Tetlock, P. (2004) "How Efficient Are Information Markets? Evidence from an Online Exchange" Harvard University. January 2004

Vogel, Harold L. (2001). *Entertainment Industry Economics: A Guide for Financial Analysis.* 5th ed., Cambridge University Press, Cambridge

Wolfers. J and Zitzewitz, E. (2004) "Prediction Markets," *Journal of Economic Perspectives* Spring 2004

Woodland, L. and Woodland, B. (2001) "Market Efficiency and Profitable Wagering in the National Hockey League: Can Bettors Score on Longshots?" *Southern Economic Journal* 67, 983-95

End Notes:

[1] Tradesports.com Drive Calculator (http://www.tradesports.com/aav2/misc/NFLDriveCalculator.jsp)
[2] American Gaming Association Industry Fact Sheet (http://www.americangaming.org/Industry/factsheets/statistics_detail.cfv?id=7)
[3] Ibid.
[4] Top Grossing Films Worldwide (http://www.the-movie-times.com/thrsdir/alltime.mv?world+ByWG)
[5] Jaeger, James. *The Movie Industry* (http://www.mecfilms.com/moviepubs/memos/moviein.htm)
[6] Vogel, Harold L. (2001) *Entertainment Industry Economics: A Guide for Financial Analysis.* 5th ed., Cambridge University Press, Cambridge
[7] Motion Picture Association Worldwide Market Research *US Entertainment Industry 2004 MPA Market Statistics*
[8] Jaeger, James. "The Movie Industry"
[9] Robertson, David. *Making Monster Money* InsideEdge Magazine (http://www.hsx.com/about/insideedge_p1.htm)
10 Litman, B.R. and Ahn, H. (1998). "Predicting Financial Success of Motion Pictures" appearing in *The Motion Picture Mega-Industry*. By B.R Litman. Allyn Bacon Publishing Inc. Boston, MA.
[11] Ibid.
[12] Litman, B.R. (1983) "Predicting Success of Theatrical Movies: An Empirical Study," *Journal of Popular Culture*, Vol. 16, No. 9, pp.159-175.
[13] Sharda, R. and Delen D. (2002) "Forecasting Box-Office Receipts of Motion Pictures Using Neural Networks" (http://orcs.bus.okstate.edu/sharda/research/Forecasting%20Box-Office%20Receipts100102.pdf)
[14] Sawhney, M. and Eliashberg, J. (1996) "A Parsimonious Model for Forecasting Gross Box-Office Revenues of Motion Pictures" (http://knowledge.wharton.upenn.edu/papers/785.pdf)
[15] De Vany, A. and Walls, W. D. (1999) "Uncertainty in the Movie Industry: Does Star Power Reduce the Terror of the Box Office?" presented at the annual meeting of the American Economic Association, New York, January 1999
[16] (http://www.biz.uiowa.edu/iem/)
[17] "Prediction market." *Wikipedia, The Free Encyclopedia*. 15 Jan 2006, (http://en.wikipedia.org/w/index.php?title=Prediction_market&oldid=35290123).
[18] Berg. J., Forsythe, R., Nelson, F., and Rietz, T. (2000) "Results from a Dozen Years of Election Futures Markets Research." University of Iowa. November 2000
[19] (http://www.tradesports.com)
[20] Tetlock, P. (2004) "How Efficient Are Information Markets? Evidence from an Online Exchange." Harvard University. January 2004
[21] Ibid.
[22] (http://buzz.research.yahoo.com/bk/)
[23] Servan-Schreiber, E., Wolfers, J., Pennock, D. and Galebach, B. (2004) "Prediction Markets: Does Money Matter?" originally published in *Electronic Markets*, 14-3, September 2004.
[24] "Prediction market." *Wikipedia*
[25] Cowgill, Bo. (9/21/2005) "Putting crowd wisdom to work" (http://googleblog.blogspot.com/2005/09/putting-crowd-wisdom-to-work.html)
[26] Suriweski, James (2004) *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations.* Random House.
[27] Wolfers. J and Zitzewitz, E. (2004) "Prediction Markets," *Journal of Economic Perspectives* Spring 2004
[28] Grossman, S. (1981) "An introduction to the theory of rational expectations under asymmetric information." *Review of Economic Studies*, 48(4) 541-559
[29] Pennock, D., Lawrence, S., Giles, C. L., and Nielsen, F. A. (2001) "The Power of Play: Efficiency and Forecast Accuracy in Web Market Games" *Science 291: 987-988* 2/9/2001
[30] Ibid.

[31] Hanson. R., Oprea, R., and Porter, D. (2004) “Information Aggregation and Manipulation in an Experimental Market” Interdisciplinary Center for Economic Science George Mason University. September 2004
[32] Hanson. R. and Oprea, R. (2004) “Manipulators Increase Information Market Accuracy.” George Mason University. July 2004
[33] Pennock, D., Debnath, S., Glover, E., and Giles, C. L. (2002) “Modeling Information Incorporation in Markets, with Application to Detecting and Explaining Events.” E-business Research Center Working Paper. September 2002
[34] Shannon, C., and Weaver, W. (1963) *Mathematical Theory of Communication.* University of Illinois Press, Urban Ill.
[35] Jullien, B. and Salanie, B. (2000) “Estimating Preferences under Risk: The Case of Racetrack Bettors,” *Journal of Political Economy*, 108, 503-30
[36] Woodland, L. and Woodland, B. (2001) “Market Efficiency and Profitable Wagering in the National Hockey League: Can Bettors Score on Longshots?” *Southern Economic Journal* 67, 983-95
[37] Dortch, Shannon. (1996) “Going to the movies – demographic aspects of movie attendance” *American Demographics*, Dec. 1996 < http://www.findarticles.com/p/articles/mi_m4021/is_n12_v18/ai_18894246>
[38] Pennock, D. et al. “The Power of Play …”
[39] Pennock, D. et al. “The Power of Play …”
[40] Jaeger, James. “The Movie Industry”
[41] Einav, Liran. (2002) “Seasonality and Competition in Time: An Empirical Analysis of Release Date Decisions in the U.S. Motion Picture Industry” based upon Harvard University dissertation 2002 <http://www.stanford.edu/~leinav/Release_Dates.pdf>
[42] Eliashberg. J., Elberse, A., and Leenders, M. (2005) “The Motion Picture Industry: Critical Issues in Practice, Current Research & New Research Directions” < http://www.hbs.edu/research/pdf/05-059.pdf>
[43] <www.imbd.com>
[44] <http://www.mpaa.org/FilmRatings.asp>
[45] De Vany, A. and Walls, W.D. (1999)
[46] Prechter, R. (1999) *The Wave Principle of Human Social Behavior*, New Classics Library. 1999
[47] De Vany, A. and Walls, W. D. (1999)
[48] Levy, Todd. (2005) *Buying and Selling Movies: An Analysis of the Hollywood Stock Exchange* an ORFE Princeton University Senior Thesis
[49] Debnath, S., Pennock D., Giles, C. L., and Lawrence, S. (2003) “Information Incorporation in Online In-Game Sports Betting Markets” *Fourth ACM Conference on Electronic Commerce*, pp. 258-259
[50] Grove, Lloyd. (2006) “Lucas: Big pics are doomed” Daily News March 5, 2006 <http://www.nydailynews.com/front/story/397167p-336664c.html>