

# First You Get the Money, Then You Get the Reviews, Then You Get the Internet Comments: A Quantitative Examination of the Relationship Between Critics, Viewers, and Box Office Success

Jeremy Berg & M. Jordan Raddick

To cite this article: Jeremy Berg & M. Jordan Raddick (2016): First You Get the Money, Then You Get the Reviews, Then You Get the Internet Comments: A Quantitative Examination of the Relationship Between Critics, Viewers, and Box Office Success, Quarterly Review of Film and Video, DOI: [10.1080/10509208.2016.1156426](https://doi.org/10.1080/10509208.2016.1156426)

To link to this article: <http://dx.doi.org/10.1080/10509208.2016.1156426>



Published online: 28 Apr 2016.



Submit your article to this journal [↗](#)



Article views: 98



View related articles [↗](#)



View Crossmark data [↗](#)

## **First You Get the Money, Then You Get the Reviews, Then You Get the Internet Comments: A Quantitative Examination of the Relationship Between Critics, Viewers, and Box Office Success**

Jeremy Berg and M. Jordan Raddick

The American movie industry has been much studied, as befits what is not only a cultural institution in its native country, but a major export to other nations as well. Academic articles on film have frequently been critical or historical studies focused on one or a few films.<sup>1</sup> At the same time, film studios and distributors, and the consultants they hire, have compiled detailed sales and marketing analyses of movie performance, but these studies remain proprietary and do not appear in the open literature.

We submit that there is a clear need for studies of American movies as a whole, looking for trends in both the film industry and popular culture, and including as many movies as possible. Quantitative methods are ideal for such a study, because they make relatively few theoretical assumptions and easily scale up to incorporate many data sources. The key to such studies is to identify measurable quantities of interest (variables) that provide insight into research questions, and to use as large and complete a dataset as possible.

The main research question we seek to answer is: What, if any, are the relationships between a movie's perceived artistic merit as ranked by critics, a movie's quality as ranked by viewers, a movie's gross box office take, and its release date? The relationship between these four categories is an important one. As much as one may decry the ties between art and commerce, they always exist, and perhaps nowhere more so than in film. Even as the word "film" itself becomes increasingly obsolete with the rising prominence of digital video, movies remain fantastically expensive, with professional product well beyond the reach of almost any one person to fund. All arts become expensive when skilled professionals need to be paid and when professional grade equipment is purchased, but none more so than the

---

Jeremy Berg is a cataloging librarian at the University of North Texas. He received a Master's Degree in Library and Information Science from the University of Illinois, a Bachelor's Degree in Creative Writing from Johns Hopkins University, and has been an avid moviegoer since his teens. He grew up in South Jersey, has lived in Baltimore, Milwaukee, and Normal, IL, and currently resides in Texas.

M. Jordan Raddick is the Associate Director for Education at the Institute for Data-Intensive Engineering and Science (IDIES) at Johns Hopkins University. He received a Master's Degree in Writing About Science from Johns Hopkins and a Bachelor's Degree in Geophysics from Brown University. His passion is bringing big data management and analysis methods to the wider world.

Color versions of one or more of the figures in the article can be found online at [www.tandfonline.com/gqrf](http://www.tandfonline.com/gqrf).

© 2016 Taylor & Francis Group, LLC

movies. Paint, canvas, sculpting materials, and musical instruments can all cost a great deal of money, and recording studio time operates at hundreds of dollars an hour. But nowhere in music, writing, or the visual arts has a budget of five million dollars ever been considered fantastically cheap, as it is for movies. Even independent features with volunteer casts and crews cost sums into five and six figures.

Because movies are so expensive, they frequently require corporate backing to be funded, or at least to be distributed. Because movies are backed by corporations, the amount of money that they make is very important not only to whether more work is given to the primary artists involved (director, screenwriter, actors), but also to the whether more movies of that type are made at all. Furthermore, financial considerations are the primary driver of studio decisions about which movies to release and when to release them—decisions which can become self-fulfilling prophecies when, for example, a movie predicted to do poorly at the box office is narrowly released and poorly marketed at a time when theater attendance is low, leading to poor box office performance. Cultural theorists have been particularly interested in the phenomenon of “dump months,” when studios supposedly release low-quality movies.<sup>2</sup>

These financial considerations have an obvious impact not only on the film industry, but also on the cultural products the industry produces. Cultural critics have often decried the phenomenon that movies derided by film critics and audiences frequently perform very well at the box office,<sup>3</sup> as well as what they see as a divergence between the judgments of perceived experts—professional movie critics—and consumers.<sup>4</sup> These perceptions have become the conventional wisdom in cultural discussions of film—but like all conventional wisdom, these ideas need to be tested against the available evidence.

The genesis of this article was a *Slate* article,<sup>5</sup> reported and discussed in the A.V. Club,<sup>6</sup> claiming that movies released in February were measurably of lower quality than those released in other months. Discussing these articles, and particularly some of the comments in the A.V. Club article, inspired a collaboration between the two co-authors, a librarian and cultural studies researcher (Berg) and a data scientist (Raddick). The research expanded when the authors realized that the claims could not be tested from the aggregate data presented in the *Slate* article, so a new dataset would need to be constructed. That dataset, containing financial data and ratings from both viewers and critics from 6,820 films released between 2000 and 2013, is a major outcome from this research. The dataset is described in more detail in the Data and Methods section herein. We hope that it will be useful to researchers in many fields for answering a wide variety of research questions.

## Literature review

While we found no previous research that comprehensively charted the monetary success of a film in terms of its connection to release dates, critical reception, and

viewer evaluation with a quantitative data set, researchers in several different fields have addressed elements of this topic.

Basuroy, Chatterjee, and Ravid<sup>7</sup> studied the relationships between critics' reviews of films and those films' box office performance, seeking to investigate whether critical reviews exert influence over audience demand, or merely predict what audience demand will be. Reinstein and Snyder<sup>8</sup> examined how the timing of movie reviews published by the famous critics Gene Siskel and Roger Ebert correlates with those movies' box office performance. Plucker, Kaufman, Temple, and Qian<sup>9</sup> compare ratings of the same films among critics, online movie fans, and college freshmen. Like us, they use IMDb as their source for viewer data and Meta-critic as their source for critic data.

The phenomenon of “dump months” in particular has been a popular topic of study in the academic literature around film. McClintock<sup>10</sup> chronicles the increasingly crowded summer release schedule that is arguably a logical consequence of studios trying to put most films out during a prime season. She offers no real judgments and was unable to draw any conclusions, since she reported on an upcoming schedule, rather its results. However, she does quote DreamWorks Chief Jeffrey Katzenberg on the idea that unlimited free time for students over the summer allows sufficient room for all films released then.<sup>11</sup>

Sochay<sup>12</sup> offers a valuable predecessor to our research, by looking into both the release date and critical reception's relation to economic success. Sochay also found a paucity of prior research.<sup>13</sup> Unlike our study, Sochay's primary research tools were the length of time a movie played in theaters and rental data;<sup>14</sup> nonetheless, his conclusions and synthesis of previous studies also show that Christmas and summer are the prime movie-going seasons, with the two together accounting for 51% of all ticket sales as of 1989.<sup>15</sup> Additionally, he cites Litman and Kohl,<sup>16</sup> who reported a trend that the Christmas season was decreasing in financial importance while summer increased.<sup>17</sup>

Rottenberg<sup>18</sup> suggests that the importance of the Christmas season lies in the fact that every day is open for movie-going.<sup>19</sup> He also brings in the Oscar factor, discussing Christmas as a prime release date for Oscar hopefuls.<sup>20</sup> The importance of Oscar release strategies to the ebb and flow of both critical reception and box office take is considerable. Bart<sup>21</sup> agrees, considering even September to be “early” for releasing an Oscar contender.<sup>22</sup>

Cartier and Liarte<sup>23</sup> point to summer as a high-income season: “Nearly 50% of the biggest box office successes are released during the summer. Sustained demand can be seen between week 20 and week 29 [between mid-May and mid-July].”<sup>24</sup> They also studied the effects of concentrating release dates in the summer on box office take for the movie industry as a whole. Their conclusions suggest an *ouroboros* effect, with movies performing better during summer, but also that so many movies being released at the same time contributes to that superior performance. Similarly, while almost half of all box office successes are released in the summer,<sup>25</sup> the amount of films out at once makes it necessary to outspend competitors in advertising to rise above the pack.<sup>26</sup>

Cartier and Liarte also raise the intriguing possibility that a well-reviewed film increases all movie attendance, but that this effect caps at a certain number of simultaneous releases.<sup>27</sup> Einav<sup>28</sup> expands on the issue of oversaturation with a call for releases to be spread out throughout the year. Darling<sup>29</sup> raises the possibility that year-round major releases are already an emerging trend, although there is insufficient evidence to chart the long-term, or even short-term, success of such a plan as of writing. Burr<sup>30</sup> argues that the modern release schedule with its ups and downs was engineered, whether on purpose or not, by the modern movie business as opposed to being a naturally occurring ebb and flow.

Nearly all authors agree that the most desirable periods for new theatrical releases are during the summer and the Christmas season. Most discussion of release timing effects in the literature have taken the opposite approach, looking for the periods that studios find least desirable for new releases. These periods have been termed *dump months*.

The name comes from studios “dumping” movies that they do not expect to make money into an established set of months where they are left to founder with little in the way of promotion; primarily January but also, to a lesser extent, February, August, and September. Although we were unable to find many scholarly articles on the subject, dump months are well known to film writers; Bernstein (2007), Burr (2013), Darling (2014), Dietz (2010), LaMarche (2013), Legel (2013), Meslow (2012), Raymond (2013), Salisbury (2013), and Singer (2013) all corroborate January’s long-held position as the premiere dump month. Maslin<sup>31</sup> catches January-as-dump-month as an emerging phenomenon.

There are several theories for why January became a dump month. The most popular, as advanced by Burr, Dietz, Legel, and Meslow, is that January is ceded to Oscar contenders that see wide release in January following a limited release in December for purposes of eligibility. Dietz also lists winter weather as a limiting factor in people’s desire to go out to the movies in January, which Burr alludes to without specifically naming the month. Legel<sup>32</sup> and Salisbury<sup>33</sup> emphasize the youth audience by citing the return to school in January as an issue. This tradition of January as a dump month goes at least as far back as the 1930s:

In fact, there’s a long-established studio history of holding a prestige release until the very end of the calendar, with premieres in New York and L.A. in the final weeks of December to qualify for the Oscars, then rolling it out to the rest of the country in the new year. MGM did it with the Greta Garbo classic *Queen Christina* in 1933, and again with *Gone With the Wind* in 1939.<sup>34</sup>

While January has traditionally been identified as the most important dump month, authors have identified several other months as “dump months” as well. Darling<sup>35</sup> and Kirk & Thompson<sup>36</sup> finger February as a dump month, and LaMarche alludes to it by saying the “first months of the year”<sup>37</sup> are dump months. Using the concept of “The August Movie,” Burr, Darling, and Singer all confirm August as a dump month, and Darling includes September as well. Rottenberg<sup>38</sup>

refers to spring as “historically something of a dumping ground.”<sup>39</sup> Darling cites the return to school as the reason August is a dump month as well; however trends over the past eight years show August moving away from dump month status as documented by Darling and “The August Movie.”<sup>40</sup>

Financial and, even more rarely, critical success does occasionally occur in the dump months. Salisbury<sup>41</sup> and Singer<sup>42</sup> both point out that the low expectations for a film’s performance can lead to interesting and different films getting a chance for theatrical release. “August Movies,” Bernstein, Burr, and Salisbury all discuss dump month movies that unexpectedly made money.

Finally, it is important to remember that the changing nature of the American film industry can quickly render some study methods out of date. Sochay<sup>43</sup> used rental data as an important measure of a movie’s financial success; but today, multiple online streaming and cable on-demand options have rendered it, at best, a vastly incomplete and likely biased picture of a film’s consumption. Cartier and Liarte,<sup>44</sup> although only three years old, also suffer from the impact of changing technology. They wrote, “To study the launch strategies of films in other distribution channels, we focused on DVD sales, as other distribution channels are negligible.”<sup>45</sup> This as well is no longer the case. Perhaps future scholars reading this paper will note that our methods are charmingly outdated and inapplicable to understanding the current state of the film industry.

### Implications for the future: Seasonal release schedules

As is demonstrated by both our literature review and our own data, there is a definite seasonal pattern to the financial success of movies, with summer and late fall (November and December) faring the best. Given that fact, it is interesting to look at some of the consequences of, and prognostications for, this trend. In McClintock,<sup>46</sup> DreamWorks CEO Jeffrey Katzenberg opines that these prime seasons can accommodate as many movies as the studios care to release,<sup>47</sup> with movie-going theoretically as plentiful during the week as it is on weekends at these times. If Katzenberg is wrong, then the potential issue McClintock outlines would bolster Einav’s<sup>48</sup> argument that films should be released on a more evenly-spaced schedule year-round. While Einav’s argument follows from the mathematical discipline of game theory, we think that he has failed to account for other factors that make such a plan unworkable. While theoretically, an evenly spaced, year-round release schedule would give all movies room to breathe, this idea ignores the aforementioned noticeable rises and falls in box office take, year after year. If Katzenberg is correct, then summer and Christmas are as elastic as they appear in their viewer capacity.

However, Burr<sup>49</sup> mentions an intriguing bit of history. “After [the 1940s], [good movies released in January] not so much. With the breakdown of the studio system in the 1950s, release patterns began to clump more formally around big weekends, warmer weather and national holidays.”<sup>50</sup> We do not yet have the data to test

Burr's conclusions, but the empirical results of Cartier and Liarte<sup>51</sup> do support his observation that demand during peak seasons is artificially created by movie studios. It is also possible that changes in the film industry are causing monthly release schedules to change. If, as Burr points out, a schedule in which such an established trait as the January dump month has not always existed, then the current trend likely has been artificially created and may change again.

As discussed above, several authors have noted a change in attitude toward August releases. Darling<sup>52</sup> quotes several people who think that even more of the year may come into play as prime release time, and says that the upcoming *Batman V. Superman* release will be a test for this idea:

A few shock waves rippled through Hollywood when Warner Bros. announced recently that the heavily anticipated *Batman V. Superman: Dawn of Justice* will be hitting theaters in March 2016, not in the May sweet spot as originally planned. The move was made to get the jump on the next Captain America sequel from Marvel. "The reality now is there really isn't a bad week to open a movie," Dan Fellman, president of domestic distribution at Warner Bros., told Entertainment Weekly. "We'll be the first one up [in 2016], which is very important, and we'll have six weeks before Captain America comes in."<sup>53</sup>

Though the release of such a high-profile potential blockbuster shows that studios are taking the idea of a larger spectrum of release dates seriously, it is impossible to know right now how this new strategy will pan out, or even if enough movies will be released at unusual times for it to take hold as a full strategy at all.

While the summer season may be expanding, Oscar/prestige film season shows no signs of moving. In a statement echoed by other film writers, as listed in the literature review, Levy explains the clustering of prestige films in late fall: "their [sic] are eight months between May and January, and members [of the Academy] tend to forget good movies. It's conventional wisdom in Hollywood that movies that open early in the year get overlooked at Oscar time."<sup>54</sup> Rottenberg agrees that Christmas is "the traditional climax of Hollywood's year-end prestige season,"<sup>55</sup> but also notes the industry's continuous financial success, giving the example of *The Exorcist*, a Christmas release that was neither a Christmas movie nor an Oscar hopeful.<sup>56</sup>

## Hypotheses and methodology

Both our intuition and our literature review led us to develop expectations about what we would find. These expectations can be formalized into hypotheses, or specific predictions to be tested using empirical data. Clearly stating research hypotheses is a critically important step in quantitative research; otherwise, it is all too easy to for the researcher to see only what s/he wants to see (a situation known as "confirmation bias"), leading to conclusions that have the appearance of scientific support but are merely self-delusions.<sup>57</sup> Once hypotheses are clearly stated, they usually suggest what data collection methods and statistical tools are needed to test them.



Stated in conceptual terms, our hypotheses were as follows:

1. We would find no relationship between viewers' and critics' ratings of the same movies.
2. We would find no relationship between either type of rating and the financial success of a movie.
3. We would find that movies released during the peak periods of summer and year-end perform better financially than those released during so-called "dump months."
4. We would find that movies released during peak periods would be rated more highly by both viewers and critics than those released during "dump months."

These hypotheses suggest certain methods for testing. The first two questions look for relationships between variables, suggesting correlation and regression analysis. The last two questions look for differences in average values among different groups, suggesting a statistical hypothesis test such as a t test. These methods are explained in introductory statistics courses and textbooks. A recommendation for practical advice for working researchers—now out of print but available in many university libraries—is *Reasoning with Statistics* by Frederick Williams and Peter R. Monge.<sup>58</sup> Most quantitative social science researchers are happy to work with researchers in other departments to build collaborations, so they can provide an important resource as well.

Having identified appropriate methods, the final piece required to test our hypotheses is to choose an operational definition for "dump month"—an unambiguous, measurable statement of what that term means in the context the current study. We experimented with several operational definitions, and chose to use the period most commonly quoted in the literature: January, February, August, and September. We also defined the related concept of "peak month"—the months in which it is most profitable to release new titles—as May, June, July, November, and December.

With these pieces in place, we are ready to state specific, testable hypotheses:

1. We would find no meaningful correlation between viewers' and critics' ratings of the same movies.<sup>59</sup>
2. We would find no meaningful correlation between either of the following and gross box office receipts:
  - a. Viewer rating
  - b. Critic rating
3. We would find that the average box office receipts of films released in peak months are significantly<sup>60</sup> higher than those released in dump months.
4. We would find that the average values of each of these two ratings are significantly higher in peak months than in dump months:
  - a. Viewer rating
  - b. Critic rating

With our research questions now translated into clear statistical hypotheses, we are ready to describe our dataset, which we hope will be as useful to others as it has been to us.



## Data and methods

Our research questions focus on the interaction between four variables: release date, gross box office receipts, critical opinion, and public opinion. It was therefore important to find sources of quantitative data to measure each of these constructs. Gross box office receipts and release dates are easily quantified and unambiguous, so assembling a dataset is simply a matter of finding sources. Opinions of film quality are more subjective; but fortunately, film critics and audiences have a long tradition of expressing their opinions in easily-quantifiable forms, such as a scale of 1 to 5 stars, or a thumbs-up/down classification. The fact that these ratings systems are a well-established part of the discourse around Hollywood movies makes film particularly amenable to quantitative study as a cultural product.

It is important to construct a dataset that is as large and complete as possible in our study period - that is, containing as close as possible to 100% of the movies released to U.S. theaters in that period. Our dataset is only as complete as the underlying data sources it comprises, but spot-checking our dataset with release information from the Movie Insider website (<http://www.movieinsider.com>; unrelated to the sites from which we collected our data) shows that, at least for major studio releases, our data do appear complete throughout our study period.

We used three online sources to construct our dataset. Box office receipts and release dates come from Box Office Mojo,<sup>61</sup> a website owned by the Internet Movie Database (IMDb) focused on financial data and news about the movie industry. We obtained data by copying rows from their “Movies A-Z” section.

Viewer ratings come from the Internet Movie Database (IMDb),<sup>62</sup> the largest movie website on the Internet, with more than 200 million unique web visitors per month. The site includes user ratings for each movie in its massive database ( $n = 304,984$ ), rated on a 1-10 scale. Only registered users may vote, and each user may only vote once,<sup>63</sup> so by analyzing vote distributions, we can draw conclusions about individual users’ opinions, by the thousands or even millions. Our final dataset includes more than 268 million votes, an average of 39,400 per movie. The site calculates a summary rating from these individual votes using proprietary methods. We downloaded the site’s complete movie ratings data from a text file available on the site.<sup>64</sup> IMDb also provides distributions of individual user votes for each movie, but we save that data source for future analysis.

Critic ratings come from Metacritic ([www.metacritic.com](http://www.metacritic.com)), a popular aggregator site for movie ratings. We considered Rotten Tomatoes, another popular aggregator and the data source for the original *Slate* article,<sup>65</sup> but chose Metacritic because it provides more finely detailed ratings. Metacritic assigns a percentage rating to individual critical reviews (e.g. a two-star review might count as 40%), then calculates a “Metascore”—a weighted average of critic scores calculated using proprietary methods.<sup>66</sup> As with Box Office Mojo, we obtained our data by copying rows from the site’s “Browse A-Z” list and pasting them into Microsoft Excel.

Our desire to study a dataset as complete as possible was the primary determinant of our study period. Both the Metacritic and Box Office Mojo datasets become much less complete for films released before 2000; balancing these completeness-related factors with our desire to use as many titles as possible led us to consider movies released between 2000 and 2014, inclusive. In addition, we removed from the dataset all re-releases, focusing only on first theatrical runs.

Because our research questions deal with the relationships among movie financial data, public opinion, and critical opinion, it is important to look at titles that have at least some data available in all three sources. However, this task was made more complicated by the fact that movies could have different titles and/or release years in different datasets. This was a particular problem for foreign films, since Box Office Mojo used U.S. titles and release dates, while IMDb used international ones; for example, the 2006 martial arts film starring Jet Li that is listed in Metacritic as *Fearless* and in Box Office Mojo as *Jet Li's Fearless* was listed in IMDb under its original Chinese title, *Huo Yuanjia*.

Our task was therefore to cross-match films, making sure that we matched up the same film in all datasets. This required some Internet research, and was the most time-consuming part of this project. Cross-matching required creating a unique identifier for every film in our dataset; after much experimentation, we used an alphanumeric “key” consisting of the first seven letters of the film’s title, the two-digit year, and the last 4 letters. Cross-matching was done using SPSS’s “MATCH FILES” command,<sup>67</sup> and keys were updated in Microsoft Excel.<sup>68</sup> We iterated our cross-match three times, after which we were confident that either all titles were cross-matched, or else we understood why the cross-match was unsuccessful.

After finishing the cross-match, we were left with a dataset containing release dates, IMDb user ratings, and Metacritic critic ratings for 6,820 movies; of these, 6,673 movies also had gross box office receipts. A last step was to create a new variable containing box office receipts corrected for inflation, calculated by multiplying each film’s receipts by a month- and year-dependent factor from the Consumer Price Index (CPI)<sup>69</sup> of the U.S. Bureau of Labor Statistics.

All analyses were done using IBM SPSS statistical software, version 22 for Mac. SPSS is a powerful statistical analysis program used by many social scientists that is relatively easy to learn. Another key advantage of SPSS is that all commands come with a “Paste” button that automatically preserves a reproducible history of commands in a text file. The main disadvantage is cost, but academic licenses are available,<sup>70</sup> and many campuses have existing site licenses.

Equally powerful free alternatives are available, such as R and Python, albeit with much steeper learning curves. Another alternative is that all analyses described here can be done, in a more roundabout way, using Microsoft Excel, which comes standard on many computers. Our complete dataset is available by contacting the primary author (Berg).

**Table 1.** Number of films released per year in our dataset.

Year	Movies
2000	312
2001	314
2002	400
2003	396
2004	461
2005	475
2006	503
2007	495
2008	438
2009	397
2010	424
2011	502
2012	535
2013	579
2014	589

## Results

### *Financial data: Release times and box office receipts*

Table 1 shows the number of movies in our dataset by year of release. Most new films are released into theaters on Fridays ( $n = 5,970$ ; 88%), with the only other day with an appreciable number of releases being Wednesday ( $n = 725$ ; 11%).

Table 2 shows the number of movies released by month over our entire study period. The number of theatrical releases per month is roughly the same throughout the year, as is required to keep a fresh supply of movies available in theaters year-round. Slightly fewer movies are released in January, and slightly more are released in the fall.

The Box Office Mojo dataset is missing some values for gross box office receipts; the number of films with receipts listed is 6,673, which is 98% of all movies in the dataset. A quick review of the few titles that are missing shows, as expected, they are all small releases whose data would be highly unlikely to change any of our conclusions.

**Table 2.** Number of films released per month.

Month	Movies	Percent of All Releases
January	361	5.3
February	471	6.9
March	615	9.0
April	639	9.4
May	563	8.3
June	540	7.9
July	515	7.6
August	642	9.4
September	703	10.3
October	693	10.2
November	555	8.1
December	523	7.7
Total	6820	100.0

**Table 3.** Number of films released per studio in our entire dataset.

Studio	Movies	Top-Grossing Title	Total Adjusted Gross (\$ Billion December 2014)
Warner Brothers	296	<i>The Dark Knight</i>	\$23.3
Buena Vista	226	<i>The Avengers</i>	\$21.5
20th Century Fox	236	<i>Avatar</i>	\$18.7
Universal	231	<i>Despicable Me 2</i>	\$17.4
Sony	200	<i>Spider-Man</i>	\$15.2
Paramount	167	<i>Iron Man</i>	\$13.3
New Line	88	<i>The Lord of the Rings: The Return of the King</i>	\$5.3
Paramount / DreamWorks	39	<i>Transformers: Revenge of the Fallen</i>	\$5.1
Lionsgate Entertainment	122	<i>The Hunger Games: Catching Fire</i>	\$4.3
DreamWorks	46	<i>Shrek 2</i>	\$4.1
Miramax	142	<i>Chicago</i>	\$3.0
Screen Gems	65	<i>The Vow</i>	\$2.6

After correcting for inflation to express all values in December 2014 dollars, the movies in our dataset combined for gross box office receipts of more than \$160 billion, underscoring once again the enormous importance of the American film industry. What companies are producing these films? [Table 3](#) shows the top twelve studios represented in our dataset, in terms of total adjusted gross for all movies released.

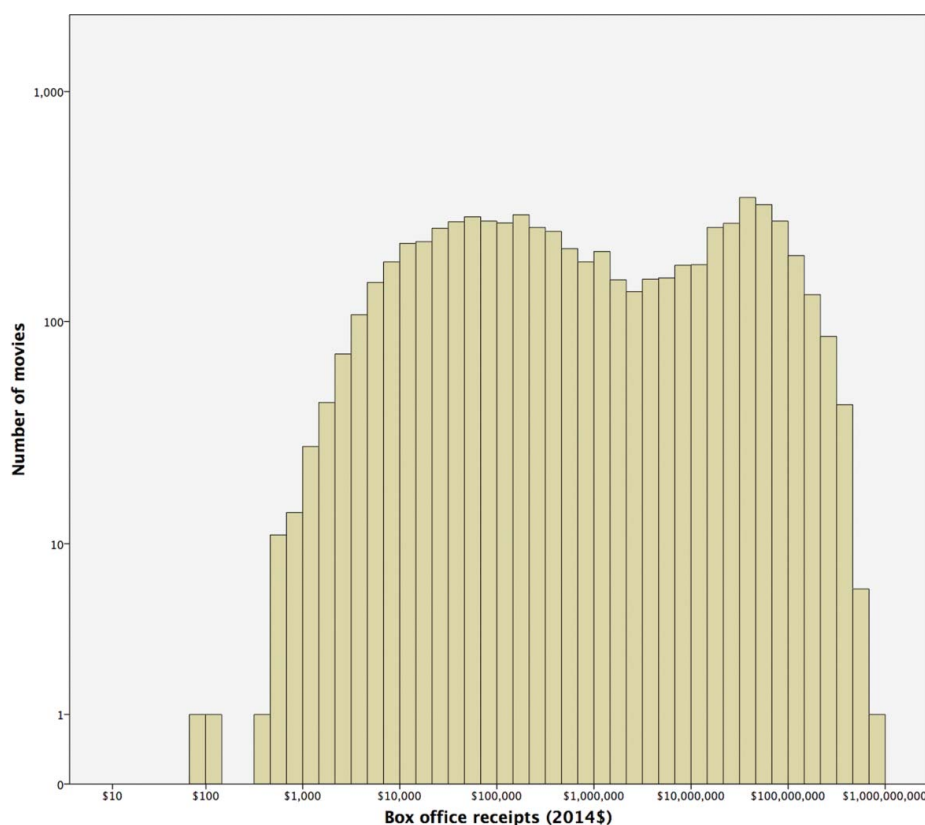
These 12 studios together are responsible for more than 83% of all gross box office receipts during our study period. The limitations of our dataset preclude making comparisons between studios; for example, Screen Gems is a division of Sony Pictures, but they are listed as separate rows in [Table 3](#). However, [Table 3](#) does lend support to the common observation that only a few studios contribute most of the industry's financial returns.

This observation shows that, like many economic and social phenomena, film box office receipts are extremely top-heavy, with a few titles responsible for most receipts, and a large number of titles generating much smaller receipts. This fact means that the analysis of movie receipts data is clearer if receipts are sorted into bins of unequal size. [Table 4](#) shows the number of movies in each of four bins (again, figures are in inflation-adjusted December 2014 dollars):

1. Box office receipts less than \$100,000
2. Greater than (or equal to) \$100,000 but less than \$10 million
3. Greater than (or equal to) \$10 million but less than \$100 million
4. \$100 million or more

**Table 4.** Number of films in our dataset for which the gross box office receipts reach four different levels (bins).

Bin	Movies	Percent of All Movies
Less than \$100,000	2,138	31.3
\$100,000–\$10 million	2,428	35.6
\$10–100 million	1,646	24.1
More than \$100 million	461	6.8
Total	6,673	97.8
Missing	147	2.2



**Figure 1.** Logarithmic histogram of box office receipts ( $n = 6,673$ )

Another way to show such top-heavy data is to plot a histogram of the logarithm (base 10) of both observed values and frequency counts. [Figure 1](#) shows this plot for box office receipts in our dataset. A histogram is a graph showing a range of data values along the bottom (the x-axis). The range is divided into smaller groupings called “bins”; above each bin is a bar whose height shows how many times the values in that bin occur in the dataset.

In most histograms, all the bins have equal widths, but this histogram is somewhat different. Each tick mark on both the horizontal and vertical axes shows values a factor of ten larger than the previous tick. Thus, each bin along the x-axis actually contains ten times the range of data values of the previous bin, and each bar that reaches an additional tick mark on the y-axis contains ten times as many films as the lower bar.

The bar to the far left shows the one film with receipts less than \$100: *Storage 24*, a 2013 British sci-fi horror film that brought in \$73 in the United States. The bar on the far right is *Avatar*, which brought in more than \$800 million (inflation-adjusted).

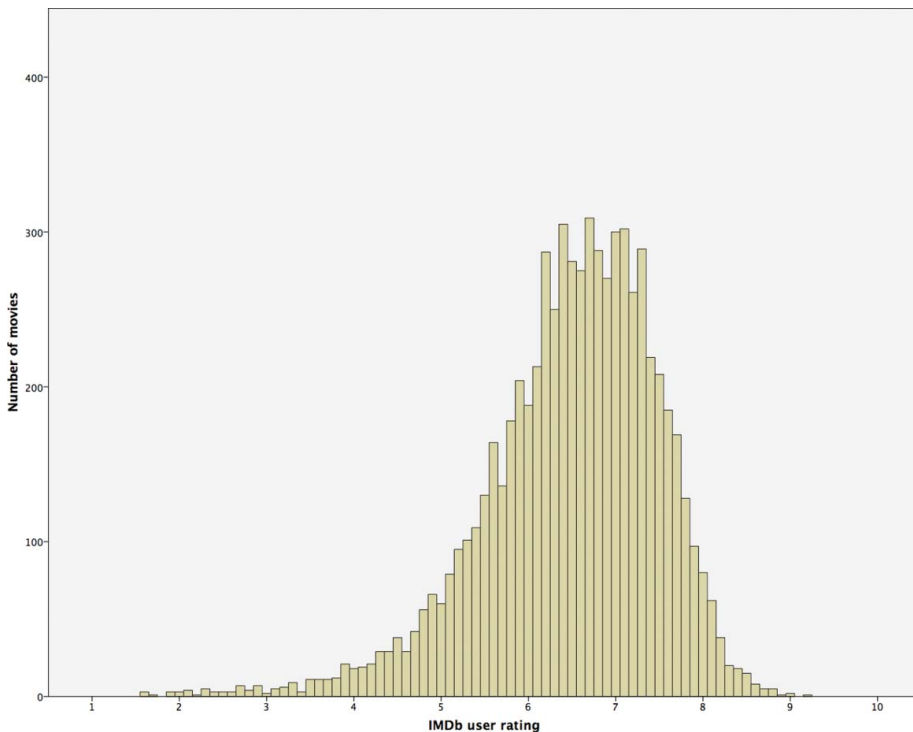
Keeping in mind the caveat that logarithmic scales introduce distortions in plots, we note the fascinating bimodal (double-peaked) distribution of receipts, with one peak of around \$200,000 and another of around \$30,000,000. We

suspect this is due to the distribution gap between wide release and limited release movies, with the \$200,000 peak representing the average for a film that only screened in a few places and the \$30,000,000 peak representing the average for wide release movies, but more careful research is needed to draw conclusions.

### **Critical judgment data: User and critic ratings**

Financial returns are the data most important to studio decisions, but are only one part of the story of how our culture understands the film industry. The other important consideration is the perceived artistic merit of the films. We do not make any artistic judgments ourselves; rather, we consider the opinions of two groups: viewers—specifically, registered users of the Internet Movie Database (IMDb)—and professional film critics. Data sources for each are explained in the Data and Methods section.

As described in the Data section, IMDb provides user ratings only for those movies that have received more than five user votes; but many movies have received many more votes, all the way up to *The Dark Knight*, which has more than 1 million votes. Our dataset of user ratings contains 268,713,871 individual votes. As described in the Data and Methods section, we know that within an



**Figure 2.** Histogram of IMDb user ratings ( $n = 6,820$ )

individual title, each user can vote only once; but of course users may vote on more than one movie, so we have no way of knowing how many people contributed these 268,713,871 votes. But it is safe to say that our dataset of user ratings is one of the largest ever used in a study of this type.

Figure 2 shows the histogram of user scores. User scores from 1 to 10 are shown along the bottom of the graph (on the x-axis). The number of movies with that score is shown along the left side (on the y-axis). The leftmost bar shows the three movies with the minimum user score of 1.6. The rightmost bar shows the two movies with the maximum user score of 9.0.<sup>71</sup>

In spite of the uneven distribution of the number of votes per title (from five to more than one million votes contributing to the ratings), the user ratings show the expected pattern of a bell curve, more precisely described as a “normal distribution.” The mean of the distribution—in other words, the average user rating for a movie on IMDb—is 6.49 out of 10. The standard deviation of the distribution of user ratings—a common measure of how much the distribution clusters around the mean value—is 1.00. The definition of standard deviation means that about two-thirds of ratings lie within one standard deviation of the mean; that is, between 5.49 and 7.49.

What about the ratings given by professional film critics? Figure 3 shows the distribution of “metascores” from the Metacritic website for films in our dataset ( $n = 6,820$ ). The histogram is shown in the same way as in Figure 2, but note that the

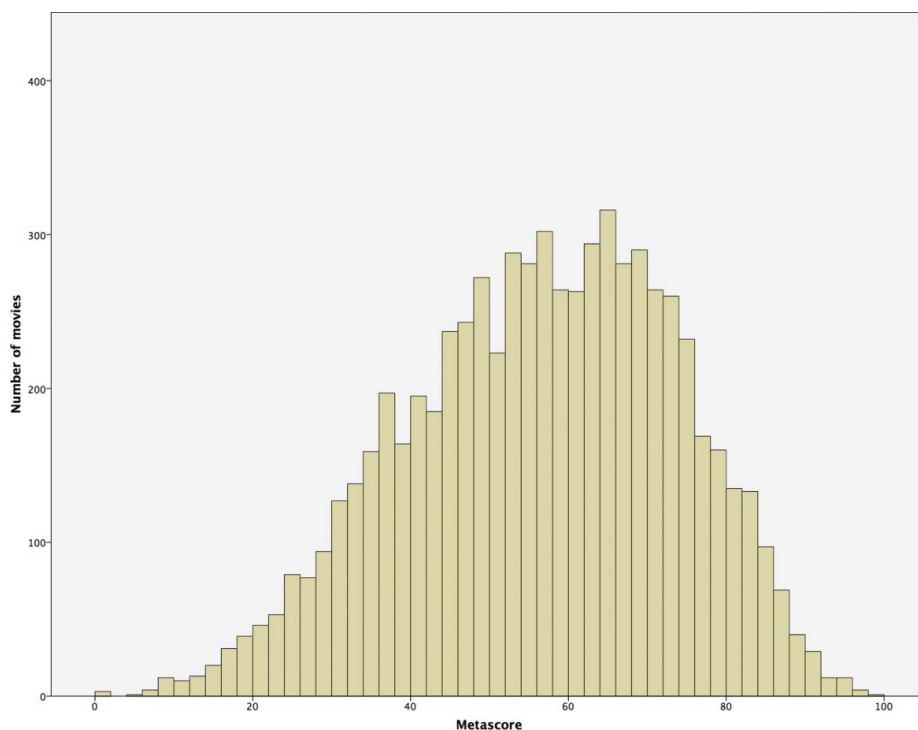
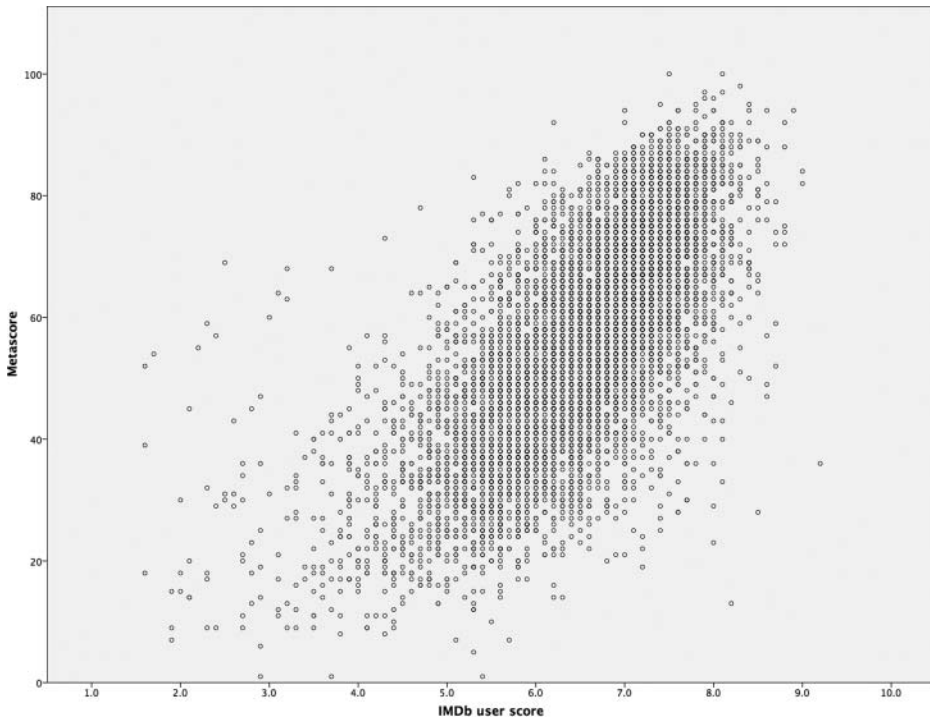


Figure 3. Histogram of Metascore critic ratings ( $n = 6,820$ )





**Figure 4.** Scatterplot of ratings by users and critics (n = 6,820)

x-axis (Metascore) now runs from 1 to 100 instead of 1 to 10. Three films have Metascores of 1 and two have Metascores of 100.<sup>72</sup> The mean Metascore is 56, and the distribution has a standard deviation of 17, meaning that about two-thirds of films have Metascores between 39 and 73.

Comparing the distributions of ratings given by viewers and critics (Figures 3 and 4) reveals some fascinating trends. The mean film rating given by critics (56/100) is lower than the mean rating given by viewers (6.49/10).<sup>73</sup> Furthermore, the spread of ratings assigned by critics (standard deviation of 17/100) is wider than the spread of ratings assigned by viewers (1.00/10). These observations mean that, on average, critics assign lower scores than viewers, and that they are more likely to give scores that are higher or lower than their average for all movies.

There could be a number of possible explanations for these observations. Perhaps critics have higher artistic standards for film, and are better able to distinguish small differences in quality. Perhaps critics have learned, consciously or unconsciously, that controversy sells copy, and are therefore more likely to assign scores closer to the extremes. Perhaps the process of becoming a professional film critic selects for people with more extreme opinions. Perhaps the effect is due to some inherent difference in the way in which IMDb and/or Metacritic calculates summary scores, since the specific methods the sites use are not publicly available. Quantitative methods are not ideal for answering such questions of underlying causation, and certainly our study

cannot distinguish between these possibilities. We encourage other researchers to continue studying possible reasons for the observed differences in the means and standard deviations of viewers and critics.

### ***How do critical and public opinions of movies correlate?***

Having assembled our dataset and understood its basic statistical properties, we now consider our research questions and hypotheses. Our first research question compares critical and viewer opinions to one another; our second compares both types of opinions to box office receipts. The ideal methods for answering such questions—those that ask how data values change in relation to one another—are the methods of correlation and regression.

A first step in any such study is to graph the variables of interest against one another to look for trends. [Figure 4](#) shows such a graph—a scatterplot with IMDb user scores along the bottom (x-axis) and Metascores along the left side (y-axis). Each point on the graph represents a single movie, showing both types of scores. For orientation, points in the top right corner represent movies that scored highly with both web users and critics; films in the middle left scored moderately well with critics and poorly with users. The graph shows, at a glance, the relationship between user and critic scores for our entire big dataset at once.

The clearest trend in [Figure 4](#) is the broad cloud of points in the middle, trending roughly diagonally from lower left to upper right. That cloud shows that, while the relationship is by no means perfect, ratings by users and critics do tend to go together.

We hypothesized that user and critic scores would not agree, based on the financial success of such critically reviled movies as Michael Bay's *Transformers* series. However, [Figure 4](#) shows that this is clearly not the case.

Can we say anything more quantitative about the relationship between user and critic scores? Yes, using the methods of correlation analysis, which are essentially a formalized way of measuring the strength of the linear trend shown graphically in [Figure 4](#). The correlation is a number ranging from -1 to 1 giving the relationship between two variables. A correlation of 1 indicates that the two variables go together perfectly—when one increases in value, the other always increases at a predictable rate. A correlation of -1 means that the two variables go together in a perfectly opposite way—when one increases, the other *decreases* at a predictable rate. A correlation of 0 means that the values are completely unrelated. There are no hard-and-fast rules for interpreting correlation values, but a common guideline is that a correlation greater than 0.5 (or less than -0.5) is considered a large correlation.<sup>74</sup>

The correlation between Metacritic critic scores and IMDb user scores in our dataset ( $n = 6,820$  movies) is .684. That value is by no means a perfect correlation, but it is large enough that we can make a clear case that critics and the public tend to give the same movies similar scores.

**Table 5.** Correlations among user score, Metascore, and gross box office receipts for films in our dataset.

Variable	IMDb user score	Metascore	Gross Box Office Receipts	
			Raw	Inflation-adjusted
IMDb user score	—	.684	.059	.055
Metascore	—	—	.010	.077

This fact is all the more intriguing given that the results of our next research question showed variables that do *not* correlate: neither the user score nor the critic score correlates to any appreciable degree with gross box office receipts, whether considered as raw amounts or adjusted for inflation. Table 5 shows the correlation matrix (the correlation between every variable and every other variable) for user score, critic score, and both types of box office receipts.

Thus, while the statistical evidence in our dataset proved our first hypothesis incorrect, our second hypothesis was supported.<sup>75</sup> Despite the fact that viewers tend to agree with critics, poor critical reception is not associated with poor box office performance. Since viewers agree with critics a far greater amount of the time than they fail to see a movie at all because of what the critics say (because to rate a film, viewers must have seen it first), we can infer that there is virtually no relationship between what the critics think of a movie and whether or not it makes money at the box office, although we cannot draw conclusions about why.

Cartier and Liarte say that, just by being out, a highly rated movie gets people to go to the movies more.<sup>76</sup> However, rather than contradict our findings, this underscores the idea that there is no connection between critical appraisal and the decision to watch a particular movie. Despite Cartier and Liarte's observation that the viewing public shows some interest in seeing critically acclaimed movies, the fact that that interest boosts attendance for *all* films still demonstrates the lack of importance given to the critical appraisal of any one movie.

### ***The significance of release times to box office take, on average—AKA “dump months”***

We next look at how the timing of a film's release is associated with both its box office performance and its artistic merit as measured by users' and critics' ratings. Table 6 shows the data we use in our analysis throughout this section, giving the number of movies included and the means of each variable.

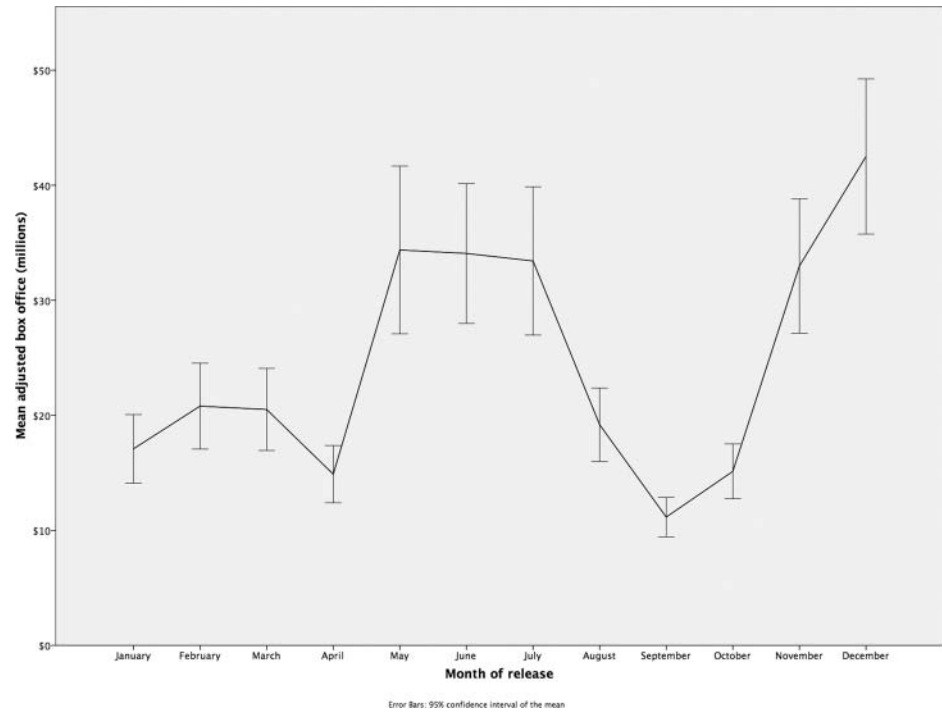
Figure 5, which plots the mean box office receipts for films released in each month, clearly shows the different levels of audience demand in different months. Error bars show 95% confidence intervals for the mean gross box office receipts for all films in our dataset that were released during a given month. Note that not only are mean receipts lower in some months than others, but the difference is so big that the error bars do not

**Table 6.** Films in our dataset by month of U.S. theatrical release.

Month of Release	Titles Included	Gross box office (\$ Millions December 2014)	IMDb User Score	Metascore
January	348	\$17.1	6.30	51.4
February	462	\$20.8	6.37	53.3
March	606	\$20.5	6.50	55.2
April	627	\$14.8	6.40	55.2
May	548	\$34.4	6.53	57.4
June	528	\$34.1	6.49	57.4
July	513	\$33.4	6.53	57.9
August	623	\$19.2	6.35	54.7
September	690	\$11.2	6.46	54.6
October	674	\$15.1	6.54	56.2
November	543	\$33.0	6.59	57.9
December	511	\$42.5	6.72	59.5
Overall	6,673	\$24.1	6.49	56.0

overlap. The clear difference in receipts shown in [Figure 5](#) certainly provides strong evidence for the idea that different months have very different levels of audience demand.

Does the difference in box office receipts translate into to a difference in perceived artistic quality? There is indeed a relationship between release month and ratings, but not in a straightforward way. The pattern of monthly variation of scores is very similar in the user and critic scores, but both are different from pattern of box office receipts. Neither of these observations is surprising given our observation, described in the previous section, of a fairly high correlation between



**Figure 5.** Mean box office by month of release ( $n = 6,673$ )

user and critic scores, and of the lack of correlation between either type score and gross box office receipts.

Finding these differences among movies by month of release are important, but our research questions are somewhat more specific. We investigate the phenomenon of “dump months.” As noted in the introduction, studios face the challenge of releasing their stable of films such that some movies are available to play at U.S. theaters throughout the year, even though they know that underlying audience demand varies widely over the course of the year.<sup>77</sup> Conventional wisdom in cultural studies says that they solve this problem by “dumping” certain movies—releasing those expected to do poorly at the box office—during months that are less profitable for the introduction of new titles.

The so-called “dump months” have been identified as different months by different authors. January is the most notorious of all of them, as evidenced by the amount of articles in the literature that stress January as *the* dump month.<sup>78</sup> The *Slate* article that inspired this research identified February as a key dump month,<sup>79</sup> and other authors have identified April<sup>80</sup> and September.<sup>81</sup> Authors have disagreed about whether March, August, and/or October should be included as well.<sup>82</sup> Regardless, conventional wisdom and scholarship readily agree about which months are NOT dump months: the summer (May/June/July) and the end of the year (November/December). We experimented with a few definitions of dump months and found the clearest signal came from the latter definition: operationalizing “dump months” as the seven months not in the summer or year-end seasons: January through April and August through October.

With that definition in hand, we are ready to perform statistical tests on our hypotheses related to the timing of film releases. Our hypotheses, restated from the hypothesis section are:

- Hypothesis 3. The mean adjusted gross box office receipts will be lower during dump months (January through April and August through October) than during peak months (May through July and November through December).
- Hypothesis 4a. The mean IMDb user score will be lower during dump months than during peak months
- Hypothesis 4b. The mean critic score on Metacritic will be lower during dump months than during peak months

Hypothesis (3) means that films released during dump months bring in less money (at least in terms of gross receipts) than films released during peak months—although our tests cannot make any claims about why this would be the case (e.g., due to decreased audience demand, lower-quality films, reduced studio marketing, or any other factors). Hypothesis (4a) means that films released during dump months are perceived to be of lower quality by IMDb users; again, our tests cannot give any evidence about why this should be the case. Hypothesis (4b) means that the films released during dump months are perceived to be of lower quality by critics (again with the same interpretive caveats).

**Table 7.** Data for movies released during either dump months and peak months.

Group	Number of Movies Included	Box Office Receipts (\$ Millions 2014)	IMDb User Score	Metascore
Peak month	2,643	\$35.4	6.57	58.0
	Standard deviation in group:	\$76.2	0.99	16.7
Dump month	4,030	\$16.7	6.43	54.6
	Standard deviation in group:	\$35.2	1.01	17.4
All months combined	6,673	\$24.1	6.49	56.0
	Standard deviation in group:	\$55.9	1.00	17.2
<i>p</i> value		<.001	<.001	<.001

The way to test these hypotheses is to compare the means of each variable in each of the two groups (dump month or peak month), then use statistical procedures explained below to evaluate the significance of those differences. Table 7 shows the observed means of each variable in both groups, as well as their observed standard deviations.

Given the observed differences in means between the two groups for each of the three variables, what conclusions can we draw? Statistical hypothesis testing allows us to draw tentative conclusions about the likelihood that the results we observe are due to chance. The concept behind the approach is as follows: if we assume for a moment that the underlying mean receipts and scores between the two groups are actually the same (this is referred to as the “null hypothesis”), then what is the probability that we would have observed a difference at least as large as the one we actually observed? Importantly, note that this is *NOT* the same as the probability that the assumption of underlying equal means is true; rather, it means that either we have witnessed an unlikely event under that assumption, *OR* that our assumed null hypothesis was incorrect and the underlying means actually are different.

Traditionally, researchers identify a threshold probability at which they will assume that the observed differences are just too unlikely to be due to chance, and they will choose the second possibility—that the assumed null hypothesis of equal means was incorrect. This is referred to as “rejecting the null hypothesis.” When researchers reject the null hypothesis, they often state the probability under the null hypothesis of observing at least the differences they observed; this probability is known as a “*p* value.” Running a one-tailed *t* test on our observed data in SPSS allows us to reject the null hypothesis of equal group means at the  $p < .001$  level for all three variables we tested. We therefore consider all our hypotheses to be supported.

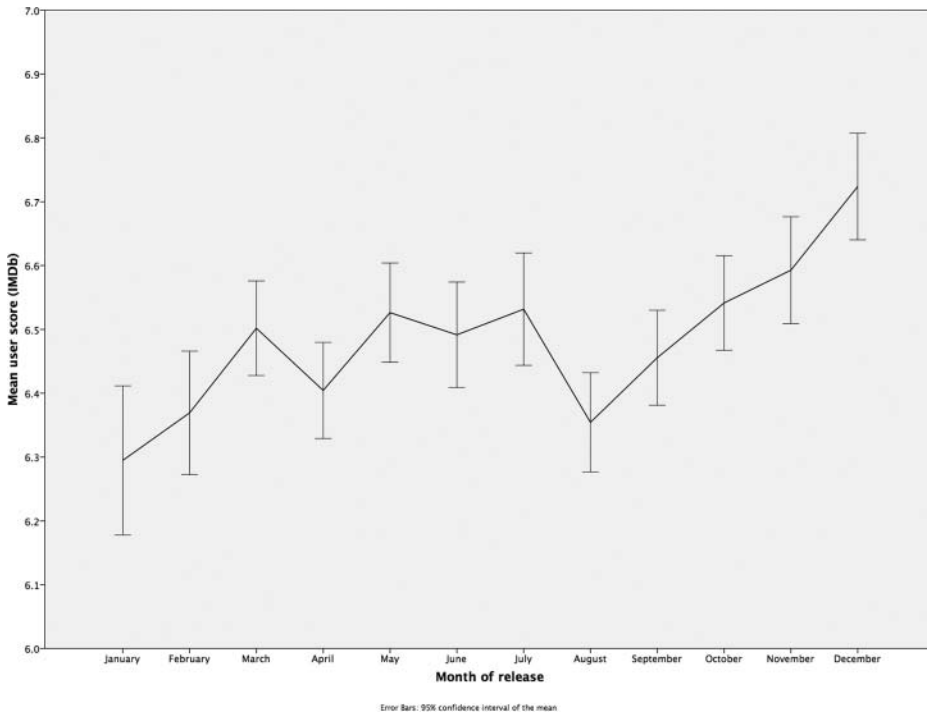
Statistical significance testing is a fundamental consideration in quantitative research, but it does not provide the entire story. First, statistical significance will always improve with additional data (quantitatively, the *p* value will always decrease as sample size increases); with 6,820 movies in our dataset, our impressive-looking *p* values could simply be an artifact of large sample size. Second, “statistically significant” is not at all the same as “scientifically significant” or “important” or “interesting.” Those ideas are less easily quantified, must be

evaluated using multiple lines of evidence. But our results are certainly strong enough to clearly indicate that “dump months” are a real phenomenon, and that the month of a film’s release is associated with real differences in its financial performance and artistic quality.

What can we say about the American film industry from these results? Although the literature on dump months tends to emphasize January, our research shows that January, or at least its financial fortunes, is not exceptional. Instead, there is a pattern of dump periods alternating with more successful movie release periods. Not only do films released during dump periods tend to bring in less money at the box office, they also tend to be rated lower by both critics and users.

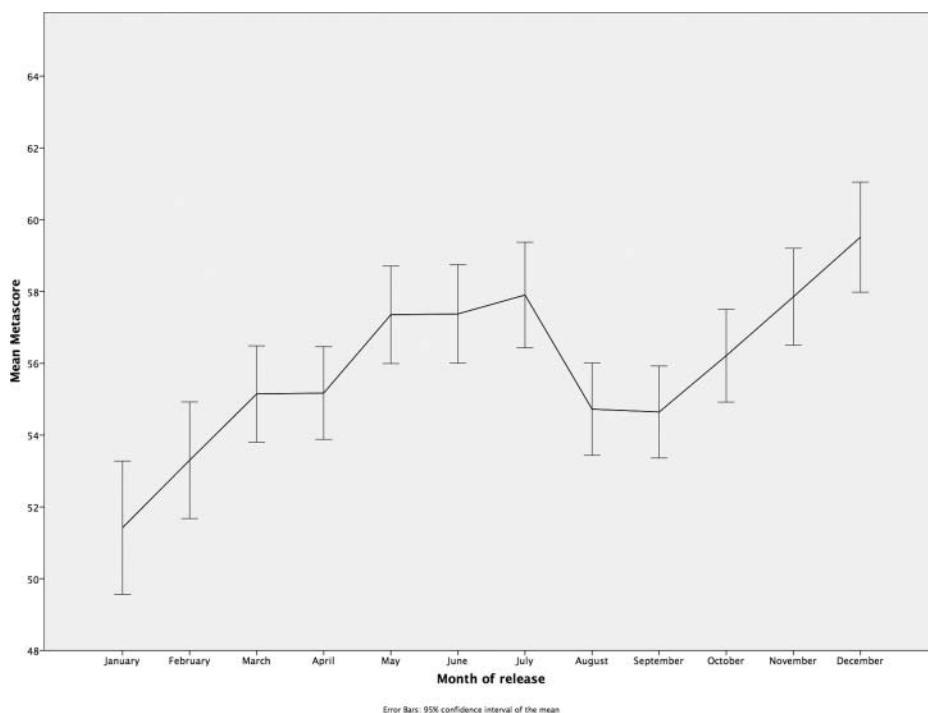
The *Slate* article that inspired this paper claimed that, statistically, February was the worst month critically; our research gives that dubious honor to January instead. February is not the worst commercial month either; there is no clear evidence that the mean domestic box office gross receipts in February are unusually low. In fact, [Table 6](#) shows that the mean domestic box office (in millions) in February (\$21 M) is higher than in January (\$17 M), April (\$15 M), September (\$12 M), or October (\$15 M).

While [Table 6](#) shows that all of the dump months take a critical drubbing, January comes out the worst in all of our assessments. Although we cannot perform a direct comparison between our results and those of *Slate*,<sup>83</sup> we do have multiple results that do not agree with the *Slate* data.



**Figure 6a.** Mean user score by month of release ( $n = 6,820$ )





**Figure 6b.** Mean critic score by month of release ( $n = 6,820$ )

IMDb user ratings by release month (Figure 6a) show some interesting variability. The mean user ratings are highest in December (6.8) and lowest in January (6.2). Metacritic critic ratings by month (Figure 6b) show similar patterns to IMDb user ratings by month. Again, the highest month is December (mean 59/100) and the lowest is January (mean 51/100). This observation is again not surprising, since we know there is a strong correlation between user scores and critic scores (as noted previously), and it provides us with both critical and viewer assessments of January as home to worse movies than February. If we are correct, then January has matching abysmal critical and commercial performances, while by both measures films improve somewhat in February. There is also an increase in both box office take and critical assessment in September and October.

This positive correlation for films in September and October is especially interesting because, as we have previously noted, there is an overall lack of correlation between critical assessment and commercial success. Though we cannot say for certain why critical assessment and box office take both increase in those months, there are several possibilities to consider. One is that the movies dumped in those months are of such poor quality that they do not appeal to the general public, whether or not the critics like them. This would serve as evidence that studio executives can indeed spot a movie unlikely to make money, as Burr says.<sup>84</sup> However, it is important to note that because those movies are pre-emptively deemed failures, they receive minimal advertising and promotion, possibly resulting in a self-fulfilling prophecy. Cartier and Liarte note the importance of

advertising: “[advertising] budgets [sic], therefore, even more than the number of films, is the variable used by the studios to capture demand,”<sup>85</sup> but these are “the films no studios believe in or care about—the stuff that doesn’t get screened for critics, the stuff that barely gets promoted beyond blurbs from obscure websites and suspicious raves from local TV chefs and weathermen.”<sup>86</sup>

The water is further muddied by August, which does not feature this joint increase. It is possible that the ascendant trend of viewing August as a legitimate release period starting in 2007—as described in our literature review—has skewed the box office take during the back half of our sample period; that is to say that the data does not behave as it does in other dump months because August is not a dump month over half of our study period. Other changes in monthly performance over our study period are possible; we will explore this question more in the future.

Given the nature of dump month films, it is worthwhile to ask why they are released in theaters at all. Meslow explains that some films have a contractual obligation for theatrical release.<sup>87</sup> However, the stronger motivation seems to be the unexpected financial success of some dump month movies, January movies in particular. *You Got Served*, *Hostel*,<sup>88</sup> *Paul Blart: Mall Cop*, *Taken*,<sup>89</sup> and *Chronicle*<sup>90</sup> are all relatively recent movies dumped in January that unexpectedly made a lot of money.

## Results/interpretation/conclusions

Our research has provided quantitative data on a number of trends regarding movies and their theatrical release dates, financial performance, and critical and viewer reception. First, we have identified a correlation between critical opinion and viewer opinion. This was a surprising result, as the continued presence of movies that receive poor critical reception argues in favor a model in which poor reviews do not impact economic success. We had extrapolated from that idea that critics and viewers judged movies differently.

However, this turned out not to be the case, which gave tremendous importance to a second relationship we quantified, the lack of correlation between the critical reception and financial success of a movie. As previously stated, we had been anticipating this, but the why of the matter became a mystery when we also found that user ratings do not correlate with box office receipts. Thus, we cannot explain the lack of correlation between critic ratings and box office receipts by reasoning that viewers think differently from movie critics. The data gathered by our study do not permit us to know why this is. However, it does seem clear that, while critical assessment would serve a majority of the movie-going public well as a barometer for whether or not they will enjoy a film, they either ignore or do not consult it.

We have also added quantitative proof to the anecdotal evidence given by film researchers and critics that January is a critical and commercial dead zone, the worst of the dump months. Our data places January’s box office take and mean critical rating as the worst of the year. However, we have also charted other months that habitually feature poor performance, both financially and critically. Some of

these months were expected, such as September, which (although not as infamous as January) arose in the literature review.

However, April is rarely spoken of in such dire terms, and yet we found that year after year, box office receipts dropped precipitously in that month. April is certainly a lull in between winter and summer releases, but the critical ratings are not nearly as bad as they are for January or September. Why ever that may be, April certainly deserves a place of study alongside January, February, and September when considering lulls in the annual release strategies of movie studios. There is some evidence that studios consider this as well: Rottenberg says that the release of *American Pie* was moved from spring to summer when it did unexpectedly well in test screenings.<sup>91</sup>

### **Acknowledged possible issues with/limitations of our data**

Although we have done our best to account for disparities and imbalances, there are some factors affecting financial performance that we did not take into account in our data. One is the matter of how many screens a movie is released on. Art-house films, which traditionally have a high critical rating but low box office take, are released on significantly fewer screens than mainstream pictures. This is something of a self-fulfilling prophecy, with studios assuming the film will not make much money, so they do not give it a mass release, which helps guarantee that it will not make much money. We have since obtained data on the number of screens on which movies were shown, but we save that analysis for a future paper. For this work, though, focusing on gross receipts without considering screens unavoidably leaves some movies on an unlevel playing field.

Another potential issue is the nature of IMDB ratings. IMDB ratings are limited to the people who make the effort to register for the website, which is not necessarily an unbiased sample of U.S. moviegoers. There is a perceived wisdom that people who bother to comment on websites are the happiest and most unhappy about things—if so, we are unavoidably missing the in-between viewer opinion.

Titles are an issue as well; movies are sometimes retitled for different markets. While we made the best possible effort to collate the disparate names (as described in the Data section), there may be foreign movies released in the U.S. under different titles that slipped through our information gathering net, thereby introducing small unknown biases into our data and analysis.

### **Directions for future research**

Although we were able to reach a number of interesting conclusions through our research, many intriguing questions came up that we were unable to answer—as is the case in all good quantitative research. The following are some directions for future research drawn from the results of this article.

1. In line with most traditional discussions of Hollywood release schedules, we have considered releases by month, exploring the phenomenon of “dump months.” But, as Cartier and Liarte<sup>92</sup> point out, weeks might be a better unit of analysis for

- film releases. This might particularly be the case for holiday releases; for example, we may learn more about the impact of release schedules by analyzing Memorial Day releases as their own week rather than grouping them with the rest of May.
2. A sociological look at why the viewing public apparently does not consult critics in deciding what to watch—a question that is all the more worth asking now that we know their opinions tend to converge. Panned movies make money, but viewers end up agreeing with critics. Why is this the case?
  3. The increasing prevalence of simultaneous release dates for theater and on-demand will surely alter the way in which box office take is counted, and possibly how the financial success of a movie is evaluated altogether. It will be worth reevaluating our research in light of this model once these changes become more apparent.
  4. Burr asserts that there used to be artistically successful<sup>93</sup> January releases during the studio system, up through the 1940s; and that therefore the current release schedule, with its dump months, big seasons, and lulls, was manufactured by the studios and is not a natural pattern. It would be interesting to run a variation of this study for the studio system days to see if he is correct; and if so, just how much of a variation there is between patterns then and today. Data for these older films should be available online, although building a complete dataset may prove much more difficult.
  5. Darling says that the annual release structure is changing, specifically citing August, but also some other months. Such changes could potentially throw off the pattern of dump months and good months we observed. Tracking whether these changes pan out, and in what fashion, over the next several years could be very interesting.
  6. Sochay cites Litman alone and Litman and Kohl together as saying that “each additional star [rating] had a significant positive impact on rentals.”<sup>94</sup> These datasets do have issues—Sochay does not explain Litman and Kohl’s methodologies (most pressing, where do the star ratings come from, and how can they prove they increased rentals?), and the studies date back to 1983 and 1989, with Sochay’s assessment of them going back to 1994. Also, the rental structure has changed drastically in the intervening years, in ways previously discussed. Nonetheless, the idea that critical assessment plays a greater role in movie selection post-theatrical release is a very intriguing one that is worth looking into.

## Notes

1. Bordwell and Carroll, *Post-Theory: Reconstructing Film Studies*.
2. Burr, “January Is Hollywood’s Very Own Leper Colony,” in *New York Times Magazine*.
3. Kuchera, “How Does a ‘Terrible’ Movie Make \$300 Million in Three Days?” in *Polygon*.
4. Emerson, Jim. “Why Don’t Critics, Oscars & Audience Agree?” *Scanners: With Jim Emerson* Roger Ebert.com.
5. Kirk and Thompson, “Confirmed: February Movies Suck,” in *Slate*.

6. O'Neal, "February Statistically Proven to Have the Shittiest Movies," at A.V. Club website. February 6, 2014.
7. Basuroy, Subimal and Ravid, "How Critical Are Critical Reviews? The Box Office Effects of Film Critics, Star Power, and Budgets," in *Journal of Marketing*.
8. Reinstein and Snyder, "The Influence of Expert Reviews on Consumer Demand for Experience Goods: a Case Study of Movie Critics," in *Journal of Industrial Economics*.
9. Plucker, Kaufman, Temple, and Qian, Meihua. "Do Experts and Novices Evaluate Movies the Same Way?" in *Psychology and Marketing*.
10. McClintock, "Summer's Brutal Animation War," in *Hollywood Reporter*, p. 83.
11. *Ibid.*, p. 2.
12. Sochay, "Predicting the Performance of Motion Pictures," in *Journal of Media Economics*.
13. *Ibid.*, p. 2.
14. *Ibid.*, p. 1.
15. *Ibid.*, p. 8.
16. Litman and Kohl, p. 4.
17. *Ibid.*, p. 4.
18. Rottenberg, "The Imperfect Science of Release Dates," in *New York Times Magazine*, p. 64.
19. *Ibid.*
20. *Ibid.*
21. Bart, "Fall Film Flurry Stirs Arty Anxieties," in *Variety*.
22. *Ibid.*, p. 4.
23. Cartier and Liarte, "Impact of Underlying Demand and Agglomeration of Supply on Seasonality: The Case of the Hollywood Film Industry," in *International Journal of Arts Management*.
24. *Ibid.*, p. 21.
25. *Ibid.*
26. *Ibid.*, p. 21, p. 25.
27. *Ibid.*, p. 19.
28. Einav, "Not All Rivals Look Alike: Estimating an Equilibrium Model of the Release Date Timing Game," in *Economic Inquiry*.
29. Darling, "August: Now a Hot Month For Blockbusters," at dfw.com.
30. Burr, January.
31. Maslin, "Is January the Cruellest Month?" *New York Times*.
32. Legel, "Why is January a Cinematic Dumping Ground?" at Film.com.
33. Salisbury, "Why Oscar Season Is Hollywood's Bad Movie Dumping Ground," at Hollywood.com.
34. Burr, January, n.p.
35. Darling, August.
36. Kirk and Thompson, Confirmed.
37. LaMarche, "A Cinematic Dumpster Dive," in *Vegas Seven*.
38. Rottenberg, Imperfect Science.
39. *Ibid.*, p. 66.
40. "The August Movie: A Theory of Awfulness," at *Vulture*.
41. Salisbury, "Why Oscar Season Is Hollywood's Bad Movie Dumping Ground," Hollywood.com.
42. Singer, "Why I Love Going to the Movies In January," in *Criticwire*.
43. Sochay, Predicting.
44. Cartier and Liarte, Impact.
45. *Ibid.*, p. 20.
46. McClintock, Summer's Brutal.
47. *Ibid.*, p. 83.

48. Einav, Not All Rivals.
49. Burr, January.
50. Ibid., n.p.
51. Cartier and Liarte, Impact.
52. Darling, August.
53. Ibid., n.p.
54. Levy, *Oscar Fever: The History and Politics of the Academy Awards*, p. 311.
55. Rottenberg, Imperfect Science, p. 64.
56. Ibid., p. 1.
57. Not all quantitative research papers have a separate “Hypotheses” section or use the term “hypothesis,” but clearly stating what predictions are to be tested is a critical component of good quantitative research.
58. Williams and Monge, *Reasoning With Statistics: How To Read Quantitative Research*.
59. To more clearly specify what we mean by “no meaningful correlation” in the first two questions: we would find that the Pearson correlation between these variables would be less than 0.5 ( $r < 0.5$ ).
60. “Significant” has a specific technical meaning in statistics, explained in the Results section below. We mean “significantly” in this technical sense.
61. <<http://www.boxofficemojo.com/>>
62. <<http://www.imdb.com/pressroom/about/>>
63. The site’s ratings methodology is explained at <[http://www.imdb.com/help/show\\_](http://www.imdb.com/help/show_leaf?votestopfaq)  
[leaf?votestopfaq](http://www.imdb.com/help/show_leaf?votestopfaq)>
64. This file is available for download from several sources at <<http://www.imdb.com/interfaces>>
65. Kirk and Thompson, Confirmed.
66. The Metacritic ratings are described in the site’s Frequently Asked Questions at <[www.metacritic.com/faq](http://www.metacritic.com/faq)>
67. Using SPSS v22 for Mac.
68. Excel 2008 for Mac.
69. <<http://www.bls.gov/cpi/>>
70. <<http://www-01.ibm.com/software/analytics/spss/academic/>>
71. The three movies to score 1.6 are: *Justin Bieber’s Never Say Never*, *Justin Bieber’s Believe*, and Kirk Cameron’s *Saving Christmas*. The two to score 9.0 are *Stray Dogs* and *The Dark Knight*.
72. Metascore 1: *InAPProprate Comedy*, *10 Rules for Sleeping Around*, *Not Cool*; Metascore 100: *Best Kept Secret*, *Boyhood*.
73. This observation is supported by statistical tests; the concepts behind such tests will be explained in the next section. Specifically, we tested for differences in mean and standard deviation between Metascore and user score rescaled by multiplying by 10, using a *t* test and Levene’s test respectively. Both tests were statistically significant at the  $p < .001$  level.
74. Cohen, These guidelines apply to the simplest mathematical form of correlation, the “Pearson product-moment correlation coefficient,” which is the one we use in this article.
75. We consciously choose the asymmetric language of “proved incorrect” vs. “supported” to indicate a desire to proceed cautiously—hypotheses should be easier to falsify than to prove. This falsification-based approach is practiced by most quantitative researchers, although details differ. Entire journals are devoted to these questions, both in statistics and in philosophy of science.
76. Cartier and Liarte, Impact, p. 19.
77. Ibid., p. 17.

78. Dietz, "The Nightmare After Christmas: Are January Movies Really That Bad?" in *Metacritic*. In January, a large concern is the attention placed on Oscar hopefuls from December entering wide release; see the literature review for corroborating sources.
79. Kirk and Thompson, Confirmed.
80. Rottenberg, Imperfect Science.
81. Darling, August.
82. Although August is beginning to claim ground as a legitimate release month, as per Darling and "The August Movie."
83. We are unable to perform this comparison because of the previously discussed difference in granularity between Metacritic, which we used, and Rotten Tomatoes, which they used.
84. Burr, January, n.p.
85. Cartier and Liarte, Impact, p. 21.
86. Bernstein, Why January Is a Good Month to Bury Bad Movies," in *The Guardian*.
87. Meslow, "January: Dumping Ground for Terrible Movies Like 'Contraband,'" in *The Atlantic*.
88. Bernstein, Why January, n.p.
89. Burr, January, n.p.
90. Salisbury, Why Oscar Season, n.p.
91. Rottenberg, Imperfect Science, p. 66.
92. Cartier and Liarte, Impact.
93. Defined here using critical consensus.
94. Sochay, Predicting, p. 8.

## Works cited

- "The August Movie: A Theory of Awfulness." *Vulture*. July 31, 2008. <[http://www.vulture.com/2008/07/august\\_movie\\_awfulness.html](http://www.vulture.com/2008/07/august_movie_awfulness.html)>
- Bart, Peter. Fall Film Flurry Stirs Arty Anxieties. *Variety*. September 10–16, 2007: 4, 106.
- Basuroy, Suman, Chatterjee, Subimal, and Ravid, S. Abraham. "How Critical Are Critical Reviews? The Box Office Effects of Film Critics, Star Power, and Budgets." *Journal of Marketing* 67:4 (2003), 103–117.
- Bernstein, Jonathan. Why January Is a Good Month to Bury Bad Movies. *The Guardian* January 8, 2007. <<http://www.theguardian.com/film/2007/jan/08/awardsandprizes.features>>
- Bordwell, David, and Noël Carroll, eds. *Post-Theory: Reconstructing Film Studies*. (Madison, WI: University of Wisconsin Press, 2012).
- Burr, Ty. January Is Hollywood's Very Own Leper Colony. *New York Times Magazine* January 20, 2013.
- Cartier, Manuel, and Liarte, Sebastien. Impact of Underlying Demand and Agglomeration of Supply on Seasonality: The Case of the Hollywood Film Industry. *International Journal of Arts Management* 14:2 (2012), 17–30.
- Cohen, Jacob. *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. (Mahwah, NJ: Lawrence Erlbaum Associates, Inc. 1988). Print.
- Darling, Cary. "August: Now a Hot Month For Blockbusters." *dfw.com* August 13, 2014. <http://www.dfw.com/2014/08/13/917066/august-movies-blockbusters-guardians.html>
- Dietz, Jason. The Nightmare After Christmas: Are January Movies Really That Bad? *Metacritic*-January 5, 2010. <<http://www.metacritic.com/feature/are-january-movies-really-that-bad?page=0>>
- Einav, Liran. Not All Rivals Look Alike: Estimating an Equilibrium Model of the Release Date Timing Game. *Economic Inquiry* 48:2 (2010), 369–390.



- Emerson, Jim. Why Don't Critics, Oscars & Audience Agree? *Scanners: With Jim Emerson* Roger Ebert.com, February 10, 2012. <<http://www.rogerebert.com/scanners/why-don't-critics-oscars-and-audience-agree>>
- Feinberg, Scott. "The Perils of Being an Early Oscar Frontrunner." *Hollywood Reporter* December 20, 2013, 46.
- IMDb Votes/Ratings Top Frequently Asked Questions. *Internet Movie Database* <[http://www.imdb.com/help/show\\_leaf?votestopfaq](http://www.imdb.com/help/show_leaf?votestopfaq)>
- Kirk, Chris, and Thompson, Kim. Confirmed: February Movies Suck. *Slate* February 6, 2014. <[http://www.slate.com/articles/arts/brow\\_beat/2014/02/february\\_movies\\_are\\_bad\\_here\\_s\\_statistical\\_proof\\_of\\_it.html](http://www.slate.com/articles/arts/brow_beat/2014/02/february_movies_are_bad_here_s_statistical_proof_of_it.html)>
- Kuchera, Ben. How Does a "Terrible" Movie Make \$300 Million in Three Days? *Polygon*. June 30, 2014. <<http://www.polygon.com/2014/6/30/5857506/Transformers-bay-300-million>>
- LaMarche, Una. A Cinematic Dumpster Dive. *Vegas Seven* January 3, 2013. <http://vegasseven.com/2013/01/03/cinematic-dumpster-dive/>
- Legel, Laremie. Why is January a Cinematic Dumping Ground? *Film.com*. January 9, 2013. <<http://www.film.com/movies/january-box-office-slump>>
- Levy, Emanuel. *Oscar Fever: The History and Politics of the Academy Awards* (New York: Continuum, 2002).
- Litman, Barry R., and Kohl, Linda S.. Predicting Financial Success of Motion Pictures: The '80s Experience. *Journal of Media Economics* 2:2 (1989), 35–50.
- Maslin, Janet. Is January the Cruellest Month? *New York Times* February 5, 1989.
- McClintock, Pamela. Summer's Brutal Animation War. *Hollywood Reporter* June 21, 2013, 82–84.
- Meslow, Scott. "January: Dumping Ground for Terrible Movies Like 'Contraband.'" *The Atlantic* January 13, 2012.
- O'Neal, Sean. February Statistically Proven to Have the Shittiest Movies. A.V. Club website. February 6, 2014. <<http://www.avclub.com/article/february-statistically-proven-to-have-the-shitties-107578>>
- Plucker, Jonathan, Kaufman, James C., Temple, Jason S., and Qian, Meihua. Do Experts and Novices Evaluate Movies the Same Way? *Psychology and Marketing* 26:5 (2009), 470–478.
- Raymond, Adam K. The Cruellest Month: Just How Bad of a Movie Month Is January? *Vulture* January 7, 2013. <<http://www.vulture.com/2013/01/january-worst-movie-release-month.html>>
- Reinstein, David A., and Snyder, Christopher M. The Influence of Expert Reviews on Consumer Demand for Experience Goods: A Case Study of Movie Critics. *Journal of Industrial Economics* 53:1 (2005), 27–51.
- Rottenberg, Josh. The Imperfect Science of Release Dates. *New York Times Magazine* November 9, 2003, 64–66.
- Salisbury, Brian. Why Oscar Season Is Hollywood's Bad Movie Dumping Ground. *Hollywood.com* February 23, 2013. <<http://www.hollywood.com/movies/bad-january-february-movies-oscar-season-57156667/>>
- Singer, Matt. Why I Love Going to the Movies In January. *Criticwire*. Indiewire.com, January 23, 2013. <<http://blogs.indiewire.com/criticwire/january-movies>>
- Sochay, Scott. Predicting the Performance of Motion Pictures. *Journal of Media Economics* 7:4 (1994), 1–20.
- Williams, Frederick and Monge, Peter R. *Reasoning With Statistics: How To Read Quantitative Research* 2nd ed. (Fort Worth, TX: Wadsworth Publishing, 2000).