Investing in Movies

Mark J. Ferrari and Andrew Rudd[1]

August 2006

Introduction

The world economy is increasingly devoted to the production of intangible assets. "The economic output of the United States has become…predominantly conceptual," according to Greenspan (2003). Borod (2005) states, "The sheer volume of intellectual property worldwide is staggering." A large part of this intellectual property is filmed content created by the major movie studios. Despite its increasing importance, analysis of intellectual property as an asset class remains difficult because of the scarcity of systematically collected data. Filmed content is emerging as an exception. Reasonably reliable information about the creative content and financial performance of feature films is available to those willing to assemble it from a variety of sources. The goal of this article is to use such data to understand the returns to investing in movies.

Movie projects are priced in an incomplete and inefficient market where valuation and arbitrage are difficult if not impossible, suggesting an opportunity for active management. However, there are at least four reasons why this task is quite different

---

[1] Mark J. Ferrari is director of research and Andrew Rudd is managing partner at Procinea Management LLC, Emeryville, California. Without implicating them in any errors and omissions, we thank our colleagues at Procinea Management who have collaborated at all stages in this research project and who have made truly significant contributions to this study. We would particularly like to thank Adam Cao, Paul Jung, Mark Rozells, Bob Rubin, and Scott Ryles. In addition, Gordon Rausser and Bill Balson directed the initial stages in the project; without their leadership we would have traveled neither as far nor as fast. An earlier version of this article was presented at the spring seminar of the Institute for Quantitative Research in Finance, Palm Beach, Florida, April 4, 2006.

from investing in more conventional securities. First, there are fewer investment opportunities in the movie industry. Each year the major U.S. studios release just over one hundred movies that might be candidates for investment, compared to several thousand common stocks listed on the U.S. exchanges. Second, movie projects are far less transparent than listed equities. No regulatory organization mandates financial disclosure at the project level, so the nature of the investment opportunities is unclear even in retrospect. Third, there is a very large body of research and a sound theoretical basis underlying the analysis of public securities. Despite a great deal of academic work, there is no generally accepted model relating movie revenue to any particular set of factors. Finally, each movie is an individual project which offers the investor a single decision: whether or not to commit at the project's inception. After this decision is made, the investment is illiquid and there is little opportunity to revise the holding. Stock markets offer nearly continuous pricing and liquidity, providing the equity investor with a range of entry and exit points and the ability to mark portfolios to market.

We analyze the revenue streams from major feature-length films in order to construct a revenue-forecasting model and to determine whether it can be the basis for a successful active strategy. It is well known that the revenue generated from filmed content has been growing steadily for decades. In fact, since 1990, nominal revenue growth has exceeded 6.5% per year, and aggregate revenues are now approaching $100 billion annually. In the past ten years, this growth has accelerated, driven by the geographic expansion of audience and by new technologies that allow each title to be exploited through additional channels. Although domestic box office results are the most well known measure of a film's appeal, today domestic box office is only a minor component of revenue. The revenue earned from exhibition of movies in the home, whether by DVD, video, or pay-per-view, has become significantly larger than the revenue from theatrical exhibition. In addition, the movie industry has become global. Box office revenues and DVD and pay TV penetration continue to grow internationally.

At the same time that revenues have been increasing, so too have the costs of movie production. Hollywood has responded by seeking outside financing partners to share risk. The six major studios typically release in total 100 to 120 titles per year with average production and domestic distribution costs of approximately $100 million each. This cost translates to an annual capital requirement of approximately $10 to $12 billion.[2] Rising production costs and marketing (usually referred to as prints and advertising, or P&A) expenses, coupled with capital constraints placed on the studios by their parent companies, are among the factors driving the ongoing need for capital.

Historically a large percentage of films have not generated revenue greater than the cost of production and distribution. Defining return on investment (ROI) as revenue in excess of costs, divided by cost, we estimate that while the average ROI for all films in recent years has been positive, the majority of films yielded negative returns. Our research attempts to apply quantitative modeling to improve this performance. This paper describes a multi-factor model based on audience appeal that forecasts revenue for individual movies, defining an appropriate capital investment for each title. An investor able to avoid movies whose costs exceed their discounted expected revenues could achieve a risk-adjusted return significantly better than that of the asset class.

[2] Production and domestic distribution cost averages are taken from the Motion Picture Association (2006b). We estimate that accounting for foreign distribution costs would increase the annual capital requirement to approximately $12 to $14 billion.

In a private market such as movie investment it is not sufficient to have an asset selection process. Equally important is a structure and contracting process to ensure that the interests of investors and the studios are aligned. We will not cover these aspects here, except to state our belief that investment in film projects should occur only before principal photography begins. At this time, the studio has the least information advantage relative to the investor and adverse selection is minimized.[3]

In the next section, we provide a brief review of the existing literature. Subsequent sections discuss our movie database, the financial characteristics of filmed content, our approach to modeling, and the performance of movies selected according to these ideas.

Prior Literature

There is a substantial academic literature on the movie industry, undertaken typically in management science, marketing, or economics departments. The importance of extreme and idiosyncratic events to the industry's overall performance is widely recognized. "Ten percent of films generate 50 percent of the box office," according to Vogel (2004). Correlations among movie projects, if they are significant at all, are likely to be negative as a result of competition for audiences. Although defining a movie index is somewhat problematic, Chance, Hillebrand, and Hilliard (2005) report that movies show negligible correlation with the S&P 500.

Much of the academic literature analyzes a movie's performance using factors

[3] Goettler and Leslie (2004) compare co-financed movies with those in which the studio retains full rights. They find no significant difference in return on investment or in risk, suggesting that studios are unable to predict performance before the start of production.

that can be known only near or after its release date. For example, Eliashberg *et al.* (2000) predict revenue and the effect of changes in advertising spending based on surveys of potential moviegoers who have viewed trailers. Elberse and Eliashberg (2003) model the dynamics of a movie already in release based on its number of screens and revenue in prior weeks, as well as other factors. Ainslie, Drèze, and Zufryden (2005) and Einav (2006) estimate self-consistent models in which movies compete for market share. Such models can provide insights, but they would be of limited usefulness to an investor evaluating a movie prior to production because they depend crucially on inputs not known at that time. For example, studies which include the number of screens generally find it to be the variable with the most explanatory power.[4] Although opening screens could in principle be contractually fixed in pre-production, in reality this variable is determined after the quality of the finished movie can be estimated with some confidence, as described by Hayes and Bing (2004). Using the number of screens in forecasting would amount to "snooping ahead" to the outcome of the production process. Advertising expenditures correlate with revenue but are disallowed to us for the same reason.[5]

Some academic work attempts to forecast performance using only fundamental characteristics available as the project goes into production. Simonoff and Sparrow (2000) study domestic box office revenue for 311 movies released in the United States during 1998 and find significant effects for genre, rating, whether the movie was released

---

[4] Results which are dominated by the number of screens include those of Neelamegham and Chintagunta (1999), Walls (2005), and Sharda and Delen (2006).

[5] Positive correlation between P&A expenses and movie revenue has been demonstrated by many researchers; see Elberse and Eliashberg (2003) and citations therein. Advertising creates demand and also reflects movie quality. Vanderhart and Wiggins (2004) and Elberse and Anand (2005) distinguish these effects.

in summer, and whether it contained actors with a history of large revenues. The presence of stars is positively correlated with production budget; Simonoff and Sparrow omit the latter factor, which we believe overstates the importance of the former. Hsu (2005) investigates the effect of genre on appeal, as measured by the number and favorableness of online reviews. She finds that the more genres a story attempts to combine, the greater the number of moviegoers who will be interested in it, but the less enthusiasm each one is likely to exhibit.

No consensus exists on which factors might influence movie success or how they should be measured. According to De Vany (2004), "Most movies are unprofitable. Large budgets and movie stars do not guarantee success. Even a sequel to a successful movie may be a flop." After studying budget, stars, sequels, genre, rating, screens, box office life, and year of release, he concluded, "There are no formulas for success in Hollywood." Reviews of the factors considered by various workers appear in Elberse and Eliashberg (2003) and in Terry, Butler, and De'Armond (2005). Sawhney and Eliashberg (1996) find greater revenue for sequels, critically acclaimed movies, and those with major stars; however, without controlling for production budget, the effect of stars could merely be a proxy for other costly factors of production. Ravid (1999) summarizes various studies on the influence of individual factors on movie financial performance including stars, sequels, rating, and budget. He finds that stars have a positive univariate effect on film profitability, which however is accounted for by covariance with production budget (i.e., spending more money explains variability in film revenues whether the spending is on stars or on other factors).

The academic literature uses a wide variety of dependent variables to quantify success: dollar revenue, logarithmic revenue, dollar profit, return on investment, and discrete functions of these. Results from different researchers therefore can be difficult to compare. With the notable exception of Ravid (1999), the only revenue measure used in most studies is the increasingly unimportant domestic box office. In contrast, we study total revenue from all channels. Surowiecki (2005) gives examples of movies whose performance hinged on DVD revenues, and foreign box office can be equally decisive.

Quantitative research on the determinants of movie success has generally employed linear models. However, evident nonlinearities and interactions among factors have inspired other techniques, such as the neural network of Sharda and Delen (2006). Our work bridges the gap between these approaches. We use a kernel estimator (locally weighted regression) that is responsive to nonlinearities and interactions but, unlike neural networks and other machine learning techniques, provides interpretable coefficients for each factor and confidence intervals at every point.

It is common wisdom among experts in the movie industry that revenue is unpredictable. When Goldman (1983) wrote “Nobody knows anything,” he coined one of Hollywood’s most repeated aphorisms. Academic researchers tend to interpret their results in this atmosphere of pessimism. Swami, Eliashberg, and Weinberg (1999) built a revenue model based on genre, rating, and indicator variables for sequel, major star, and major distributor, and described its $R^2 = 0.28$ as “relatively low.” De Vany (2004) modeled dollar profits on a historical dataset of 2,015 films, concluding, “The equation is a very poor fit, with an R-squared value of just 0.118.” However, in comparison to quantitative equity investing, these models show great explanatory power. An information coefficient IC above 0.2, so high that it “usually signals a faulty backtest or imminent investigation for insider dealing" according to Grinold and Kahn (2000), corresponds to only $R^2 \geq 0.04$. An investment strategy based on movies would have a breadth BR, equal to the number of independent bets per year, of roughly one hundred. Thus despite the high information coefficient, the information ratio predicted by Grinold's fundamental law of active management, $\mathrm{IR} = \mathrm{IC}\sqrt{\mathrm{BR}}$, for movie investment would be comparable to that available from quantitative equity strategies.

Description of Data[6]

Attributes and Financials. We have collected movie information from a variety of third parties; we also obtain directly some data which are not generally available. Our sources include industry standard vendors such as Nielsen EDI, Adams Media Research, and Box Office Mojo. These data are supplemented with both free and fee-based online sources, including entertainment industry publications, media research reports, and equity analyst reports. Procinea has also defined and compiled movie attributes (e.g., cast billed order, story elements) not provided by vendor or industry sources but which we believe add value in the modeling process. Collectively, these sources provide financials (e.g., box office, production cost) and attributes (e.g., rating, genre, cast) for movies released theatrically in the North American market since 1980.[7] Our database has roughly 70 data items for each film and covers more than 7,800 films.

Compiling a database from multiple sources allows us to validate and clean many data items. For example, we have summed our video revenue data by format and verified

[6] We thank Adam Cao, who was responsible for building Procinea's movie database, for his contribution to writing this section.
[7] In Hollywood parlance, North America consists of the United States and Canada.

that the totals match published industry aggregates, and that they accurately reflect the decline in VHS sales and the rise of the DVD. We compare values for individual films across different sources and flag discrepancies for manual investigation and correction. For missing data, we estimate values based on movies with available data, research publications, and insights from industry insiders. As a result of these efforts, we believe our database is more accurate and has wider coverage than any single available source.

With the exception of box office gross, which is reported daily, most movie financial data are available only as aggregated amounts. However, film revenues are earned and costs are paid across different windows spanning several years, e.g., production, post-production, theatrical distribution, home video, and television. A reasonable estimate of the timing of the cash flows in these various channels is necessary for valuation. We have therefore reconciled a number of public and proprietary sources to produce a timing grid that specifies the initial month for each channel, the number of months comprising each channel, and the fraction of total revenue or cost expected in each month.

Investing before the commencement of principal photography raises several important issues of timing and data availability. By and large, public and commercially available databases contain substantial information on movies only as of their release date or later. Accordingly, we assume that the data captured and reported are sufficiently accurate and similar to the movie characteristics we model before production begins. When backtesting our model, we must be careful to include appropriate lags so that decisions are based only on data available at the time.

Estimation Universe. As in quantitative equity investing, our model is fit to an estimation universe smaller than the broad universe of assets, to exclude unrepresentative movies, but larger than the investable (target) universe, for a more robust fit. We define a *broad universe* consisting of all movies with a production cost of $2m or more which were released theatrically in North America, excluding only unrated and foreign productions. The broad universe contains 1,627 movies released from 1997 to 2004, the sample period we have chosen for this paper. The broad universe contains movies which would not be considered “institutional” in character. Accordingly, the *full estimation universe,* which contains 1,394 movies over the same period, excludes animation, sequels, documentaries, and NC-17 ratings. Studios are less likely to seek co-financing for sequels and franchises.[8] Animated films should be modeled separately because the relevant attributes for these projects are probably different from live-action films.[9]

The regular *estimation universe*, which we use when production cost is not an explicit factor in the model being fit, is the subset of the full estimation universe whose costs lie in the range from $15m to $125m. It contains 836 movies. Finally, the *target universe* is the subset of the estimation universe financed (or co-financed) and distributed by a major studio, including specialty subsidiaries. The target universe includes 588 movies released 1997-2004.

The Market for Filmed Entertainment

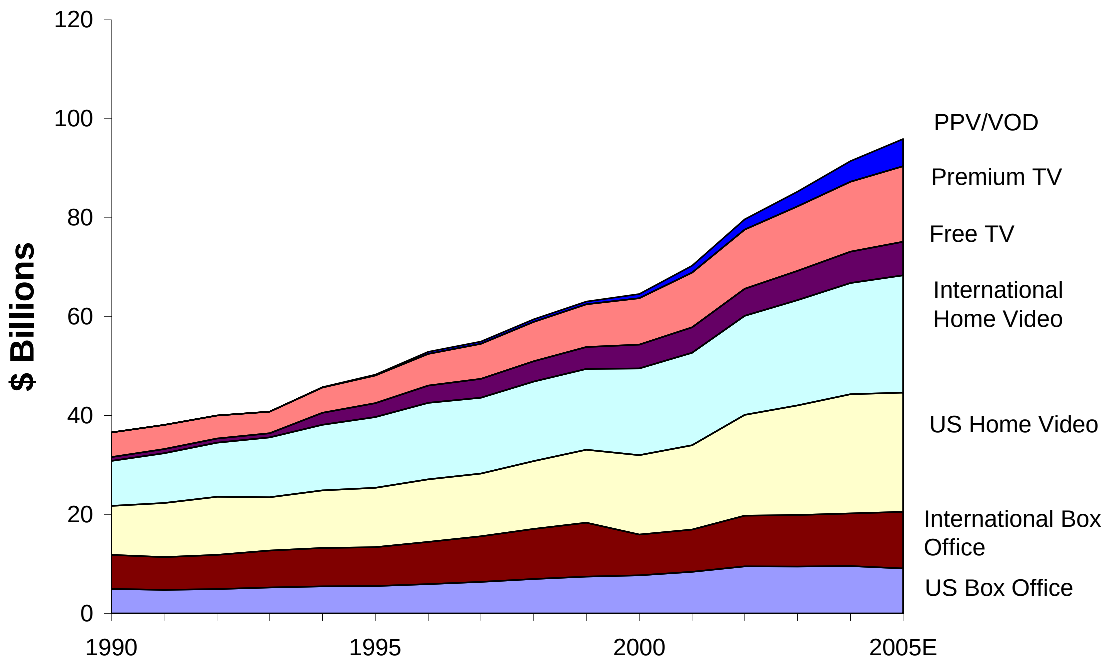
Revenue Growth. As described above, revenue from filmed content has been

---

[8] It is common wisdom that sequels and franchise films are very profitable. This causes the mean ROI of our target universe to be slightly less than that of a larger universe including these movies.
[9] For example, the impact of the cast is likely to be different when it provides only voice talent, and the typical production cycle is much longer for an animated film.

steadily growing. Nardone (1982) presents evidence that this growth is counter-cyclical as consumers see more movies during weaker economic periods as a substitute for more expensive leisure spending. At the consumer level, the market is currently approaching $100 billion in annual revenues on a global basis.

Figure 1. Worldwide Filmed Entertainment Revenue by Window



Source: Company Reports, JP Morgan and Procinea Estimates

In order to maximize revenues and profits, distributors make films available for exclusive periods of time (commonly referred to as windows) in each of the various distribution channels following initial release. Historically, this strategy has allowed distributors to minimize cannibalization of existing revenues and to continue steady revenue growth despite a maturing market at the box office. Home exhibition of movies, including DVD, video, and television (pay-per-view, cable, and free TV), has in recent years become a significantly larger source of revenue than theatrical exhibition as seen in

Figure 1. Home exhibition revenues are now almost three times theatrical exhibition revenues. Since its introduction in 1997, DVD penetration has increased to almost 80% of U.S. households, representing the fastest adoption of consumer technology in history and changing the economics of feature film exploitation. Video-on-demand (VOD) and other emerging digital distribution technologies are expected to create additional opportunities to exploit filmed entertainment content.

In addition to the opening of new windows, revenue growth has been driven by the geographic expansion of the audience base. Our data show that box office revenues from outside North America now exceed domestic box office for roughly half of movies, and foreign box office is growing at twice the domestic rate. Foreign DVD and pay TV penetration are behind the U.S. and should drive international growth as they catch up to North American levels.

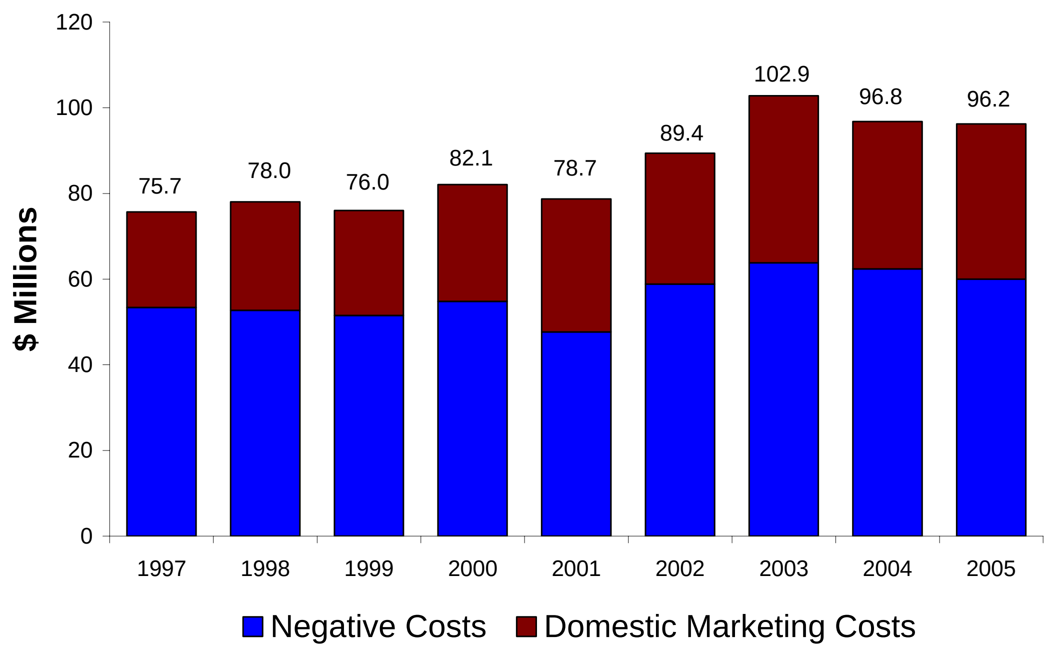
Despite the recent media focus on disappointing box office results and slowing DVD sales, many analysts believe the fundamentals for continued revenue growth for filmed content remain sound. The drivers are generally accepted to be continuing growth internationally and the emergence of new technology and distribution channels which enhance opportunities to exploit content. The compression of windows, causing a title to move more quickly from one channel to the next, should lead to only slight cannibalization of revenue and may have benefits in marketing the movie.[10]

Ongoing Need for Risk Sharing Capital. Hollywood has a long history of seeking outside financing partners to share risk and sustain the number of titles produced. The necessary capital is raised through foreign territory pre-sales, tax driven

[10] Wang, Blackledge, and Chew (2005) describe the movie industry from the perspective of an institutional equity analyst.

structures, and co-finance and co-production structures on a single or multi-picture basis. Eisbruck (2005) reviews the securitization of movie receivables. However, there has been no efficient, flexible, and reliable source of risk sharing capital to the studios. As shown in Figure 2, the average production and domestic distribution cost of a feature film from one of the major studios (Disney, Fox, Paramount, Sony, Universal, Warner, and their subsidiary brands) has been trending upwards and now stands at roughly $100 million. The 100 to 125 titles released by these studios each year create $10 to $12 billion in annual funding needs.

Figure 2. Average Cost of Major Studio Feature Films for Films Released in 1997 - 2005

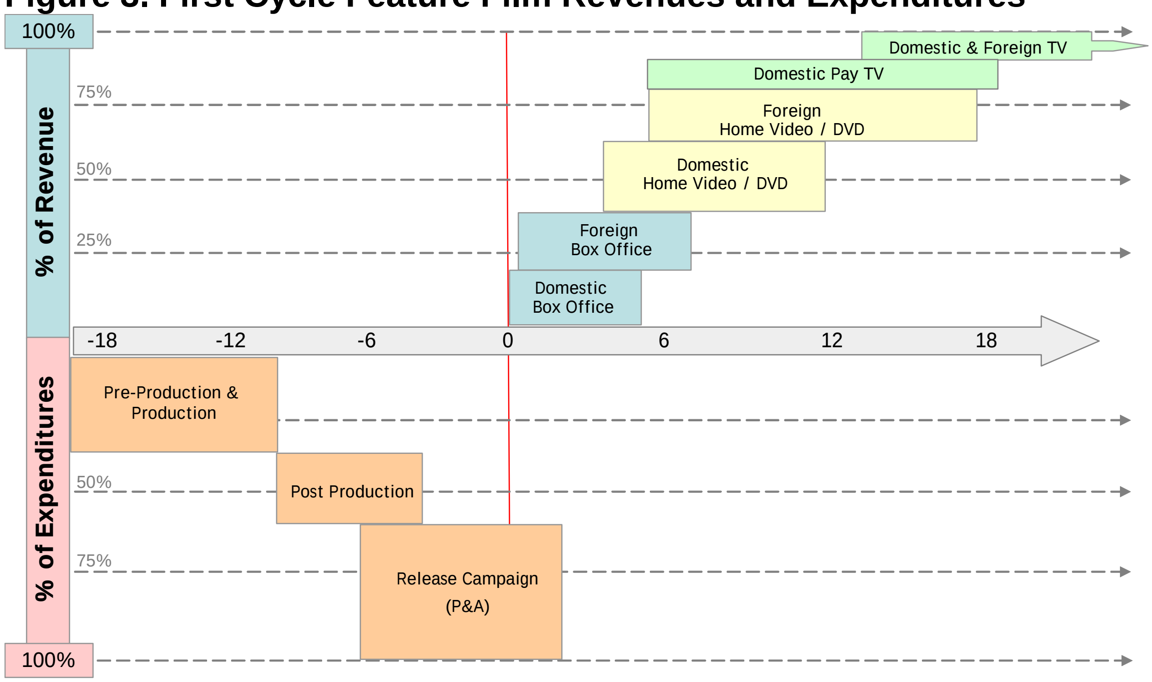


Source: Motion Picture Association of America

In recent years, a combination of developments has increased the "financing gap" in Hollywood. While both production and distribution costs have been rising, the supply of available investment capital has been constrained. Industry consolidation has reduced the number of studios (major, mini-major, and independent) and resulted in all of the major studios being owned by large media conglomerates. These corporate parents have not increased the amount of internal capital available to their studio subsidiaries commensurate with rising costs and an increase in the number of films produced. In addition, the availability of tax-driven financing has continued to decline over time.

Potential investors may be wary of the industry based, upon widespread publicity surrounding apparent lack of transparency in studio financial behavior, unsuccessful attempts by previous investors, and the perceived causes of these failures. Rightly or wrongly, investors fear that motion picture performance is wholly unpredictable, that studios saddle investors with "losers" (which would seem to contradict the prior concern), that participation accounting strongly favors the studio, and that investors must bear studio error, bad judgment, and profligate spending.

Figure 3. First Cycle Feature Film Revenues and Expenditures



Source: Procinea Management, LLC

Short Duration Cash Flows. Feature films have unusual cash flow characteristics. Typically, investment in production and distribution takes place over a 12- to 18-month period after which the film is released and exploited through different channels of distribution which include theatrical exhibition, home video, pay-per-view TV, premium TV, and free TV. As shown in Figure 3, the sequential exploitation of a film for the first time in all channels and in all markets is known as the first cycle and is completed in approximately seven years. Industry estimates suggest that on average roughly 60% of first cycle revenue is generated within 12 months of release, and 80% is generated within 24 months.[11] In contrast to other alternative investments such as private equity, feature film investments have a shorter duration and no realization risk tied to capital markets, the M&A environment, or other exit opportunities.

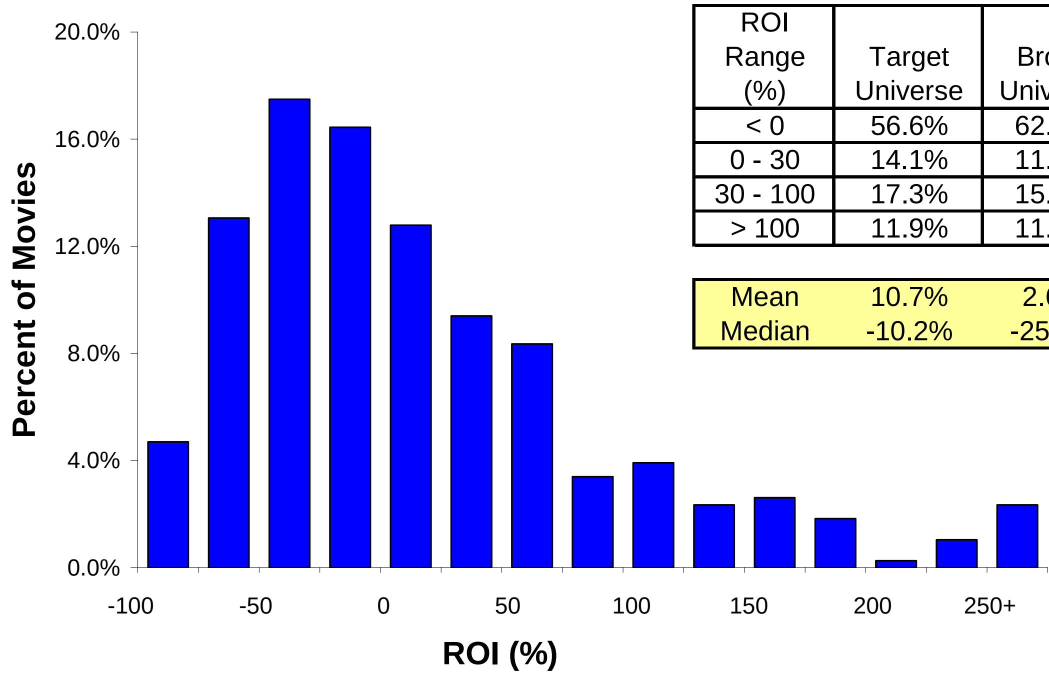
Overview of Financial Performance of Feature Films. Financing the production and distribution of films is risky. The financial performance of a film can be measured by comparing the estimated net revenue received from all windows to the estimated investment in the film, defined as production costs plus P&A expenses. To assess the economics from investing directly in films, consider the broad universe of North American releases costing $2m or more. Based upon first cycle revenues only, a large percentage of films in the broad universe do not generate a positive ROI. While the overall ROI for films released from 1997 to 2004 has been positive and improved since 2000, we estimate that 62% of the films generated negative ROI and an even higher percentage failed to generate adequate returns on capital given the risk inherent in film investing.

---

[11] These estimates are taken from marketing material produced by JPMorgan.

Figure 4 shows the distribution of ROI for films released from 1997 to 2004. The equal weighted mean ROI for the target universe is approximately 11% compared to 3% for the broad universe. The distribution is also highly skewed as evidenced by a median ROI of negative 10% for the target universe versus the positive 11% mean for the same universe of films. The results for the broad universe suggest more variability of returns. The variance of film returns and the strong positive skewness of the distribution explain why studios and investors use a portfolio approach to reduce risk, increasing the probability of capturing enough big winners to offset the losers.

Figure 4. Distribution of Target Universe ROI for Films Released from 1997 to 2004



ROI Range (%)	Target Universe	Broad Universe
< 0	56.6%	62.3%
0 - 30	14.1%	11.1%
30 - 100	17.3%	15.1%
> 100	11.9%	11.5%

Mean	10.7%	2.6%
Median	-10.2%	-25.0%

ROI = (Total Revenue / Total Cost) - 1
Movies released from 1997 to 2004

Behavioral Idiosyncrasies of the Industry. We believe the production, marketing, and distribution decisions for a given film project are influenced by a number of business considerations and economic biases that may have little to do with the financial merits of that particular project. Such considerations include cultivation of relationships, demand for content in a distribution channel, perceived artistic value, availability, and market share. Although these apparently non-economic influences may in fact bring long-term performance benefits to the studio, they can impact the returns available to investors in an individual film project and must be taken into consideration in constructing an investment portfolio. Ravid (2004) explains several Hollywood anomalies, such as the overproduction of R-rated movies, as agency problems resulting from studio executives' risk aversion.

Of particular interest is the relationship between creative management and commercial management at the studios. We observe these groups to have different definitions of "success" for a movie. While the commercial management appears to pursue traditional profitability measures, the creative management focuses on artistic value and appears driven more towards achieving critical acclaim. Our belief is that this difference in focus could lead in some instances to inefficient decision making and potentially systematic bias at the studios.

Changing Nature of the Asset Class. The industry is undergoing a number of important changes, which may impact the financial returns for feature films. As noted above, new technologies have created new revenue streams. In addition, new technologies and changing dynamics of the industry have driven margin improvements. For example, the lower production cost of a DVD has improved profit margins as video has migrated from VHS to DVD. Also, the transition from video rental to the sell-through model has improved margins. Emerging digital technologies such as VOD are expected to enhance margins further by lowering distribution costs. In the longer term, the industry can also realize cost benefits through digital distribution to theaters of new releases which can eliminate the need to manufacture and transport film prints.

The industry continues to experience window compression as films generally spend less time in theatrical exhibition and are released in the DVD market sooner. As noted above, some analysts believe window compression should be generally positive for industry economics. It is true that window compression may cannibalize box office revenue, but it also accelerates revenue streams. In addition, window compression allows the distributor to leverage the consumer awareness generated by marketing at theatrical release, spending on which has grown substantially in recent years, and which can result in higher DVD sell-through or less marketing expense during the DVD window.

Forecasting Movie Revenues

Model Development. This section describes the model we have built to forecast movie revenue. We have chosen total dollar revenue as the quantity to be modeled, rather than return on investment or a particular component of revenue such as domestic box office. There are two reasons for this, both consequences of the strategy that the model is designed to support. First, we assume that investors will contract with studios for the same share of revenue from all channels, so total revenue is the appropriate quantity for valuation. Second, investors will generally contract for the same share of cost and of revenue (most frequently 50%) for each movie. This means that an investor's portfolio of movies will be approximately cost weighted.[12] Any model's error in

[12] The movie portfolio is analogous to a long-only equity portfolio that implements alphas by holding each asset either at index weight or not at all. In the terminology of Clarke, de Silva, and Thorley (2002), this suboptimal portfolio construction will introduce a transfer coefficient less than unity in the fundamental

predicting the portfolio revenue is proportional to the sum of the dollar errors for each movie, and therefore the model should be designed to minimize the forecast error relative to dollar revenue.

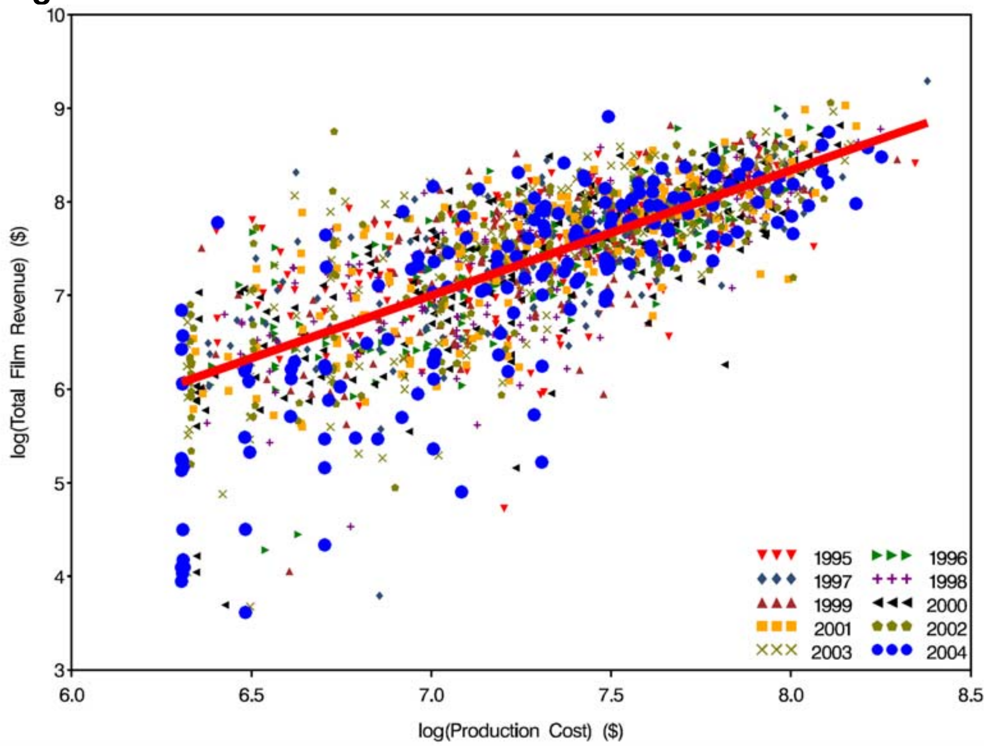
Our research proceeds in three stages. First, we identify economically plausible factors. Second, we construct appropriate measures of these factors. Third, we price them in a multi-factor model. The result is a revenue forecasting model which can identify films that are more likely to earn a positive return on investment and thereby improve the returns from the asset class. The model is based upon the assumption that similar movies should, on average, generate similar revenue streams. We acknowledge that it is not possible to identify and model all the relevant dimensions of similarity; however, any missing factors can be addressed through portfolio diversification. The approach, therefore, is to build a model which relates the revenues of movies to a set of their attributes. This reduces the problem of forecasting the revenue for a new movie project to first identifying its attributes and then pricing these attributes by analyzing the performance of other movies with similar attributes.

We will describe our model one factor at a time to illustrate our research process and show how nonlinearities and interactions between the factors are handled. A good starting point is production cost. Figure 5 plots total revenue as a function of production cost for movies released from 1995 to 2004 in the full estimation universe. Revenue is as reported through March 2005, and all amounts are in constant dollars as of that date. The red line is an ordinary least squares fit to the data. Its slope is slightly greater than unity, indicating that the relationship between revenue and production cost is described by a

law of active management.

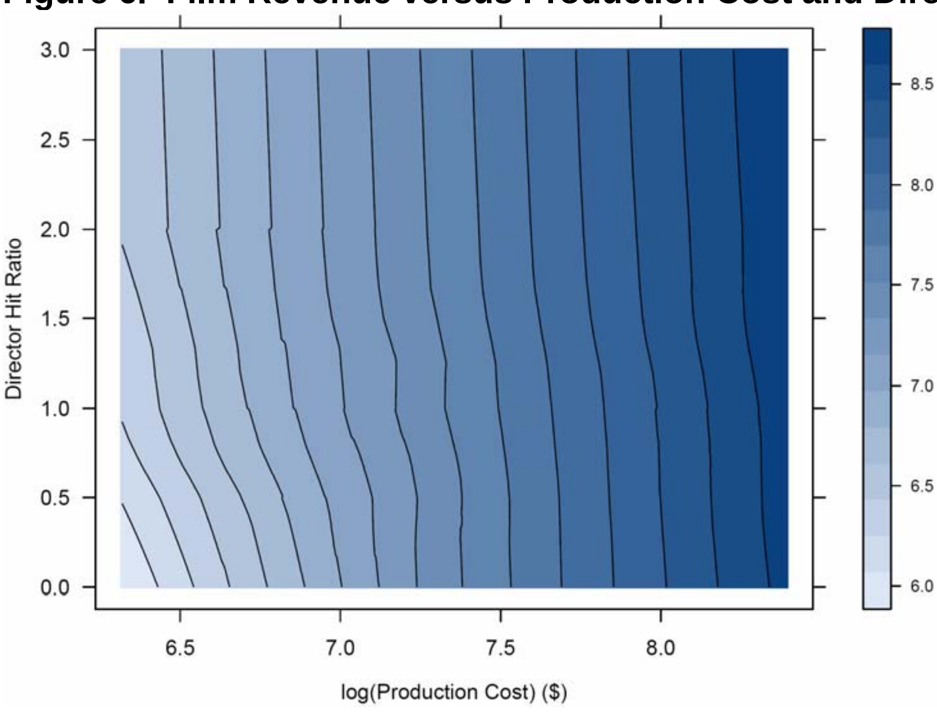
convex power law. The relationship is generally stable from year to year, although a disproportionate number of negative residuals appear in 2004 for movies which have not yet had their video release. This artifact underscores the importance of DVD revenues, which must be handled carefully in model estimation. Note also the obvious heteroskedasticity in Figure 5. Smaller movies have more volatile revenues than larger ones. We will return to this issue shortly.

Figure 5. Film Revenue versus Production Cost



Although the relationship between revenue and production cost is clearly significant, it could result from a number of causes. Movies with higher production budgets also enjoy better-known actors, more talented directors, higher advertising expenditures, and so forth. We proceed to disentangle these effects by introducing a measure of director talent we call the director hit ratio D_i. The director hit ratio is a dimensionless quantity in the range from 0 to 3 calculated from the financial performance of the director's past movies with imputation rules for first-time directors. It is designed to capture a director's potential for commercial success. The correlation between production cost and director hit ratio over the full estimation universe is 0.34. Linear regression shows that production cost, director hit ratio, and their multiplicative interaction are simultaneously significant predictors of revenue. However, there is no reason to believe that such a linear model best captures the interaction between the director's abilities and the financial resources available. We develop a more flexible specification using the non-parametric technique of locally weighted regression, which is described in more detail in the Technical Appendix. The essential idea is that the forecast revenue at any point is obtained from a regression that gives greater weight to nearby movies. The forecast is not constrained to obey a linear relationship across the entire range of data.

Figure 6. Film Revenue versus Production Cost and Director Skill

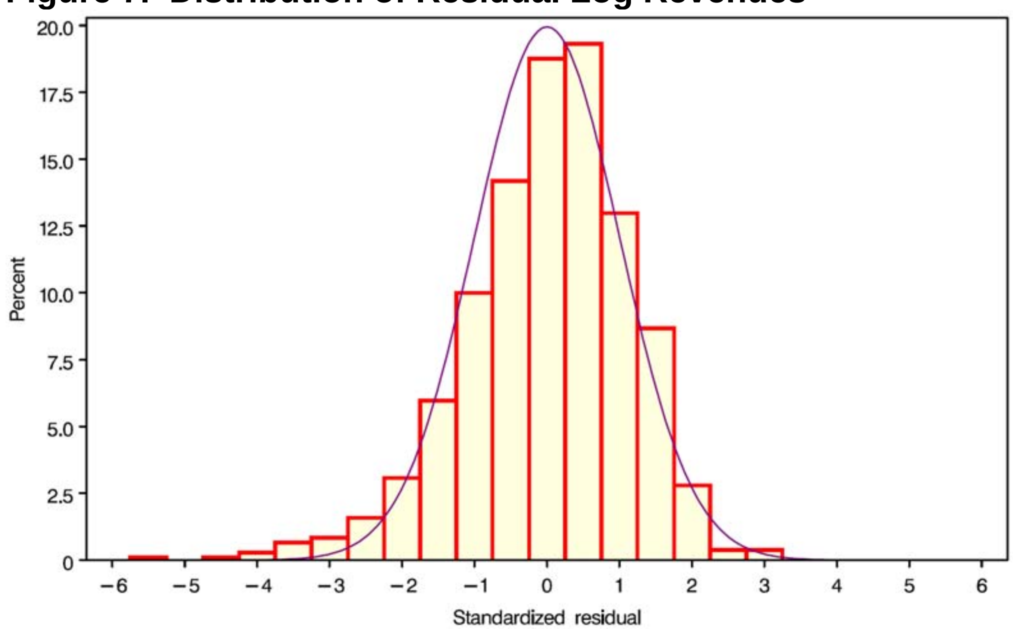


The results of our two-factor locally weighted regression are illustrated in

Figure 6. The contour plot shows forecast total dollar revenue on a base-10 logarithmic scale as a function of production cost and director hit ratio.[13] As expected, revenue increases with either increasing production cost or increasing director skill. However, the effects are neither linear nor independent; if they were, the contour lines would be parallel and straight. Performance of movies in the lower left corner (low budgets and poor directors) is worse than expected from a linear extrapolation of either effect alone. Interactions between factors can reveal a class of movies to be avoided.

Let η_i denote residual log revenue, the difference between $\log_{10}(R_i)$ and the fitted surface in Figure 6. As noted previously, this residual displays lower variance at higher production costs. A univariate local fit to squared residuals gives us an estimate of the cost-dependent standard deviation $\sigma(C)$. This is a simple risk model for movies.

Figure 7. Distribution of Residual Log Revenues



[13] The fit is done releases from 1995 to 2004 as reported through September 2005. This later reporting date minimizes the DVD artifact seen in Figure 5. The shape of the fitted surface is robust to changes in the reporting date and to the exclusion of movies that have not been released on DVD.

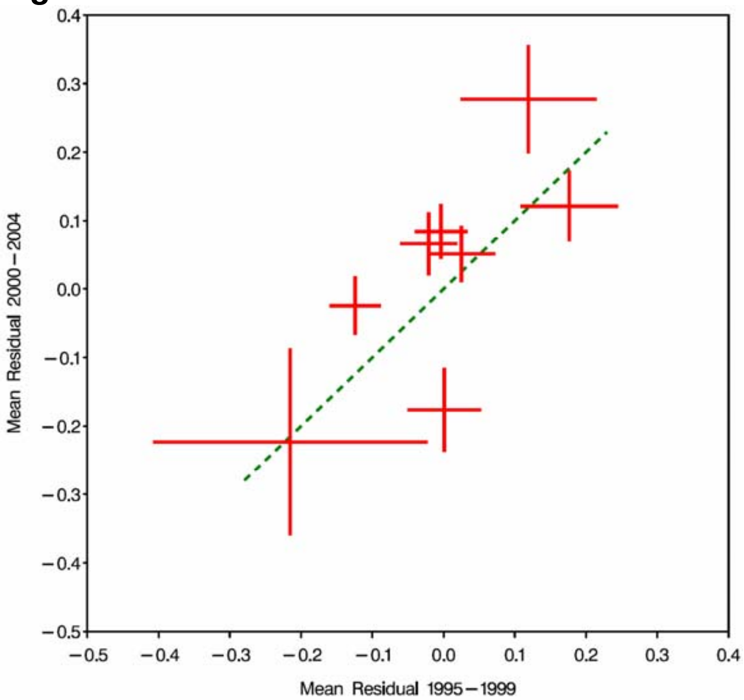
Figure 7 plots the in-sample distribution of the standardized residual $\eta_i / \sigma(C)$. The smooth curve is a standard normal distribution with no adjustable parameters. We can draw two important conclusions from this plot. First, although De Vany (2004) and Walls (2005) have fit dollar revenue to a Lévy distribution with infinite variance, simple models of mean and variance can yield a good fit to a lognormal distribution. This will be very important when we wish to convert our logarithmic forecasts to dollar forecasts, because the mean of a lognormal distribution depends on both its location and scale parameters. The second conclusion we can draw from Figure 7 is that outliers are sufficiently controlled to allow us to use these residuals in further estimation. We now continue our investigation of factors relevant to movie revenue.

Table 1. Average Residual Revenue and Fraction of Movies by Rating and Season

	Summer	Holiday	Off-Season	Full Year
A. 1995-1999				
G, PG, and PG-13	0.21	0.01	-0.03	0.04
	13%	10%	26%	49%
R	0.05	0.00	-0.08	-0.04
	9%	10%	33%	51%
B. 2000-2004				
G, PG, and PG-13	0.17	0.06	0.01	0.06
	15%	13%	34%	63%
R	0.02	-0.07	-0.13	-0.10
	5%	7%	25%	37%

Rating and season are two additional economically plausible factors. They are correlated with thematic content and with accessibility: young audiences have more free time in the summer, and are barred from R-rated movies. Table 1 shows the percentage of movies and the average residual η_i in six rating-season categories. The effect of both variables is significant at the 95% level but their interaction is not. Although the strength of the effects has changed over time, their relative magnitudes persist from the early period (1995-1999) to the late period (2000-2004). Summer is consistently the season with the highest revenues, and R ratings underperform in all seasons. Henceforth we subtract the rating effect from the residual η_i but not the season effect, since the latter may proxy for other factors we have yet to study.

Figure 8. Consistent Abnormal Revenue of Genres



Nielsen classifies movies into one of twelve primary genres.[14] We combine these in intuitively reasonable clusters (action with adventure, for example) until no category contains fewer than 30 movies in our combined period. This results in eight categories. The average residual revenue in each of these genre categories for two time periods is plotted in Figure 8. The dashed line is the diagonal on which a genre would fall if its

[14] Nielsen's primary genres are action, adventure, black comedy, comedy, drama, fantasy, horror, musical, romantic comedy, science fiction, thriller, and western. Genre correlates with season in unsurprising ways. For example, action movies tend to be released in the summer and dramas in the fall.

average residual were the same in both periods. Analysis of variance indicates that genre has significant explanatory power in each period. Furthermore, the correlation of the genre effect between the early period and the late period is 0.75 with a *p*-value of 0.03. Genre has a significant and persistent effect on revenues, an obvious inefficiency that one might expect to be arbitraged away. With far more projects available in development than could be produced in a year, why don't studios increase production of high-revenue genres to satisfy the evident demand? One possible explanation is the externalities that low-revenue genres provide to the studio; these tend to be more serious projects which cultivate relationships with actors or directors, or bring the studio awards and other indicators of status.

Table 2. Residual Revenue as a Function of Metascore

	Intercept	Metascore	Adjusted R^2
A. Raw Metascore			
Mean		49.5	
Standard deviation		17.0	
Parameter estimate	-0.275	0.00665	
t-Statistic	-5.98	7.57	6.9%
B. Metascore residual to genre			
Mean		0.0	
Standard deviation		16.4	
Parameter estimate	0.054	0.00843	
t-Statistic	3.70	9.40	10.3%

We would next like to examine factors which measure the quality of a movie's story. Before doing so, however, we seek a proof of concept to indicate whether this effort is likely to be worthwhile. The number of critically acclaimed box office failures might initially cause some skepticism, although Sawhney and Eliashberg (1996) have found a favorable effect of positive reviews. The website metacritic.com aggregates reviews from influential critics and assigns each movie a numerical "Metascore" from 0 to 100. Higher numbers correspond to more positive reviews. Table 2 shows that the Metascore explains 6.9% of the variance in residual revenues, or 10.3% if the effect of genre is first removed by regressing on genre indicator variables. Critics tend to give high scores to dramas and low scores to horror movies, for example. Their opinions are most predictive of revenue after such biases are removed by genre neutralization, which we believe is analogous to industry neutralization of a value signal.[15] Although the Metascore would obviously not be available before production begins, and is therefore not useful for a revenue forecasting model, it demonstrates that audiences are willing to pay for quality.

Table 3. Residual Revenue as a Function of Story Elements

	Young Protagonist	Male Protagonist	Happy Ending	Adjusted R^2
A. Raw story elements				
Mean	0.179	0.771	0.791	
B. De-meaned story elements				
Parameter estimate	0.145	-0.086	0.067	3.0%
t-statistic	2.76	-1.82	1.37	
C. Story elements residual to genre				
Parameter estimate	0.165	-0.087	0.065	3.4%
t-statistic	3.08	-1.75	1.26	

Given that story quality appears to be rewarded, we have gathered a great deal of

[15] Whether critics influence moviegoers' preferences or merely reflect them is an interesting question. Eliashberg and Shugan (1997) showed a significant correlation between reviews and cumulative box office revenues, but none with early revenues. They conclude, "This finding suggests that critics…appear to act more as leading indicators than as opinion leaders." This supports our interpretation of Metascore as a proxy for movie quality.

survey data in an attempt to understand what makes a story appealing. We report here some preliminary results. According to the Motion Picture Association (2006a), the population of moviegoers includes significantly more young people and slightly more males than does the general population. Hypothesizing that people prefer uplifting stories about others like themselves, the variables we choose are the age (over or under 23) and sex of the protagonist and whether the movie has a happy ending.[16] The regression results in Table 3 show higher revenues for movies with young protagonists, female protagonists, and happy endings, although only the first of these is significant. The story variables are correlated with genre (e.g., horror movies are less likely to have happy endings) but neutralizing them to genre affects the results only slightly.

As part of our research process we have examined a number of additional variables, including producer, studio, writer, target demographic, run time, the presence and track record of stars, repeated collaborations between stars and directors, and the specialization of stars by genre. In all these cases, we have found that the in-sample effect was weak, that it did not persist out of sample, or that further research or data collection were necessary to test the idea decisively. A thorough examination of all of these factors is beyond the scope of the current paper.

At least as important as the factors we have chosen to use in our attempt to forecast movie revenue is the set of factors we have deliberately chosen not to use. Some, like an interaction between rating and season, are economically plausible but not

[16] Procinea employs film students to complete a survey on specified movies. We define a happy ending as one which meets all of the following criteria: The protagonist is still alive at the end of the film, the protagonist has achieved his or her goal, an obstacle presented at the beginning of the story has been completely overcome (e.g., villain receives comeuppance, lovers reunited, disaster survived), and all conflicts have been resolved in a way that leaves no obvious difficulty for the protagonist. Cross checking among analysts shows a reassuring level of agreement about which movies have happy endings and which do not.

supported by the data. Others have formal explanatory power but are incompletely determined before production begins. In the latter category we include marketing variables such as the number of opening screens and the advertising budget. Decisions about these variables are made only after much of the ambiguity in the production process has been resolved. The studio may not have seen the final cut, but it usually knows whether it has a flop on its hands.

With the goal of predicting total revenues, we have identified several movie attributes that best define appeal. Understanding these attributes allows us to forecast the distribution of revenues of a portfolio of well-defined feature films. Our model is calibrated based on a broad universe of films already produced, released, and distributed, for which the relevant attributes are known. We believe that these data items will be available before principal photography begins, which is the time when the informational asymmetry between the investors and the studio is minimized.

Accuracy and Selectivity Out of Sample. We now test the accuracy of our model's forecasts in a strict out-of-sample environment. To do this, we derive the parameters of the model as they would have appeared over a range of months in the past. Specifically, we estimate a version of the revenue forecasting model based on data available at each month end from September 1995 through August 2003. The structure of the model is constant over this period; only its parameters are updated as new information becomes available. The corresponding historical state of the model is used to evaluate movies which were released from January 1997 to December 2004, allowing a fairly conservative 15 months from the beginning of principal photography to release.

Figure 9. Distribution of Dollar Forecast Errors from Three Models

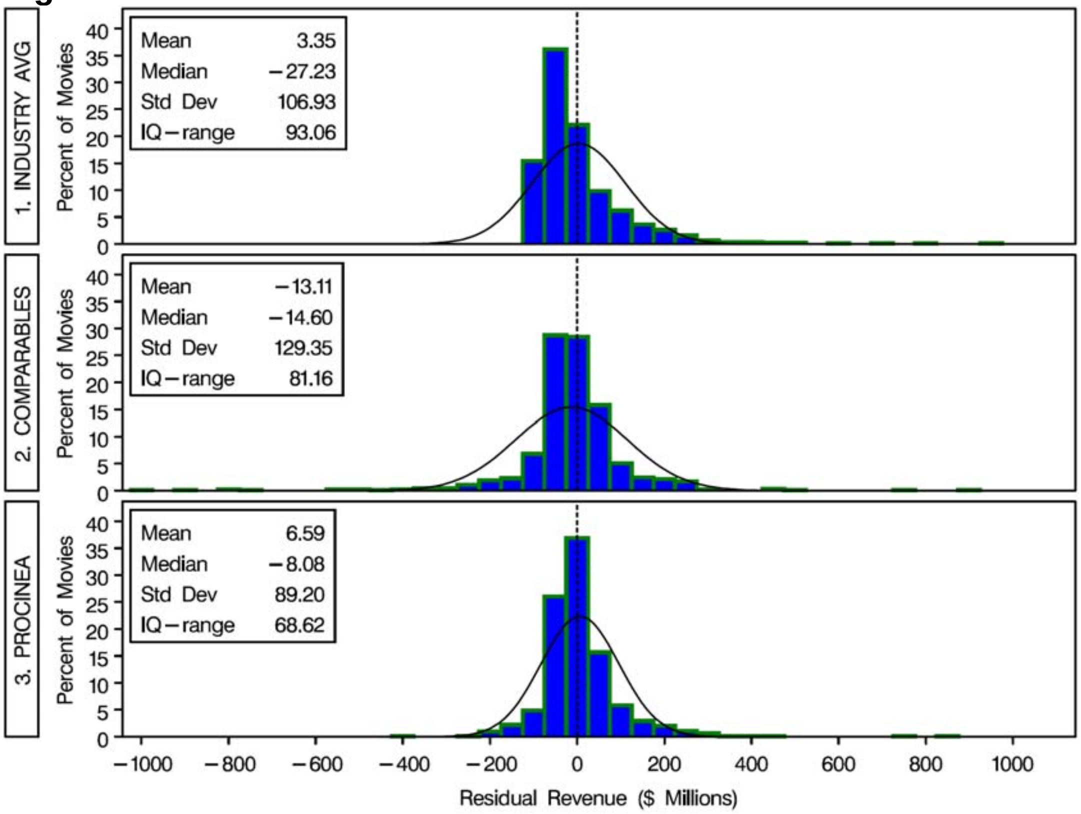
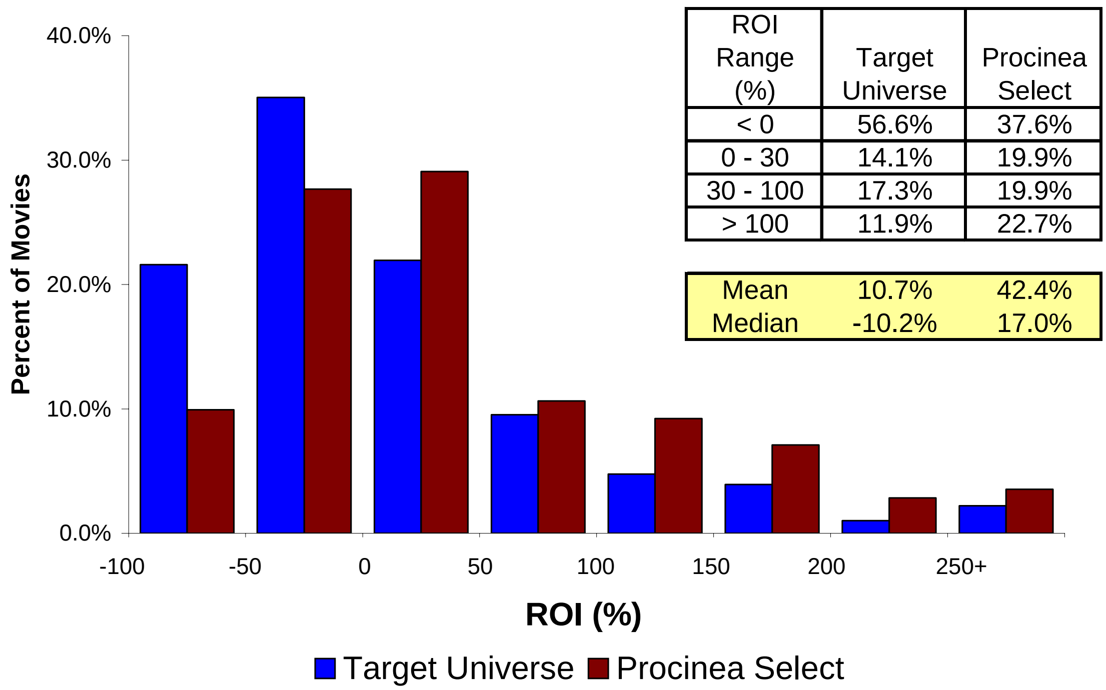


Figure 9 compares the accuracy of our model to two benchmarks. For each model, the distribution of the difference between the forecast and actual dollar revenues is shown over the estimation universe. The first benchmark is an exponentially weighted moving average of revenue to past movies in the estimation universe with a five-year half life. This "nobody knows anything" model assumes that no factors relevant to valuation distinguish one movie from another. The second model uses a clustering algorithm that groups movies into roughly 60 categories. The forecast for a new movie is its category average. This model is intended to approximate a studio's "running the numbers" to create a revenue forecast from comparable movies. All models are estimated out of sample with realistic lags, as described above. Our model is the most accurate of the three, measured by standard deviation or by inter-quartile range. Of course there are large outliers in all the distributions of dollar error because specific returns to movie assets are significant and long-tailed. Over the target universe the model we have developed has an out-of-sample R^2 of 31%.

Figure 10. Model Selectivity for Films Released from 1997 to 2004



ROI Range (%)	Target Universe	Procinea Select
< 0	56.6%	37.6%
0 - 30	14.1%	19.9%
30 - 100	17.3%	19.9%
> 100	11.9%	22.7%

Mean	10.7%	42.4%
Median	-10.2%	17.0%

We expect that revenue forecasts generated by the model could be used to estimate discounted cash flows for prospective film projects. We propose that an investor would seek to invest in projects whose discounted revenue stream exceeds the expected total of production and P&A costs. An example of the simulated results from such a strategy appears in Figure 10. We chose a hurdle rate of return that selects approximately one quarter of the movies in the target universe. The P&A expense used in the internal rate of return calculation is a function of production cost estimated from conversations with studio marketing executives, and the results are not sensitive to reasonable variations in its form. The mean return on investment of the selected movies is 42% as compared to 11% for the universe. The fraction of selected movies which fail to cover their costs is 38% as compared to 57% for the universe.

Summary

Why do we hope to succeed where studios fail? According to De Vany (2004), "There are a lot of failures and a few rare and unpredictable successes. Individuals tend to attribute causality improperly. They tend to attribute their successes to ability and their failures to bad luck." The skewed statistics of movie production make it easy to draw the wrong conclusion from success (repeat the formula) and to learn nothing from failure. With studios able to take only a few bets per year, and executives justifiably worried that one wrong decision will end a career, the opportunities for learning are restricted and the incentives to sacrifice return for comfort are strong. In this industry, a little objectivity goes a long way.

Technical Appendix

The core of the forecasting algorithm is a locally weighted regression of revenue on raw and derived attributes. This technique adapts to two effects seen in movie revenue: nonlinearity and interactions. As an example of the former, the extra revenue resulting from an increase in production cost from $20 million to $30 million need not be the same as that resulting from an increase from $100 million to $110 million. Interactions between factors can be as important predictors of revenue as the factors themselves. For example, star power might have a strong effect on revenue only when coupled with a powerful director.

The standard reference on locally weighted regression is Cleveland (1988). Consider a point (*c*, *D*) in the cost-director plane. Here, *c* is the base-10 logarithmic production cost and *D* is the director hit ratio. The Euclidean distance between the *i*th movie (c_i, D_i) and this point is $d_i = \left[(c_i - c)^2 + (D_i - D)^2\right]^{1/2}$. Given a smoothing parameter *s*, find the cutoff distance $d_{\max}$ such that the fraction of all movies with $d_i \le d_{\max}$ is equal to *s*. Assign a weight to each movie according to the tricubic function $w_i = \left[1 - (d_i / d_{\max})^3\right]^3$ if $d_i \le d_{\max}$ and $w_i = 0$ otherwise. A multiplicative age weight may also be included to reduce the influence of movies based on the time since their release date.

Using movies in the full estimation universe, estimate the regression $\log_{10} R_i = a_0(c,D) + a_1(c,D)c_i + a_2(c,D)D_i + \varepsilon_i$ where R_i is the total dollar revenue. In this regression each movie in the estimation universe is given weight w_i and the coefficients are written as functions of *c* and *D* to emphasize that the weights vary with position in the plane. The forecast logarithmic revenue at any point is then $a_0(c,D) + a_1(c,D)c + a_2(c,D)D$. For simplicity, the algorithm specified here involves only two factors, production cost *c* and director hit ratio *D*. The inclusion of additional continuous factors such as star power is a straightforward generalization, accomplished by adding quadratic terms to the definition of distance and corresponding linear terms to the regression.

Because it is a non-parametric technique, locally weighted regression does not produce a simple functional form. Instead, we evaluate the model's forecasts on a grid of values for the independent variables and store the results. Forecasts for new movies can be obtained by interpolating on this grid. This method is computationally intensive but its results can be made as precise as desired.

References

Ainslie, Andrew, Xavier Drèze, and Fred Zufryden. 2005. "Modeling Movie Life Cycles and Market Share," *Marketing Science*, vol. 24, p. 508.

Borod, Ronald S. 2005. "An Update on Intellectual Property Securitization," *Journal of Structured Finance*, vol. 11, no. 4 (Winter), p. 65.

Chance, Don M., Eric Hillebrand, and Jimmy E. Hilliard. 2005. "Pricing an Option on a Non-Decreasing Asset Value: An Application to Movie Revenue," working paper, Louisiana State University, December 16.

Clarke, Roger, Harindra de Silva, and Steven Thorley. 2002. "Portfolio Constraints and the Fundamental Law of Active Management," *Financial Analysts Journal*, vol. 58, no. 5 (September/October), p. 48.

Cleveland, William S., Susan J. Devlin, and Eric Grosse. 1988. "Regression by Local Fitting: Methods, Properties, and Computational Algorithms," *Journal of Econometrics*, vol. 37, p. 87.

De Vany, Arthur. 2004. *Hollywood Economics: How Extreme Uncertainty Shapes the Film Industry*. Routledge, London.

Einav, Liran. 2006. "Seasonality in the U.S. Motion Picture Industry," *RAND Journal of Economics,* forthcoming.

Eisbruck, Jay H. 2005. "Blockbuster or Flop? The History and Evolution of Film Receivables Securitization, 1995-2005," *Journal of Structured Finance,* vol. 11, no. 3 (Fall), p. 11.

Elberse, Anita, and Jehoshua Eliashberg. 2003. “Demand and Supply Dynamics for Sequentially Released Products in International Markets: The Case of Motion Pictures,” *Marketing Science,* vol. 22, p. 329.

Elberse, Anita, and Bharat Anand. 2005. “The Effectiveness of Pre-Release Advertising for Motion Pictures,” working paper, Harvard Business School.

Eliashberg, Jehoshua, and Steven M. Shugan. 1997. “Film Critics: Influencers or Predictors?” *Journal of Marketing*, vol. 61, no. 2 (April), p. 68.

Eliashberg, Jehoshua, Jedid-Jah Jonker, Mohanbir S. Sawhney, and Berend Wierenga. 2000. “MOVIEMOD: An Implementable Decision-Support System for Prerelease Market Evaluation of Motion Pictures,” *Marketing Science*, vol. 19, p. 226.

Goettler Ronald L., and Phillip Leslie. 2004. “Cofinancing to Manage Risk in the Motion Picture Industry,” working paper, Tepper School of Business, Carnegie Mellon University.

Goldman, William. 1983. *Adventures in the Screen Trade.* Warner Books, New York.

Greenspan, Alan. 2003. “Market Economies and Rule of Law,” at the 2003 Financial Markets Conference of the Federal Reserve Bank of Atlanta, Sea Island, Georgia, April 4.

Grinold, Richard C., and Ronald N. Kahn. 2000. *Active Portfolio Management,* second edition. McGraw-Hill, New York.

Hayes, Dade, and Jonathan Bing. 2004. *Open Wide: How Hollywood Box Office Became a National Obsession.* Hyperion, New York.

Hsu, Greta. 2005. “Jacks of All Trades and Masters of None: The Principle of

Allocation," *Administrative Science Quarterly,* forthcoming.

Motion Picture Association. 2006a. "US Movie Attendance Study," MPA Worldwide Market Research & Analysis.

_____. 2006b. "U.S. Theatrical Market: 2005 Statistics," MPA Worldwide Market Research & Analysis.

Nardone, Diane C. 1982. "Is the Movie Industry Contracyclical?" *Cycles*, vol. 33, no. 3 (April).

Neelamegham, Ramya, and Pradeep Chintagunta. 1999. "A Bayesian Model to Forecast New Product Performance in Domestic and International Markets," *Marketing Science,* vol. 18, p. 115.

Ravid, S. Abraham. 1999. "Information, Blockbusters, and Stars: A Study of the Film Industry," *Journal of Business,* vol. 72, p. 463.

Ravid, S. Abraham. 2004. "Are They All Crazy or Just Risk Averse? Some Movie Puzzles and Possible Solutions," *Economics of Art and Culture,* edited by Victor Ginsburgh, Elsevier, Amsterdam.

Sawhney, Mohanbir S., and Jehoshua Eliashberg. 1996. "A Parsimonious Model for Forecasting Gross Box-Office Revenues of Motion Pictures," *Marketing Science,* vol. 15, p. 113.

Sharda, Ramesh, and Dursun Delen. 2006. "Predicting Box-Office Success of Motion Pictures with Neural Networks," *Expert Systems with Applications*, vol. 30, p. 243.

Simonoff, Jeffrey S., and Ilana R. Sparrow. 2000. "Predicting Movie Grosses: Winners and Losers, Blockbusters and Sleepers," *Chance,* vol. 13, p. 15.

Surowiecki, James. 2005. "Disk Averse," *New Yorker,* August 8.

Swami, Sanjeev, Jehoshua Eliashberg, and Charles B. Weinberg. 1999. "SilverScreener: A Modeling Approach to Movie Screen Management," *Marketing Science,* vol. 18, p. 352.

Terry, Neil, Michael Butler, and De'Arno De'Armond. 2005. "The Determinants of Domestic Box Office Performance in the Motion Picture Industry," *Southwestern Economic Review*, vol. 32, p. 137.

Vanderhart, Jennifer, and Steven N. Wiggins. 2004. "Advertising, Quality, and Motion Picture Performance," *RAND Journal of Economics*, under review.

Vogel, Harold L. 2004. *Entertainment Industry Economics: A Guide for Financial Analysis*, sixth edition. Cambridge University Press, Cambridge, UK.

Walls, W. David. 2005. "Modeling Movie Success when 'Nobody Knows Anything': Conditional Stable-Distribution Analysis of Film Returns," *Journal of Cultural Economics,* vol. 29, p. 177.

Wang, Spencer, John Blackledge, and Aaron Chew. 2005. "The Future of Filmed Entertainment," JPMorgan Chase, December 2.