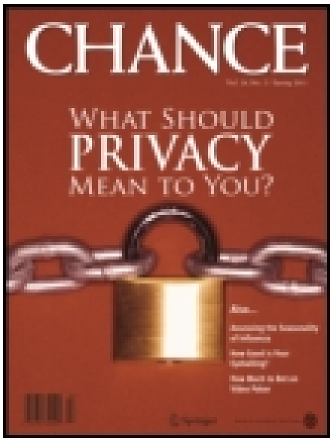


CHANCE
Publication details, including instructions for authors and subscription information:
http://amstat.tandfonline.com/loi/ucha20

Predicting Movie Grosses: Winners and Losers, Blockbusters and Sleepers

Jeffrey S. Simonoff & Ilana R. Sparrow

Published online: 20 Sep 2012.

The movie business often seems totally unpredictable. Is it?

Predicting Movie Grosses: Winners and Losers, Blockbusters and Sleepers

Jeffrey S. Simonoff and Ilana R. Sparrow

Introduction

The movie industry is a business with a high profile and a highly variable revenue stream. In 1998, moviegoers spent \$6.88 billion at the U.S. box office alone. A single movie can be the difference between millions of dollars of profits or losses for a studio in a given year. It's not surprising, therefore, that movie studios are intensely interested in predicting revenues from movies; the popular nature of the product results in great interest in gross revenues on the part of the general public as well.

In this article we examine the question of predicting domestic movie grosses from generally available information. The first version of this question relates to predictions prior to a movie's release, perhaps relatively early in production. How accurately can a studio predict the revenues for a movie before it opens? With some actors receiving as much as \$20 million per picture (plus a percentage of the gross revenues), producers obviously feel that star power leads to profit, but is that really true? Are certain types of movies more or less likely to be moneymakers? Do big budget movies make more money?

Tom Hanks, Matt Damon, and Edward Norton in *Saving Private Ryan*

A different version of the prediction question shifts the time frame to immediately after release. The opening weekend of a movie's release typically accounts for 25% of the total domestic box office gross, so we would expect that the opening weekend's grosses would be highly predictive for total gross. This, however, ignores the different release patterns of movies (some movies open on thousands of screens in the first weekend, others build slowly into wide release, and others never show on more than a few screens). Furthermore, do reviews by prominent critics have an effect on attendance?

Sometimes events long after a movie's release can have an impact on revenues. The most obvious of such events is when a movie earns awards, such as Academy Awards (or nominations). Do such awards provide a boost to revenues?

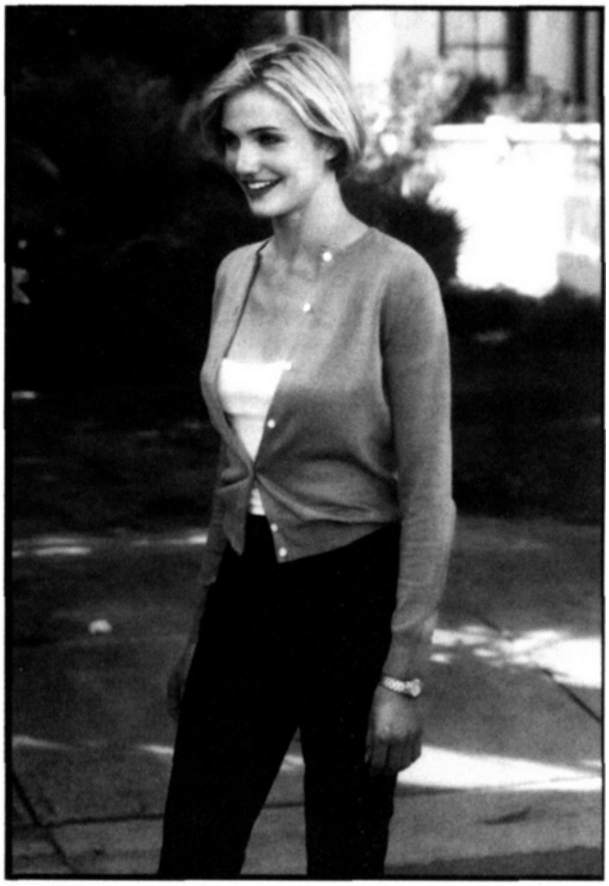
The Data

The analyses presented here are based on new movies released in the United States during calendar year 1998 for which relevant business information was available on the Internet Movie Database (*www.imdb.com*). Movies that opened on a limited number of screens late in 1997 (to be eligible for awards given in 1998) but then opened to wide release in 1998 are included in the sample; this includes movies such as *Good Will Hunting* and *Wag the Dog*. Similarly, movies that opened on a limited number of screens late in 1998 and then opened to wide release in 1999 are not included; examples of such films include *Shakespeare in Love* and *Life Is Beautiful* (*La Vita è Bella*). This yields a total of 311 films.

The response of interest here is the total U.S. domestic gross revenue for each film. Cursory examination of this variable shows that it is extremely skewed (ranging from a low of $349 for *Biker Dreams* to roughly $216 million for *Saving Private Ryan*), so the logarithm (base 10) of domestic gross will be used as the variable we attempt to explain in some analyses in the article. The predictor variables we consider include the following, versions of which have been considered in previous examinations of movie revenues and profits (although, to our knowledge, all have never been considered before in the same study):

1. The genre of the film, a categorical variable classifying the film as Action, Children's, Comedy, Documentary, Drama, Horror, Science Fiction, or Thriller. Genres were obtained from the *Videolog* catalog published by Blockbuster Video. For the few movies that were not in the catalog, genres were obtained from *Variety* magazine and *www.reel.com*, a Web site devoted to consumer purchase of movies.

2. The Motion Picture Association of America (MPAA) rating of the film, one of the ratings G (general audiences), PG (parental guidance suggested), PG-13 (possibly unsuitable for children less than 13 years of age), R (children not admitted unless accompanied by an adult), NC-17 (no one under 17 admitted), and U (unrated).



Cameron Diaz in *There's Something About Mary*

3. The origin country of the movie, classified as U.S., English-speaking (but not U.S.), or non-English-speaking.

4. Two variables attempting to measure "star power." The first is the number of actors or actresses appearing in the movie who were listed in *Entertainment Weekly*'s lists of the 25 Best Actors (August 7, 1998) and the 25 Best Actresses (November 20, 1998) of the 1990s. The second is the number of actors or actresses appearing in the movie who were among the top 20 actors and top 20 actresses in average box office gross per movie in their careers, according to *The Movie Times* Web site. The latter variable is as of the beginning of the 1998 movie season and only includes actors and actresses who had appeared in at least 10 movies at that time.

5. The production budget of the film (in millions of dollars). Cursory examination of this variable shows that it is also skewed, so the logarithm of the budget will be used in some analyses.

6. Whether or not the movie was a sequel to an earlier movie.

7. Three indicator variables identifying whether or not the movie was released before a holiday weekend (President's Day, Memorial Day, Independence Day, Labor Day, Thanksgiving, or the Christmas season), was released during the Christmas season (December 18-31), and/or was released during the summer season (Memorial Day through Labor Day).

8. The number of screens for the film's first weekend of general release. In addition to the movies mentioned earlier that opened to limited release in late 1997, some movies first opened to limited release in 1998, followed by wider release later that year (examples of such films are *The Boxer* and *Bulworth*). In all such cases, figures for the first weekend of wider (general) release are used. Cursory examination of this variable shows that it is also skewed, so the logarithm of the number of opening screens will be used.

9. The gross revenues (in millions of dollars) for the film's first weekend of general release. Cursory examination of this variable shows that it is also skewed, so the logarithm of first weekend gross will be used.

10. The rating of the movie (from zero to four stars) given by Roger Ebert, the well-known film critic from the *Chicago Sun-Times*, assuming Ebert reviewed the movie.

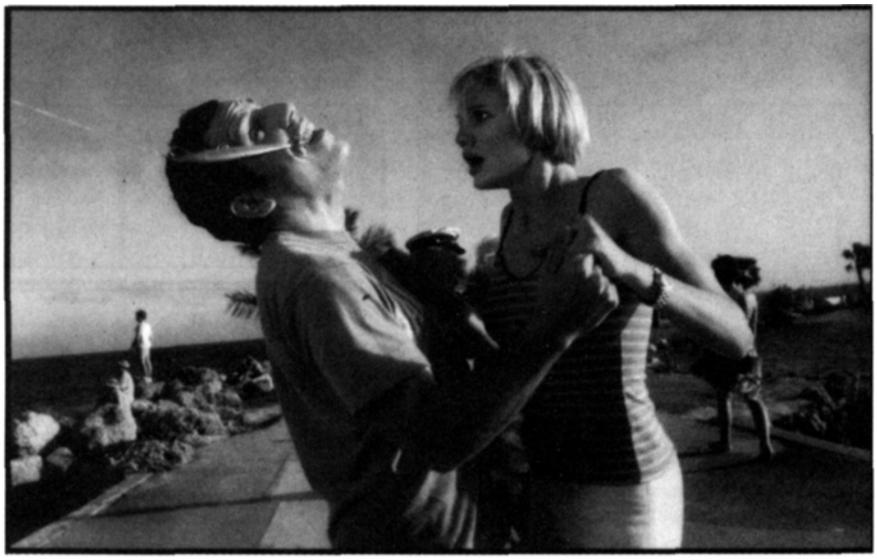
11. Academy Award (Oscar®) nominations and wins for the film. We restrict ourselves to the major categories of Best Picture, Best Director, Best Actor, Best

Actress, Best Supporting Actor, and Best Supporting Actress because these are the categories of greatest interest to the general public. Because we are examining gross theater revenues, we are only interested in nominations and wins for movies that are still in release at the time of the nominations.

Prediction of Revenues Prior to Release

We first examine prediction of movie grosses based on information available prior to release, perhaps early in production. Figure 1 gives side-by-side boxplots of total domestic gross separated by the level of different variables available before the release of a film (the width of the boxes is proportional to the square root of the number of observations in that group, so relative frequencies of different types of movies can be read from the plots as well). There are apparent differences in grosses using some of these variables. For example, the genre of the movie is predictive for grosses, with certain genres (Action, Children's, Horror, and Science Fiction) noticeably higher in revenues in general. Interestingly, there seems to be an inverse relationship between the occurrence of a genre and its revenue production, with the most popular types (comedies and dramas) performing relatively poorly at the box office. The relationship between grosses and MPAA rating is similar. As the rating becomes more "mature" from G to PG to PG-13 to R, more movies are made, typically making less money (of course, there is a connection with the genre results here, because all but one children's movie is rated G or PG, whereas more than two-thirds of comedies and dramas are rated R, NC-17, or U).

This effect can, in part, be explained by the success of well-targeted films that fit certain defined market segments. In the case of ratings, G, PG, and PG-13 movies are generally targeting children and/or teens. Similarly, certain genres are usually geared toward a narrow market segment (e.g., horror movies are often made with the teen audience in mind). Although studios may well want to produce more films in lucrative genres such as Action and Science Fiction, the barriers to entry for such films are prohibitively high — budgets can easily run from $70 to over $100 million with another $30–40 million or more for advertising and marketing. By contrast, the average cost to make and market a studio film in 1999 was considerably less at $87 million.



Cameron Diaz and Ben Stiller in *There's Something About Mary*

The high revenue of children's movies is somewhat illusory. Children's movies are actually composed of two very different kinds of films (live-action and animated) and are almost always rated G or PG. High revenues come from the animated films, which have long been considered to be one of the most profitable genres in the business (especially if the picture is by Disney). Live-action children's films, on the other hand, perform far worse than animated films (in 1998 the median total domestic gross for G-rated live-action films was roughly $10 million, but that for G-rated animated films was more than $100 million). The surprise successes of 1990's *Home Alone* and 1993's *Free Willy* notwithstanding, live-action children's films are often money-losers at the box office, and consequently, studios have scaled back their production of such films.

Although Disney has long dominated the animated film market, recent successful films from studios such as DreamWorks SKG (*Antz, The Prince of Egypt*) and films based on television shows (*Doug's 1st Movie, Pokémon the First Movie*), suggests that the lucrative animation market might get more attention in the future. Still, the economics of producing and releasing animated films are prohibitive. An animated movie often takes four years or more to produce (quite long by Hollywood standards) and involves hundreds of technically skilled artists and programmers who are not part of the studio's usual employees and would therefore need to be hired on as contractors. The market for such highly skilled people is dominated by a few top niche firms (such as Pixar and Industrial Light & Magic). Pixar, for example, is already in an exclusive relationship with Disney. Furthermore, as Dade Hayes, a *Variety* reporter based in Los Angeles, points out, "the marketing of such films requires a highly developed infrastructure of business partners such as toy makers and fast food chains, which many studios do not already have in place."

Other relationships are as would be expected. Movies made in English-speaking countries (including the United States) make more money (note that 35 of the 39 non-English movies are rated R or U, so there is a rating effect here as well). Even with a PG-13 rating, non-English-speaking movies have to overcome one of the toughest obstacles any movie could have, that of finding U.S. distribution. Non-English-speaking movies are a risky endeavor for most

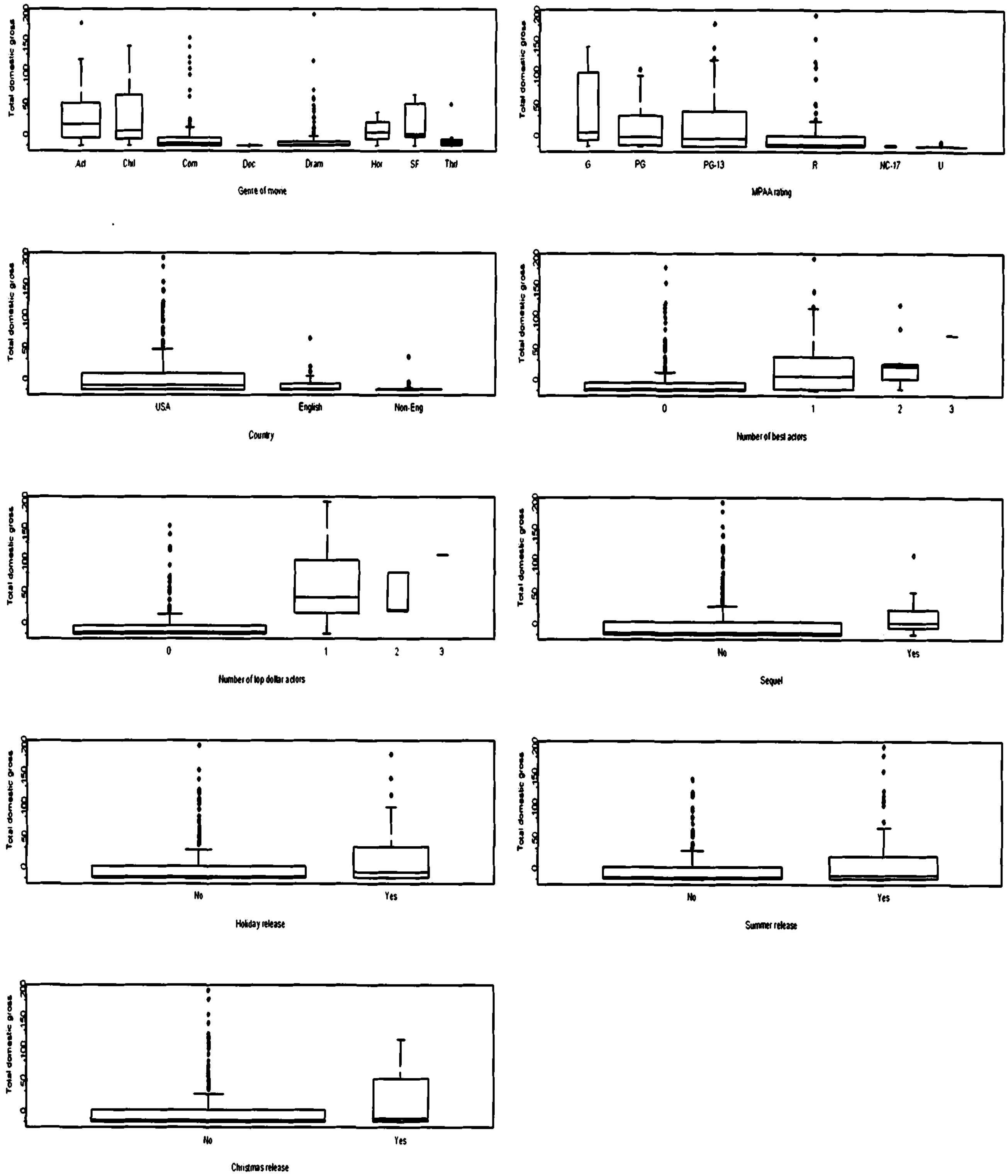


Figure 1. Side-by-side boxplots of total domestic gross in millions of dollars separated by levels of variables available before the release of a film.

distributors, because audiences are notorious in their dislike for subtitles. The market for foreign films is much smaller than that for U.S. films because its core market is restricted to an older, college-educated segment. Movies with more of the "best" actors and top-dollar actors make more money, an encouraging result for believers in "star power" (although obviously many movies with no big-name actors did very well at the box office). Sequels generally perform better than nonsequels, presumably reflecting the success of the earlier film(s), although with few exceptions sequels never perform as well as the original. It should be noted, however, that 10 of the 13 sequels came from the Action, Children's, Horror, and Science Fiction genres, the highest revenue groups. The three timing variables (holiday, summer, or Christmas release) apparently have less relation to revenues, although in all three cases the third quartile of grosses is noticeably higher for the seasonal releases than the nonseasonal releases. Movies that are expected to appeal to the youth market are usually released during the summer and holidays in hopes of getting high initial and repeat business, with other films moved to other times because of limited screen availability and the desire to avoid direct competition with a potential blockbuster.

These scheduling issues put a natural limit on the number of big-budget films (with potentially high revenue but also high risk) studios will release in any given year. Release dates are one of the most important factors in determining the success of a film. It is common practice in Hollywood to stake out a particular release date for a major film over a year in advance and to leak the information to the trades in the hopes of scaring off the competition. In general, it is unlikely to find two major blockbuster movies going head-to-head on one weekend. Because there are a limited number of screens and weekends, some films have been held back an entire season and marketing strategies rethought simply because a prime release date was already claimed by a rival studio. Interestingly, this is as equally true of small-budget pictures as of the typical blockbuster event film.

The graphs of Fig. 1 identify the effects of individual variables. To better predict gross revenues, we use linear regression models that incorporate some or all of the variables at the same time. As was noted earlier, we use regression modeling to model the logarithm (base 10) of the total gross. Table 1 summarizes the fit of our best prerelease model. The variables in the model include genre, MPAA rating, number of best actors, number of top-dollar actors, and whether or not the movie was a summer release. Note that the coefficients of the categorical predictors genre, MPAA rating, and summer release are constrained to sum to 0 across all of the categories. All of these variables have effects significantly different from 0 at traditional levels. Standard errors of the coefficients are also given.

Table 1 — Summary of Model Fitting Using Variables Available Before the Release of a Film

Variable	Effect	Coefficient	
Constant		.394	(.164)
Genre	Action	.401	(.178)
	Children's	-.030	(.335)
	Comedy	-.189	(.126)
	Documentary	-1.248	(.264)
	Drama	-.408	(.120)
	Horror	.513	(.261)
	Science Fiction	.693	(.300)
	Thriller	.267	(.275)
MPAA rating	G	.534	(.445)
	PG	.380	(.204)
	PG-13	.312	(.191)
	R	-.079	(.175)
	NC-17	-.118	(.591)
	U	-1.028	(.208)
Summer release	No	-.150	(.065)
	Yes	.150	(.065)
Best actors		.400	(.110)
Top-dollar actors		.712	(.157)

Note: Standard errors of coefficients are given in parentheses next to the coefficients.

Predictions from the model are made by summing the constant, the appropriate effects for genre, MPAA rating, and summer release, and the products of the coefficients for the actors with the actual numbers for the film. For example, *Dancing at Lughnasa* was a PG-rated drama not released during the summer with one best actor and no top-dollar actors, so its predicted log gross is $.394 - .408 + .380 - .150 + (1)(.400) + (0)(.712) = .616$, and its predicted gross is $10^{0.616} = \$4.13$ million.

The coefficients in Table 1 (which take all other variables in the model into account) correspond closely to the effects seen in the boxplots of Fig. 1. That is, although action, horror, and science fiction movies generate more revenue than average, dramas and (especially) documentaries do worse than average. The poorer performance of movies rated R and NC-17 is also clear, as is the benefit of a summer release. The two strongest effects (given all else in the model) are the "star-power" variables, with each additional cast member from the best actors list multiplying estimated gross revenue by roughly 2.5 ($10^{0.4} = 2.51$) and each additional cast member from the top-dollar actors list multiplying it by more than 5 (holding all else fixed). Note, by the way, that, because these actors no doubt are paid high salaries, these increased revenues do not necessarily translate into increased profits.

Overall, the model explains 44.6% of the variability in log domestic revenue. The standard error of the regression is .982, which indicates the weakness of the model for predicting grosses. The ends of an approximate 95% prediction interval for logged domestic gross are the estimate $\pm$ (2)(.982) = 1.964, or roughly $\pm$ 2. This means that predictions of total grosses for an individual movie can be expected to be off by as much as a multiplicative factor of 100 high or low.

Table 2 summarizes observed and predicted total grosses (in millions of dollars) for selected 1998 movies that illustrate the properties of the model. Predictions based on the model in this section are given under the heading "Prerelease model" (the other two models are discussed in later sections). The grosses of *Air Bud* (a G-rated children's movie released during the summer) and *Stepmom* (a PG-13-rated drama costarring Julia Roberts, Susan Sarandon, and Ed Harris) are well predicted by the model. The model is less successful for

Table 2 — Observed Total Domestic Grosses and Predicted Grosses in Millions of Dollars Based on Three Models for Selected 1998 Films

Movie	Total gross	Prediction (pre-release model)	Prediction (first weekend model)	Prediction (Oscar model)
3 Ninjas: High Moon at Mega Mountain	.308	10.589	.269	.278
Air Bud: Golden Receiver	10.215	11.183	6.354	6.416
Godzilla	136.023	18.084	306.910	299.889
Good Will Hunting	138.339	7.400	44.591	131.984
The Horse Whisperer	75.370	1.405	63.932	59.391
Hurlyburly	1.796	18.571	1.373	1.310
Saving Private Ryan	216.119	14.791	194.622	358.237
Stepmom	91.030	114.535	97.783	90.513
There's Something About Mary	176.483	1.893	56.234	55.628

the other films in the table. It is clear that the model seriously underestimates the grosses of big-budget films like *Godzilla* and *Saving Private Ryan*. The tremendous success of two of the sleeper films of 1998, *Good Will Hunting* (an R-rated drama costarring Robin Williams) and *There's Something About Mary* (an R-rated summer comedy), is completely missed by this model using prerelease properties (which can be viewed as precisely what made them sleepers).

The failure of the model for the big-budget movies mentioned earlier begs the question of why the (logarithm) of the budget is not used in the model. It is true that gross and budget are positively related. Unfortunately, this relationship may not be what it seems. First, most of the budget numbers (almost 60%) are not available. More importantly, the numbers that are available are suspect at best. Budget figures are notoriously difficult to obtain. In addition, the production budget only represents part of the cost of making a movie (it ignores, for example, advertising and distribution costs). Even if a number is available, it is very difficult to evaluate it, as the specifics surrounding the financing of a given film are cloaked in secrecy. For example, a budget figure may be deceptively low given the star and director power (*Saving Private Ryan*, with a $65 million budget, is a good example of this) because the director and star waived their normal salary requirements for a participation in the film's gross receipts. Unfortunately, it is impossible to find out the exact amount of "back-end" paid out under such circumstances. Thus, even though studio heads would know the budgets for their own films, we cannot build a prediction model based on budgets. For this reason, we will not consider the budget as a predictor (for the same reason, we are focusing on revenues here, rather than the perhaps more interesting profits or returns on investment).

Prediction of Revenues After the First Weekend of Release

The performance of a movie during its first weekend of release is watched closely by people both inside and outside the film industry. Leonard Klady, formerly a *Variety* entertainment reporter, describes a film's opening weekend as "a bellwether," adding that "by and large, most films tend to open at their height." It is on the strength of the opening weekend of general release that all major decisions pertaining to a film's ultimate financial destiny are made. Because competition for movie screens is fierce, movie theater owners do not want to spend more than the contractually obligatory two weeks on a film that doesn't have "legs." Should a film lose its theatrical berth so quickly, chances are slimmer that it will have significant play internationally (if at all), and it is unlikely that it will make it to pay-per-view, cable, or network television. This all but guarantees that ancillary revenue streams will dry up, making a positive return on investment very difficult to achieve because ancillary deals are predicated on domestic box office gross. Movie theater owners often make the decision to keep a film running based on the strength of its opening weekend.

Of course, there are certain circumstances in which a film that does poorly at the outset may yet be saved. Perhaps, for example, the studio will throw more money into the ongoing promotional campaign targeting a different audience than was targeted up front. Tim Noonan, a business development executive formerly at New Line Cinema, describes the first weekend gross as "predictive of what the movie will do overall...it's a signal to us to extract the most value as an asset. Can you do more with it? Take it wider?" The first weekend gross helps a studio learn if the marketing strategy was dead-on or wrong, and if wrong, if it is salvageable. A recent example Noonan cited was the case of *Drop Dead Gorgeous*, a 1999 New Line release that was given a national wide release but performed disappointingly in its opening weekend. Looking at its latest market research results, New Line discovered that, although Middle America didn't get the movie, a more narrow demographic — sophisticated, gay urbanites — loved it. As a result, the studio reevaluated *Drop Dead Gorgeous*' positioning and decided to pull all national ad campaigns and pour the money into targeted regional efforts instead. This quick shift in strategy managed to keep the film in theaters longer, giving it a better shot at long-term profitability.

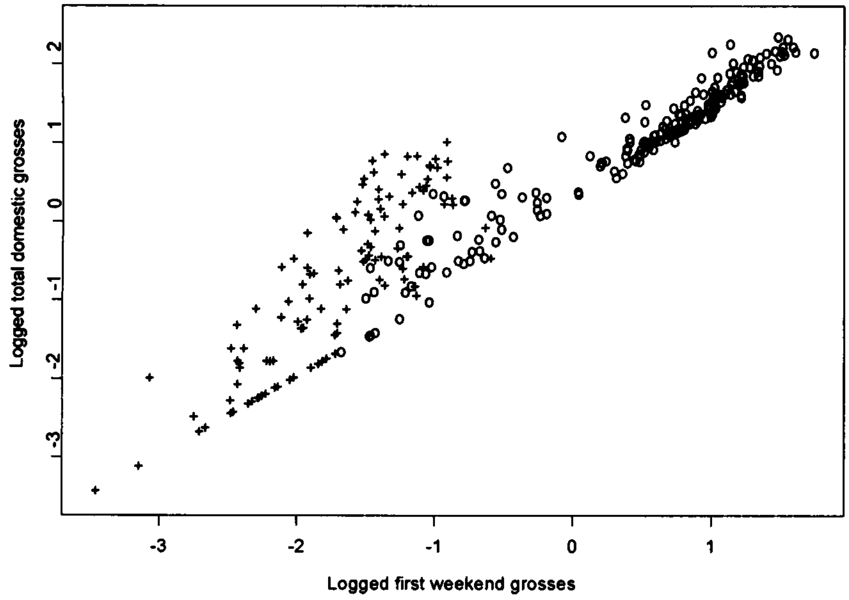


Figure 2. Scatterplot of logged total domestic gross versus logged first weekend gross, with movies labeled by whether they opened on 10 screens or less (pluses) or more than 10 screens (circles).

Clearly it is believed that the first weekend's performance is highly predictive for ultimate performance, but is this actually true? Figure 2 suggests that it is, but only to a certain extent. Although the logged first weekend gross is highly correlated with logged total domestic gross ($r = .93$), it is apparent that the strength of the relationship is different for movies with lower grosses versus movies with higher grosses. In fact, it turns out that the number of opening screens is crucial. This is clear in Fig. 2, where movies are identified by whether or not they opened on more than 10 screens. It is apparent that the relationships in the plot are different for the two groups, in terms of both slope and variability. For this reason, the movies will be separated into two groups— those that open on 10 screens or less (116 films), and those that open on more than 10 screens (192 films) (the number of opening screens was unavailable for three films).

Table 3 summarizes the results for regression models predicting log gross from information available after the first weekend. The chosen model for films opening on 10 or fewer screens is based on only logged first weekend gross and logged opening screens, with an R^2 of 73.5%. The coefficients for the model are nicely intuitive. Holding the number of opening screens fixed, multiplying the first weekend's gross by 10 is associated with multiplying the total gross by $10^{1.8} = 64.3$. On the other hand, holding the first weekend's gross fixed and multiplying the number of screens by 10 (resulting in a 90% lower per screen average revenue) is associated with a decrease in the total gross by 71% ($10^{-0.54} = .29$). The standard error of the regression estimate $s = .544$ implies that predictions can be expected to typically be off by no more than a multiplicative factor of roughly 10. This is obviously much improved over the factor of 100 possible before the opening weekend.

The situation is better for the movies released on more than 10 opening screens. The chosen model is based on genre, logged first weekend gross, and logged opening screens. The coefficients for the two continuous predictors are similar to those for the movies in very limited opening release. Holding the number of opening screens fixed, mul-

Table 3 — Summary of Model Fitting Using Variables Available After the First Weekend of Release of a Film

Movies opening on 10 or less screens		Movies opening on more than 10 screens	
Variable	Coefficient	Variable (Effect)	Coefficient
Constant	2.509 (.262)	Constant	1.983 (.133)
Logged first weekend gross	1.808 (.116)	Logged first weekend gross	1.525 (.045)
Logged opening screens	-.539 (.189)	Logged opening screens	-.598 (.052)
		Genre	
		Action	-.054 (.037)
		Children's	.114 (.046)
		Comedy	.029 (.032)
		Documentary	.009 (.152)
		Drama	.069 (.034)
		Horror	-.094 (.050)
		Science Fiction	-.047 (.056)
		Thriller	-.026 (.062)

Note: Standard errors of coefficients are in parentheses.

Table 4 — Summary of Model Fitting Using Variables Available After the First Weekend of Release of a Film and Oscar Variables

Movies opening on 10 or less screens		Movies opening on more than 10 screens	
Variable	Coefficient	Variable (Effect)	Coefficient
Constant	2.392 (.258)	Constant	1.961 (.129)
Logged first weekend gross	1.756 (.114)	Logged first weekend gross	1.512 (.044)
Logged opening screens	-.537 (.184)	Logged opening screens	-.590 (.051)
Oscar nominations	.409 (.149)	Oscar nominations	.100 (.027)
		Genre	
		Action	-.048 (.036)
		Children's	.118 (.045)
		Comedy	.031 (.031)
		Documentary	.006 (.147)
		Drama	.047 (.033)
		Horror	-.090 (.049)
		Science Fiction	-.041 (.054)
		Thriller	-.023 (.060)

Note: Standard errors of coefficients are in parentheses.

tiplying the first weekend's gross by 10 is associated with multiplying the total gross by $10^{1.53} = 33.5$, while holding the first weekend's gross fixed and multiplying the number of screens by 10 (resulting in a 90% lower per screen average revenue) is associated with a decrease in the total gross by 75% ($10^{-0.6} = .25$). Note that the effects for different genres are very different from the marginal relationships shown in Fig. 1. This is because most of the patterns noted earlier (e.g., the large revenues for horror films) are now explained by the first-weekend statistics. The coefficients do show that, for example, children's movies do noticeably better than first-weekend results would suggest (30% better given the first-weekend numbers because $10^{0.114} = 1.3$), suggesting that they build their box office slowly, whereas horror films drop off more quickly from the first weekend, finishing 20% lower given the first-weekend numbers. Richard Natale, in a June 23, 1998, *Los Angeles Times* article, reported that a 35% drop in attendance in the second week of release is considered normal for a film targeted to the adult audience, but a horror movie may drop off by 50% in its second week of release because the teen audience that tends to go to these films prefers to see a movie as soon as it opens.

The excellent predictive power of the model is obvious from the standard error of the regression estimate, $s = .17$, because it implies that predictions can be expected to typically be accurate within a multiplicative factor of roughly 2. Thus, for the typical movie at the neighborhood multiplex, knowing the first-weekend results provides an excellent forecast for total revenues.

Note that for both regression models summarized in Table 3, once the first weekend statistics are available, previously useful information becomes superfluous. That is, any benefits of star power or a child-friendly MPAA rating in predicting total gross are accounted for by the first-weekend results and don't add any further predictive power.

A return to Table 2 shows the usefulness of the added information by examining the entries under "first weekend model." The ultimately surprisingly poor showings of *3 Ninjas* and *Hurlyburly* and surprisingly good performances of *The Horse Whisperer* and *Saving Private Ryan* weren't surprising at all after the first weekend. *Good Will Hunting* and *There's Something About Mary* are better predicted than before, but the predictions are still considerably too low. Interestingly, *Godzilla* now shows up as a severe underachiever, which makes sense. It opened to very large crowds initially, but horrible reviews and word-of-mouth ultimately made it one of the biggest disappointments of 1998 (even though it was the tenth-highest domestic revenue producer).

It is worth commenting here on the absence of one seemingly obvious factor — reviews. Roger Ebert reviewed roughly 70% of the movies released in 1998. It is difficult to incorporate his reviews (by simply dropping the cases with missing data or more complex imputation schemes) because the movies he reviewed are certainly not a random sample of those released. Ebert was presumably more likely to review movies opening on more screens and more likely to review movies expected to be of higher quality (even if they had only limited release). It would be a good idea to measure critical opinion using some sort of consensus review in future work.

Do Oscars Matter?

The Academy Awards (known as Oscars) are the most sought-after prizes in the film industry. Every year movie studios spend tens of millions of dollars promoting certain films for Oscar consideration, such as in trade advertising (some studios have been rumored to spend as much on trade advertising vying for Academy recognition for a film as they have on consumer campaigns for the film). Because the nominations occur after the first weekend of release, an interesting question is whether

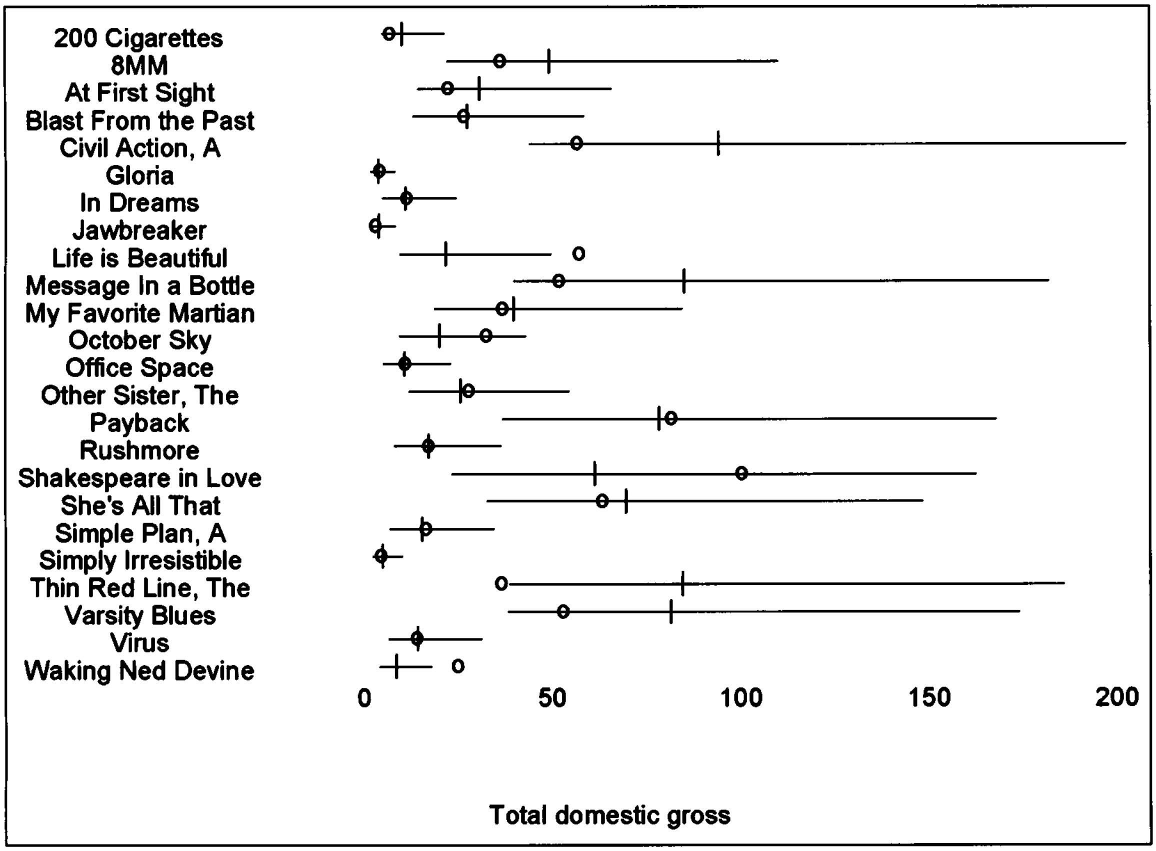


Figure 3. Predicted and observed total grosses for films released in early 1999. Predicted grosses are marked with vertical tics, observed grosses are marked with o, and prediction intervals are given by horizontal lines.

receiving nominations or awards (particularly in the major categories) provides a boost to revenues, given the revenues already earned. This question can be examined from a statistical point of view by seeing whether variables representing Oscar nominations and awards provide additional power for predicting gross revenue.

Table 4 summarizes the results of adding the number of major Oscar nominations to the regression models of Table 3. The nomination effect is appreciable for both classes of movies. For movies opening on no more than 10 screens, each additional Oscar nomination is associated with multiplying the expected gross by roughly 2.5 (holding all else fixed), but for movies opening on more than 10 screens, each Oscar nomination is associated with increasing the expected gross by roughly 30%.

The overall fit of the models is improved only slightly because only 10 of 311 movies received nominations in the major categories while still in release. This can be seen in Table 2 because the predicted grosses for all of the movies except *Good Will Hunting* and *Saving Private Ryan* (the two movies in the list with nominations) are very similar in the final two columns.

For 9 of the 10 Oscar-nominated films, predictions using Oscar nominations are more accurate than those not using the nominations, reflecting that the films benefited from being nominated (five-time nominee *Good Will Hunting* is a prime example of this). The only exception to this was *Saving Private Ryan*, which did not seem to benefit from its nominations. The reason for this is that it was released in July, nine months before the Academy Awards, so Oscar publicity came too late to boost domestic box office revenues.

Using the Models to Predict New Movie Grosses

Although the implications of the models fit in the previous sections as descriptions of the revenue process are interesting, the real usefulness of such models to a producer or a studio is in the prediction of future revenues for new projects. To investigate this, we applied our final model (using the coefficients derived from 1998 data) to the 24 movies put in wide release in the first two months of 1999. This includes five films put in Oscar Engagement release in late 1998 that were nominated shortly thereafter for major Academy Awards (*A Civil Action*, *Life Is Beautiful* (*La Vita è Bella*), *Shakespeare in Love*, *A Simple Plan*, and *The Thin Red Line*).

Figure 3 gives predictions and prediction intervals for the total domestic gross for the 1999 films, along with the observed values. In 21 of the 24 intervals, the observed gross is well within the 95% prediction interval, and many predicted values are very close to the observed values. Though this is a bit below the 95% coverage we'd hope for, the performance seems quite reasonable to us. The observed grosses for *Waking Ned Devine*, and to a greater extent *Life Is Beautiful* and *Shakespeare in Love* are higher than expected. All three films opened in general release on fewer than 650 screens, gradually widening release to (in the case of *Shakespeare in Love*) over 2,000 screens. This strategy, although not typical, is also not that unusual, especially for movies targeting a more educated demographic, where it is often successfully employed. Obviously, first weekend grosses will underestimate total grosses when this strategy is used.

Conclusion

The analyses given here illustrate that the ultimate box office performance of movies can be forecast with some accuracy given easily available information. The predictions are especially accurate after the first weekend of release for movies opening on more than 10 screens, although the tendency for some distributors to slowly widen release of a film based on word of mouth complicates matters. Oscar nominations in the major categories do seem to provide a boost to revenues, as long as the movie has not already been in release for many months when the nominations are announced.

The models discussed here have certain limitations. It is quite possible that other predictors could improve the predictive power of the models. Two movies of the same type opening at the same time can result in reduced revenues for one or the other (or both) in the short run: Does this effect carry over to long-run revenues? Clearly studios think so because a great deal of thought goes into choosing release dates. It would also be interesting to see how much subsequent weeks' revenues add to predictive power over the first weekend grosses. A clearly important potential factor is the advertising budget for the film, but as was noted earlier, these figures are extremely difficult to obtain.

The focus here on domestic revenues ignores foreign grosses, which would be more difficult to predict but are a very important revenue stream (now accounting for roughly 60–70% of the total box office gross). It is well known that certain stars (such as Sylvester Stallone and Kevin Costner) and certain subject matter (such as that in the controversial 1999 release *Eyes Wide Shut*) play better internationally than they do domestically. Other potentially important sources of revenue outside the theatre are also not considered here. These include home video, cable television, and network television, all of which could be important revenue streams to movie producers.

The surprise successes of 1998's *There's Something About Mary* and 1999's low-budget thriller *The Blair Witch Project* demonstrate that even if a model performs well overall, the movie business is characterized by occasional films that defy expectations. "Since predicting gross is extremely difficult," comments Noonan, "you have to serve up a [yearly] slate of movies and know that over time you'll have 3 or 4 to the left and 2 or 3 to the right. You must make sure you are doing things that mitigate your downside risk."

[The authors wish to thank Sam Craig, Dade Hayes, Leonard Klady, Tim Noonan, Ben Procter, and Joel Steckel for their invaluable contributions to this article.]

References and Further Reading

Dodds, J.C., and Holbrook, M.B. (1988), "What's an Oscar Worth? An Empirical Estimation of the Effects of Nominations and Awards on Movie Distribution and Revenues," in *Current Research in Film: Audiences, Economics, and Law*, ed. B.A. Austin, Norwood, NJ: Ablex, pp. 72–88.

Eliashberg, J., and Shugan, S.M. (1997), "Film Critics: Influencers or Predictors?" *Journal of Marketing*, 61, 68–78.

Neelamegham, R., and Chintagunta, P. (1999), "A Bayesian Model to Forecast New Product Performance in Domestic and International Markets," *Marketing Science*, 18, 115–136.

Radas, S., and Shugan, S.M. (1998), "Seasonal Marketing and Timing New Product Introductions," *Journal of Marketing Research*, 35, 296–315.

Ravid, S.A. (1999), "Information, Blockbusters and Stars — A Study of the Film Industry," *Journal of Business*, 72, 463–492.

Sawhney, M.S., and Eliashberg, J. (1996), "A Parsimonious Model For Forecasting Gross Box-Office Revenues of Motion Pictures," *Marketing Science*, 15, 113–131.

Wallace, W.T., Seigerman, A., and Holbrook, M.B. (1993), "The Role of Actors and Actresses in the Success of Films: How Much Is a Movie Star Worth," *Journal of Cultural Economics*, 17, 1–27.

The cast of *Saving Private Ryan*