**Abstract**

*Rotten Tomatoes* has grown in popularity especially during this age of technology. This study examines the impact of the tomato rating heuristic on opening weekend box office revenue between the years of 2009-2018. Using a cross-sectional dataset with the regression discontinuity design method revealed that overall Tomatometer ratings even with the inclusion of controls did not have a significant impact on opening weekend box office revenue. The results open the door to future research possibilities in the subject of heuristics and their impact on the movie industry in the form of Tomatometer ratings.

**The Value of the Heuristic: *Rotten Tomato* Category**

Do fully-informed consumers respond to heuristics? In today’s world, most people base their decisions off of a review, whether it’s deciding which movie to see or which restaurant to eat at. People typically use heuristics, which are mental shortcuts or rules of thumb when making decisions. Time is costly; therefore, it is highly unlikely that moviegoers will look up data and statistics when deciding which movie they want to go see. Rather, they will rely on the icons on the *Rotten Tomatoes* online movie rating website, for example, to determine which movies are good or bad. *Rotten Tomatoes* is one of the most popular film websites that includes information about movies along with critic and user reviews. The criteria for the famous Tomatometer scores, which are based on the opinions of many film critics, are as follows: Rotten means that less than 60% of the critics reviews are positive and is represented by a green splat. If at least 60% of the critics’ reviews are positive, then a red tomato is displayed to indicate fresh status. This study intends to focus on whether the heuristic tomato category has value, even when consumers are fully informed when it comes to opening weekend box office revenues. In theory they shouldn’t.

The majority of the literature focuses more around the topic of movie going by using regression discontinuity design to analyze how a movie review impacts box office revenues. For instance, the article titled “Analyzing Dynamic Review Manipulation and Its Impact on Movie Box Office Revenue” by Haoxiang Ma, Jong Kim, and Eunkyung Lee uses a 3SLS model to examine the effect of review fraud on online word of mouth, which in turn impacts box office revenues in the movie industry. Related studies use the same concept of examining the impact of reviews, but in the restaurant industry rather than movie industry. The article “Learning From the Cloud: Regression Discontinuity Estimates of the Effects of an Online Review Database” by Michael Anderson and Jeremy Magruder from the *Economic Journal,* utilizes a RD design to analyze how Yelp.com ratings impact restaurant reservation availability. A third article “Do Consumer and Expert Reviews Affect the Length of Time a Film Is Kept on Screens in the USA” by Thaís L. D. Souza, Marislei Nishijima, and Ana C. P. Fava evaluate the impact of how consumer and expert reviews effect a film’s running time in movie theaters through the use of survival regression analysis. A fourth article “The Value of a Tomato: Examining the Impact of Review Aggregation Sites on Box Office Revenues in The United States” by Sofia Licata from Harvard University is similar to this study by examining if there is a causal relationship between *Rotten Tomato* movie reviews and box office revenues in the United States by utilizing an Regression Discontinuity Design model (RDD). What sets this study apart from the rest is that it examines the value of Rotten Tomato heuristics on opening weekend box office revenue, rather than total box office revenue with a large number of observations.

The data used in this study is from *Kaggle.com, The Numbers, and Box Office Mojo*. The rotten\_tomatoes\_movies dataset is from *Kaggle.com*, which is a website that offers interesting free datasets for research purposes. The dataset includes attributes such as movie title, rating, genre, year of release, rotten tomatoes score, and Tomatometer status. This will be a cross-sectional dataset since the data for different movies are collected from the years 2009-2018. Data in the dataset containing information of opening box office revenues and movie franchises is from *The Numbers* and *Box Office Mojo*. *Box Office Mojo* is a website that tracks box office revenue. Similarly, *The Numbers* website is a movie industry data website that conducts research services and forecasts incomes of film projects. Not every movie was the same in each of the datasets and they also were not lining up. Therefore, Stata was used to line up each of the movies in the data and delete movies that were not in both of the datasets. After the two datasets were cleaned they were merged together to create a single dataset. Using Stata to line up the movies saved a great deal of time and was more efficient than trying to accomplish this task by hand. Cleaning the data took time, but in the end it was rewarding to have a clean dataset. Overall, the dataset includes over 3,300 observations as revealed by the descriptive statistics between the years 2009-2018. The maximum and minimum value that stood out the most in the descriptive statistics pertained to opening weekend box office revenue. The movie that received the largest opening weekend box office revenue of approximately $260 million between 2009-2018 was *Avengers: Infinity War*, while the film that earned the lowest opening weekend revenue of a meager $72 thousand box office revenue was, the Tom Cruise film *Oblivion*.

An RDD model is utilized to see how a movie rating a little below and above the cutoff affects box office revenues. A single cutoff of 60% will be used to create the two groups of rotten and fresh as mentioned previously. In this case the following two assumptions would have to hold: no perfect manipulation of the running variable and nothing else changes at the 60% movie rating percentage in order for the results to be interpreted as causal effects. The first assumption will hold as long as *Rotten Tomatoes* does not alter critic reviews so that they fall right above the RDD threshold. The second assumption, the continuity assumption, will hold since nothing else changes in the relationship between movie rating percentages and ticket sales at the threshold determine whether a film is rotten or fresh. In other words, the relationship between critics’ movie ratings and ticket sales do not change discontinuously at the 60% cutoff. The only thing that changes is when the critic movie rating determines whether or not a movie falls into the category of fresh or rotten.

The variables utilized in the study are defined as follows. The dependent variable in each of the regression models used in this study is the log of opening box office revenues, which is measured in thousands of dollars. A log-linear regression model is used since the distribution of opening weekend revenue is highly skewed. The indicator variable, Tomatometer rating, represents the percentage of positive professional critic reviews given to a film and is based on the opinions of hundreds of film and television critics. Tomatometer rating serves as the running variable in the study and determines if films receive or do not receive the title fresh. Rotten Tomatoes and its famous Tomatometer rating are revered as being the world’s most trusted recommendation resources for quality entertainment. Several controls are included. The dummy variable Tomatometer Status is set equal to 0 if positive critic reviews are less than 60% and is given a green splat. A value of 1 is assigned if positive critic reviews are at least 60% and is assigned a red tomato.

The chosen controls in the study are as follows. Year is controlled so that the values are in real terms and includes the time period between 2009 and 2018. The dummy variable franchise is assigned a value of 1 if the movie is from a franchise and 0 if not. Genre includes action and adventure, animation, art house, classics, comedy, documentary, drama, horror, kids and family, musical and performing arts, mystery and suspense, romance, science fiction and fantasy, and western. The four primary ratings NR, PG, PG-13, and R for each movie are represented by the rating variable.

Figure 1 presents the regression discontinuity plot for the question at hand. Movies that have a Tomatometer rating just above and below the threshold of 0.6 will be similar in characteristics such as production costs, opening weekend revenue, etc. The movies above the Tomatometer rating of 0.6 acquire the title of fresh, while the movies below don’t and earn the title rotten. Regression discontinuity design seeks to compare the movies just below and above the threshold and considers the difference in outcomes to be from the affect of the Tomatometer rating only. The x-axis is Tomatometer rating, which means of at least 60% of critic reviews for a movie are positive and the title fresh is granted to the film. The y-axis is the log of opening weekend box office revenue. The maximum value is 19.3, which is off of the chart. The graph reveals that log-opening revenues increase with Tomatometer rating, which is anticipated. The discrete jump due to the title fresh is of course expected; however, as mentioned before this is not expected to be a significant jump at the threshold, since heuristics in theory do not impact opening weekend revenue. Sure enough, the resulting graph supports this assumption. The size of this jump is the effect of receiving the title fresh. The relationship between Tomatometer rating and the outcome is approximately nonlinear.

Figure 2 presents the McCrary Density Plot for the Tomatometer rating running variable. The first aspect of the graph that stands out is the bunching of observations and discontinuity in the density of the running variable, which is an indication that there is manipulation of the running variable. In order to test if there is manipulation of the running or variable or not, a RD density test is run. The RD density test fails to reject the null hypothesis of no density change at the cutoff with a p-value of 0.1651. This serves as an indication that there is no manipulation of the running variable.

A linear regression model is estimated with the full sample, bandwidth of 0.05, 0.1, 0.2, and 0.25, without controls. In each of the five models, the dependent variable is the log of opening weekend box office revenue. TomaRate represents the generated running variable, Tomatometer rating, centered at the cutoff 0.6. Centering implies values of 0 are where Tomatometer rating equals 0.6, negative values indicate the Tomatometer rating is less than 0.6 and positive values are the Tomatometer ratings above 0.6. The treatment variable Tomatometer status is represented by tomatometer\_status. The interaction of the running variable Tomatometer rating and treatment variable Tomatometer status represented by TomaRate x TomaStat allows the running variable to vary on either side of the discontinuity.

The first Table reveals the coefficients estimates of each linear regression discontinuity model. β1 represents the effect of Tomatometer status on log opening weekend box office revenues. Based on the coefficients estimates narrowly earning the title fresh results in a 40%-1.5% increase in opening weekend box office revenue. Β2 indicates the marginal effect of a Tomatometer rating that earned film rotten status and its impact on opening weekend box office revenue. As can be seen in the table, all of the estimated coefficients for Β2 are negative for each of the varying bandwidths. Therefore, a film that is given the title of rotten contributes between a 0.8%-28% decrease in opening weekend box office revenue. Β3 serves as the marginal effect of an additional increase in Tomatometer rating and Tomatometer status on opening box office revenue for films with fresh status relative to those with rotten status. None of the coefficient estimates turned out to be statistically significant at the 5% level as revealed by the p-values. Overall, there is not a strong relationship between the running variable and outcome of interest besides the discontinuity around the cutoff.

The four coefficients estimate models from Table 2 offer a different approach from the linear regression discontinuity model used in table 1 by providing second-order polynomial equations of the running variable, Tomatometer rating, on either side of the cutoff. TomaRate2 represents the quadratic term for the running variable, Tomatometer rating. TomaRate2 x TomaStat serves as the interaction between the quadratic term and Tomatometer status. As in the linear regression discontinuity model, the bandwidths are once again 0.05, 0.1, 0.2, and 0.25. For Tomatometer status the magnitude of the full sample changed from 40% to 80%. The coefficient that changed the most for Tomatometer status was the one assigned the bandwidth 0.05. Before the coefficient estimate was 1.5%, but then it dropped to -0.3%, which is a large change. The coefficient estimates for Tomatometer status for the bandwidths 0.05, 0.1, 0.2, and 0.25 did not change too much from table 1. Once again, none of the estimates turned out to be statistically significant at the 5% level, as revealed by the p-values.

Table 3 introduces the inclusion of control variables, such as the genre categories, MPAA ratings, and movie franchises. The most popular movie genre is comedy; therefore one would predict that comedy should exert a positive impact on the log of opening weekend box office revenue. The action genre was added as well for the sheer interest of the impact it will exert on the outcome of interest. Several action movies score above 60%, which is an indication that there may be manipulation of the running variable present. The dummy variable, comedy is equal to one if a movie falls into this category, and 0 if not. The same goes for action as well. Another factor to take into consideration is the fact that MPAA ratings of a film such as, PG+13 tend to have larger audiences. As a result, one would anticipate that movies rated PG-13 or over should have higher opening weekend revenue than any other rating categories. Films that are part of a franchise are assigned a value of 1, and 0 if they are not as represented by the franchise variable. Films that are part of a franchise would earn more money on opening weekend, than those that are stand-alone movies.

The coefficient estimates in Table 3 use the same linear functions as table 1, except controls are added. Interestingly, the coefficient estimates turned out to be exactly the same as those from table 1. This is what is expected since if the coefficient estimates changed drastically it is likely that perfect manipulation is present. Controls enable the estimates to be more precise. The only coefficient estimates that changed slightly were those for Tomatometer rating. This makes sense intuitively since when critics rate movies that base it on a number of factors in general such as the actors in the film, whether its part of a popular franchise or not, etc.

Table 4 displays the coefficient estimates using the local kernel estimator, RDROBUST. The first column provides the CCT optimal bandwidth and the second through fourth columns use the same bandwidths as above, 0.05, 0.1, 0.2, and 0.25. The estimates are similar to those in the other tables in size and significance.

A shortcoming of the study is the accuracy of the data. All of the critic reviews that came out after the movies were released were close to zero; therefore it served as a strong proxy for ratings that were made just prior to opening weekend box office revenue. Critic reviews that are made before a movie is released are only available for new movies on the rotten tomatoes website. Future researchers should scrape the information for new releases, which would provide more accurate coefficient estimates. A downside is for the current year 2020 due to the Covid-19 epidemic a lot of movies are being postponed until 2021 such as the new Batman film. Researchers who would like to scrape the data for this year are going to not have a lot of films to collect information from. Not only are some of the films being delayed, but there is also not an audience to watch the films as the result of movie theaters being temporarily closed. Unfortunately, some of the data such as opening weekend revenue or critic rotten tomato reviews were missing between the years 2009-2018, as a result these films were dropped from this research project.

In light of future research, additional controls should be added to strengthen the results of the study. A control that would be interesting to add is a film’s production budget. One would expect that films with higher budgets are more likely to be better movies compared to those with lower budgets. The downside of this approach is that some film companies do not reveal their production budgets in an effort to keep this type of information hidden from competitors. Aspiring researchers have the option to drop from the films that do not provide production budgets study or pay to acquire a more complete dataset. An additional control that should be looked into is the distributor of the film such as, Warner Bros, and Walt Disney Studios Motion Pictures to examine if a distributor’s popularity impacts opening weekend revenue. An additional variable that would be interesting to research is to see if a particular actor has an influence on opening weekend box office revenues. More popular actors should bring in more movie revenue.

Some potential future research project ideas would be to add more years to the study to make it more comprehensive, conduct a fuzzy regression discontinuity design, and to use audience critic reviews instead. Rotten Tomatoes was launched in August 12, 1998. Future researchers should collect data from this year and onward to analyze how *Rotten Tomato* reviews impacted opening weekend revenue over time as it gradually gained popularity. It would be interesting to run a fuzzy regression discontinuity design so that the criteria for certified fresh versus a plain tomato can be distinguished. This will allow the cutoffs of 60% and 75% to be used. Another possible future regression model that can be run is to utilize audience reviews to examine the impact of opening weekend box office review in place of critic reviews.

In summary the table results did not reveal much statistical significance, which was initially anticipated since in theory the Tomatometer rating heuristic does not have a significant impact on opening weekend box office revenue. Even though there was bunching in the regression discontinuity plot and some changes in the local estimates, the inclusion of controls and McCrary Test did not provide evidence of manipulation of the running variable. This study provides a lot of leeway for aspiring future researchers to conduct studies with new controls, variables etc. to contribute new insights to the study of movie industry revenues.

**Resulting Tables and Graphs**

**Descriptive Statistics**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Variable | Obs | Mean | Std.Dev. | Min | Max |
| Tomato status | 3366 | .573 | .495 | 0 | 1 |
| Tomato Rating | 3366 | .611 | .272 | 0 | 1 |
| opening | 3366 | 8530000 | 2.05e+07 | 72 | 2.58e+08 |
| franchise | 3366 | .094 | .291 | 0 | 1 |
| rating | 3366 | 2.519 | 1.601 | 0 | 4 |
| genre | 3366 | 4.03 | 2.465 | 0 | 13 |
| year | 3366 | 2013.387 | 2.677 | 2009 | 2018 |
| lnopening | 3366 | 12.365 | 3.453 | 4.277 | 19.367 |
|  | | | | | |

Table 1: Linear RD Estimates

----------------------------------------------------------------------------------------------------

(1) (2) (3) (4) (5)

All h = 0.05 h = 0.1 h = 0.2 h = 0.25

----------------------------------------------------------------------------------------------------

tomatometer\_status 0.396 1.501 0.940 0.702\* 0.551

(0.247) (0.957) (0.614) (0.414) (0.358)

TomaRate -0.853 -28.94 -2.944 -1.521 -0.967

(0.577) (29.90) (8.412) (2.871) (1.933)

TomaRate x TomaStat 0.702 19.79 -3.945 -0.850 0.0867

(0.917) (35.41) (10.59) (3.605) (2.450)

----------------------------------------------------------------------------------------------------

Observations 3366 262 604 1346 1770

R-squared 0.001 0.011 0.006 0.003 0.002

----------------------------------------------------------------------------------------------------

Standard errors in parentheses

\* p<0.10, \*\* p<0.05, \*\*\* p<0.01

Table 2: Quadratic RD Estimates

----------------------------------------------------------------------------------------------------

(1) (2) (3) (4) (5)

All h = 0.05 h = 0.1 h = 0.2 h = 0.25

----------------------------------------------------------------------------------------------------

tomatometer\_status 0.837\*\* -0.358 1.053 0.898 0.869

(0.374) (1.943) (0.983) (0.637) (0.558)

TomaRate -3.652\* 177.7 -6.413 -7.979 -4.748

(2.198) (173.8) (39.42) (11.73) (8.066)

TomaRate² -4.959 4154.6 -33.92 -32.43 -14.88

(3.757) (3442.7) (376.5) (57.12) (30.82)

TomaRate x TomaStat 1.106 -230.7 -4.649 6.805 0.350

(3.555) (187.3) (46.08) (14.36) (10.00)

TomaRate² x TomaStat 10.85 -3016.9 81.79 26.29 29.04

(7.631) (3859.6) (459.5) (70.29) (38.46)

----------------------------------------------------------------------------------------------------

Observations 3366 262 604 1346 1770

R-squared 0.002 0.018 0.006 0.003 0.003

----------------------------------------------------------------------------------------------------

Standard errors in parentheses

\* p<0.10, \*\* p<0.05, \*\*\* p<0.01

Table 3: RD Estimates with Controls

----------------------------------------------------------------------------------------------------

(1) (2) (3) (4) (5)

All h = 0.05 h = 0.1 h = 0.2 h = 0.25

----------------------------------------------------------------------------------------------------

tomatometer\_status 0.462\* 1.594 0.896 0.801\* 0.604\*

(0.244) (0.983) (0.611) (0.409) (0.352)

TomaRate -0.850 -29.34 -0.986 -2.006 -1.084

(0.574) (30.59) (8.405) (2.846) (1.908)

TomaRate x TomaStat 0.720 12.09 -4.997 -0.250 0.164

(0.909) (36.52) (10.58) (3.583) (2.418)

----------------------------------------------------------------------------------------------------

Observations 3366 262 604 1346 1770

R-squared 0.041 0.085 0.059 0.046 0.045

----------------------------------------------------------------------------------------------------

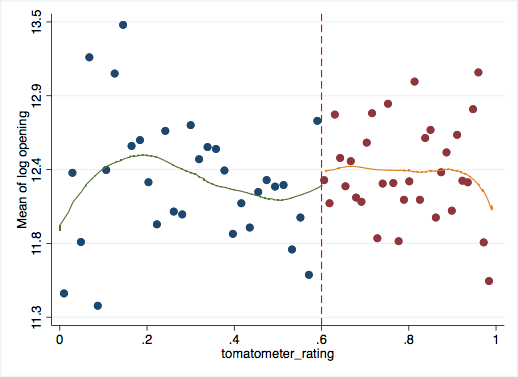
Standard errors in parentheses

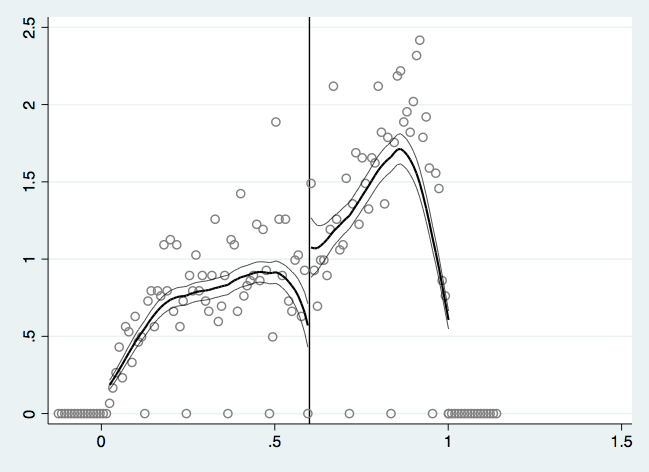
\* p<0.10, \*\* p<0.05, \*\*\* p<0.01

Table 4: Nonparametric Kernel Density RD

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) | (5) |
| VARIABLES | lnopening | lnopening | lnopening | lnopening | lnopening |
|  |  |  |  |  |  |
| RD\_Estimate | 0.841 | 1.258 | 0.978 | 0.767\* | 0.672\* |
|  | (0.514) | (1.109) | (0.687) | (0.462) | (0.402) |
|  |  |  |  |  |  |
| Observations | 3,366 | 3,366 | 3,366 | 3,366 | 3,366 |
| Bandwidth | 0.164 | 0.0500 | 0.100 | 0.200 | 0.250 |
| Obs | 654.2 | 154.1 | 351.1 | 793.2 | 1060 |

Figure 1: Log Opening Weekend Box Office Revenue by Tomatometer Rating



Figure 2: Observation Density by Tomatometer Rating – McCrary Plot

**References**

Anderson, M., & Magruder, J. (2012). Learning From the Crowd: Regression Discontinuity Estimates of the Effects of an Online Review Database. *The Economic Journal*, *122*(563), 957-989. <https://doi.org/10.1111/j.1468-0297.2012.02512.x>

Ma, H., Kim, J. M., & Lee, E. (2019). Analyzing Dynamic Review Manipulation and Its Impact on Movie Box Office Revenue. *Electronic Commerce Research and Applications*, *35*, 100840. <https://doi.org/10.1016/j.elerap.2019.100840>

Souza, T. L., Nishijima, M., & Fava, A. C. (2019). Do Consumer and Expert Reviews Affect the Length of Time A Film Is Kept on Screens in the USA?. *Journal of Cultural Economics*, *43*(1), 145-171. <https://doi.org/10.1007/s10824-018-9332-6>

Sweigart, M. (2018). *The Value of a Tomato: Examining the Impact of Review Aggregation Sites on Box Office Revenues in the United States* [Senior thesis, Princeton University]. DataSpace. <https://dataspace.princeton.edu/handle/88435/dsp01t435gg711>