The Importance of Being the First in Korean Box Office

: Causal Inference Using a Regression Discontinuity Design

JI HEE LEE

Sungkyunkwan University

Abstract

I estimate the effect on the movie performance of leading the weekend box office for the first week of release using the regression discontinuity approach. When a researcher does not observe all movie qualities that are relevant to the success of the movie, the omitted variables may bias upward estimates of the #1 effects. The regression discontinuity strategy in my approach posits that the treatment of being #1 is locally randomized for the movies with small margins of victory. The results suggest that to exploit the discontinuity in determining the highest grossing movies greatly reduces the bias from omitted variables. I find being #1 in the first week of release to account for an increase of the total audience by 15% over the movie's run in the Korean motion-picture industry. The estimates of the #1 effects nearly triple to 40% without dealing with the potential endogeneity from omitted variables.

1. Introduction

Movies with the title 'Number 1 (#1)' can come in different to consumers choosing which movie to watch. Especially, a movie that ranked #1 during the opening weekend can mean a lot to them, maybe a strong proof for its superiority. People rely on responses of the early moviegoers for the quality, because they cannot know about it prior to their consumption other than that. In addition, people can easily follow up both the updates of new movies and their rankings since the number of movie opening every day is countable. Therefore, a movie being #1 during the very first weekend will surely give a positive signal to its potential audiences. Then, how influential will it be?

The process is not that simple since it is hard to precisely quantify a quality of a movie. There are many widely used proxies for movie quality. However, omitted variable problem always exists since econometricians fail to observe and quantify the quality perfectly. Movie quality variables that are omitted are likely to have positive correlation with the #1 indicator variable, thus resulting in an overestimation of the #1 coefficient with endogeneity problem. Thus, I focus on precisely estimating the #1 effect by getting rid of omitted variable problems that were mentioned as limitations in the former studies.

My paper relates to a literature that focuses on estimating demand for movies. Many of them suggest diverse factors that influence the performance of movies (Einav (2007), Elberse (2007), Park and Jung (2009), Kim and Han (2014), Chang et al. (2009)). Moreover, a few studies have focused on awareness effects in movie demand. Moul (2007) identifies and measures the impact of word-of-mouth, discovering that 10% of the variation in consumer expectations of movies can be attributed to information transmissions. Moretti (2011) figures out 32% in sales of movies with a positive surprise are made through social learnings and peer effects. Within the framework of awareness effect, I concentrate on the awareness driven by the ranking '#1'. Particularly relevant for my research is the work by Cabral and Natividad (2016). They propose a theoretical background for understanding the #1 effect, and empirically test their predictions using the U.S. box office data. They discover that the primary channel for the #1 effect is through the greater awareness by being #1. In their empirical tests, they use two movie quality variables, *Star Power* and *Quality Rating*. They recognize their limitations on the possibility that the #1 indicator variable may be picking up heterogeneity in omitted movie quality variables. Thus, Cabral and Natividad (2016) do not claim to have verified the causal effect of being #1, but rather correlation based on their theoretical results.

My paper is differentiated from the former movie-demand-studies in that I focus on estimating the 'causal' inference of the awareness effect by being #1, rather than specifying its channels. On behalf of the empirical limitation of Cabral and Natividad (2016), I suggest applying the regression discontinuity design (RD design) to overcome the endogeneity problem between the #1 indicator and omitted movie quality variables.

The RD design has been applied in the literatures of awareness effect. Anderson and Magruder (2011) implements the RD design to estimate the effect of positive Yelp.com ratings on restaurant reservation availability. They solve the endogeneity problem of review quality and restaurant quality by using the method. However, my paper shares more context with RD designs used in literatures of political economics, such as comparing winners and losers of an election. The studies have in common that the RD design can be used for getting rid of unobserved individual heterogeneity, which has potential of omitted variable problem. In the RD design, under weak assumptions, observations that barely won or lost share similar characteristics except for the treatment effect. Therefore, using observations close to the threshold will allow for an identification of differences between a control and treatment groups, which leads to causal inference of the treatment effect. In the literatures of political economics, whether the candidate is incumbent or not is usually the treatment. Especially, my RD setting has a lot in common with Hainmueller and Kern (2008), which measures the incumbency effect in Germany's mixed electoral system using the RD design. They compute the margin of vote (or victory) in the election at $t - 1$ as an assignment variable, and vote share at t as an outcome variable. Discontinuity exists at the threshold where the winning(incumbent) or losing(non-incumbent) in the prior elections is decided. By identifying the discontinuity, they effectively measure the causal effect of being an incumbent on vote shares of next elections.

In order to overcome the limitation of former studies of movie demand and to use the methodological power the RD design possesses, I apply the RD design on estimating the #1 effect as a treatment effect. Intuitively, I expected there will be a considerable difference in performances between movies that barely got the #1 title and that barely failed, although they are similar except for the #1 title. With the raw data, I identified the discontinuity graphically at the threshold where the victory of #1 and non-#1 gets decided. The main idea behind using the RD design in a quasi-experimental setting is that the treatment of being #1 gets effectively randomized with movies having close margins of victory under relatively weak assumptions. By comparing the movies that barely become #1 or not during their opening weekend, I can deal with omitted movie quality variable problems, because those movies will largely share common properties except for the treatment. As a result, I effectively address the estimation problem of Cabral and Natividad (2016), and allow for a causal inference.

Using 2010-2016 movie-level data from Korean Film Council, I estimated the #1 effect to be 15% of the mean value of #1 movies' total audiences under the RD design setting. However, the #1 effect nearly triples to 40%, if I ignore the omitted variable problem. Therefore, by applying the RD design, I control the positive bias of the #1 coefficient and conclude that being #1 during the opening weekend has a causal effect of attracting 15% of the total audiences.

To the best of my knowledge, my research is the first to use the RD design for estimating the awareness effect in movie demands. Therefore, the main contribution of my paper is that I suggest a different perspective of handling endogeneity problem in the movie demand literatures. My model based on the RD design effectively addressed the problem arising from omitted movie quality variables and enabled a precise estimate for the causal effect of a movie being #1.

The paper is structured as follows. In Section 2, I describe my data. In Section 3, I suggest my empirical framework based on applying the RD design, and how I handled the endogeneity problem. The main results are presented in Section 4. Section 5 concludes the paper.

2. Data

My data is comprised of movies that were released in Korean theaters for 7 years, between 2010 to 2016. I used data from Korean Film Council (KOFIC) and Naver movies. KOFIC provides most of the aggregate movie-level data, including periodical number of audiences and other movie characteristics. It also provides person-level data with filmographies of actors and directors. I used ratings of a movie from 'Netizen Ratings' provided by Naver movies.

My analysis is conducted at the movie level. I focused on top 100 movies each year, with 694 movies consisting my sample. 6 movies were excluded due to missing data. There are 342 weeks in my sample. In each weekend, there exists a movie with the #1 title, but my focus is on movies that are #1 on its opening weekend. Thus, dummy variable *Number 1* is defined as equal to one when a film is the absolute winner during its first weekend. 186 movies achieved the status. From now on, I will indicate being #1 as the status of a movie that has achieved the #1 ranking during its very first opening weekend.

There are three different movie performance variables; *Total audience, Daily average audience* and *2nd week* *audience*. I gradually restrict the duration of the #1 effect by using the three outcome variables with different periods. They are all expressed in number of audiences. I set *Total audience* as the number of total audiences minus the first weekend audiences. *Total audience* is used for defining the final outcome of a movie. After all, this performance variable is the one that movie people are most likely to be interested in. Next, I restrict the #1 effect duration to a month by using *Daily average audience.* It is the number of total audiences for 30 days, divided by 30. I used the daily average term since there exists movies with less than 30 days opening, in which I used the average value of 28 - 29 days. The variable shows the performance of movies in a similar running periods, and allow to estimate the #1 effect on a same given term. Lastly, I further restrict the duration to a week, with *2nd week audience*. It is the audience number in the second weekend. This is the variable with the shortest term, and I expected that the effect of the title #1 would have the most direct causal effect on the number of audience in the right next weekend.

Control Variables To effectively segregate the #1 effect from other factors that influence movie performances, I use the following control variables: 3 different kinds of movie quality variables, opening screen, genre, rating, distributor, opening week, holiday effect, seasonal effect of summer and winter and year effect. Except for the movie quality variables and opening screen, the control variables are in the form of dummy variables.

For movie quality, I used the three quality variables that are widely used in the literatures. I defined *Star Power* as the sum of audience of all movies that three-main actors have had appearance for, over the 3 years prior to the release of a movie, divided by 3. In Cabral and Natividad (2016), *Star Power* was operationalized as the sum of box office revenue of films of each team member over the three calendar years prior to the release of a movie, divided by the number of team members. I defined *Director Power* as the average of audience of all movies that the director has directed before the movie's release. This is based on the belief that unlike a rather temporary popularity of stars, director's power can be everlasting. For the last quality variable, I used *Quality Rating* of Netizen ratings from Naver Movies. For *Open Screen*, I used the number of screen on the opening day of a movie. Table 1 provides basic summary statistics of the main variables used in the regressions.

Table 1
Summary Statistics (Movie Level)

Variable	Obs	Mean	Std. Dev.	Min	Max
Total audience	694	1,343,779	1,872,740	85,827	14,300,000
Daily average audience	694	57,550	67,591	267	556,074
2nd week audience	694	306,678	381,761	5,786	2,823,600
Number 1	694	0.27	0.44	0	1
Star Power	694	461	572	0	3,640
Director Power	694	88	112	0	822
Quality Rating	694	7.79	1.01	3.00	10.00
Open Screen	694	483	249	15	1,864

Notes: *Star Power* and *Director Power* in 10,000 people.

For specification of other control variables, I categorized 11 *genres*. Drama, Action/Thriller/Crime, Animation, Melo/Romance, Comedy, Documentary, Horror/Mystery, Performance/Musical, SF/Fantasy, Adventure/Family, and Others. For *ratings*, there exists 4 kinds of age restrictions in Korean movies, Universal, age of 12, 15 and 19.

For *distributor*, I divided it in 3 groups. One is major distributors in Korea, where CJ E&M, Showbox, Lotte and NEW are included in this group. Another one is distributors that import movies from Hollywood, where Warner Brothers, 20th Century Fox, Walt Disney Company, Universal Pictures and Sony Pictures are included. Distributors not included in the above two groups are included the last group. *Opening week* dummies capture weekly differences among movies. Furthermore, it can be used for considering different levels of competition every week. Total of 342 weeks are included in the sample. For *Seasonality,* 11 kinds of *holidays* are regarded: New Year, Lunar New Year, 3.1, Buddha's Birthday, Children's Day, Memorial Day, National Liberation Day, Chu-Seok, Foundation Day, Hangul Proclamation Day and Christmas. I also made dummies for *summer season*, from July to August, and for *winter season*, December to February. For each year 2010 to year 2016, *year dummies* were given.

In the next section, I suggest my empirical framework based on the above dataset.

3. Empirical Framework

Regression Discontinuity Design (RD design) I rely on the RD design to obtain estimates of the causal effects of being #1 during the opening weekend. Let Y be the overall performance of a movie, and let $\mathrm{Y_{1i}}$ denote the potential outcome for movie i exposed to the treatment of being #1 during the opening weekend. Then $\mathrm{Y_{0i}}$ denote outcome of movie i not exposed to the treatment. Causal inference is difficult to make since we only observe $\mathrm{Y} = \mathrm{DY_1} + (1 - \mathrm{D})\mathrm{Y_0}$. However, in the RD design setting, under certain assumptions, the average treatment effect $\mathrm{ATE} = \mathrm{E}[\mathrm{Y_1} - \mathrm{Y_0}]$ can be estimated.

$$Observed\ \#\ of\ audience_i\ = Latent\ \#\ of\ audience_i\ + \eta_i \qquad (1)$$

Observed # of audience of movie i can be decomposed into *Latent # of audience* and η as in equation (1). Let *Latent # of audience* denote to a systematic or predictable component that is a function of the movie's attributes or actions. Then η is an exogenous, random chance component with mean zero and a continuous density. The existence of η allows local randomization at the threshold (Hainmueller and Kern (2008)). The main idea of RD designs is that observations near the cutoff of victory share similar level of *Latent # of audience*, so comparing them allows econometricians to be freed from omitted variable problem. In other words, as the movies get closer to the threshold, the random part η of the *Observed # of audience* significantly decides on whether the movie receives the treatment or not. Then, the only difference between the observations just above and below the cutoff is the presence of the treatment effect which gets randomly distributed. Therefore, inference through the RD design allows us to navigate causal effects like comparing control and treatment groups in an experiment. This is why the RD design is so called as a quasi-natural experiment, since the treatment effect gets effectively randomized on observations that are near the cutoff point.

The basic intuition behind using RD designs on estimating the #1 effect is that movies which barely got #1 or barely failed to become #1 will be similar, even in aspects that econometricians fail to observe. For example, there will exist a qualitative common-taste component of successful movies that strikes tastes of public. This component is difficult for econometricians to observe and to quantitively measure. However, if they exclude the variable, the effect of it will be absorbed by the error term, while movies with high common-taste component are more likely to perform well, and this applies especially to the #1 movies. This will result in correlation between the error-term and the #1 indicator, which we denote as the 'endogeneity problem'. Thus, the coefficient of the #1 indicator will be overestimated. The problem can be solved by using samples near the cutoff of being #1 and non-#1, where they share similar levels of common-taste component. Then identifying the discontinuity in performances of the two groups will enable me to estimate the precise causal effect of being #1, since the omitted variable is no longer a problem. Based on this intuition, I applied the RD design on estimating the #1 effect on movie demand.

RD design Variables There are 3 main variables to concern when applying the RD design; the outcome variable, the assignment variable and the indicator variable for treatment effect. First, for the outcome variable in the RD setting, I used the 3 performance variables, *Total audience, Daily average audience* and *2nd week audience.* Specifications of these variables were made in the Section 2. For the second main variable, which is the assignment variable, I defined it as the margin of victory (MV) during the opening weekend. For the winners (#1), MV is defined as the margin of its opening weekend audience with the #2 movie in the same weekend. In this case, the #2 movie does not have to be in its opening weekend, though in some cases both #1 and #2 movies share the same opening weekend. The MV for second-ranked movies (#2) is defined as the margin of the first weekend audience with the same weekend's #1. Finally, MV for third or higher ranked movies (#3 or bigger) is the margin with the #1 in the same weekend. Also, in both cases (MV of #2 and #3 or bigger), the #1 movie does not have to be in its opening weekend. Under this assignment variable setting, the threshold is where the value of MV is zero. Therefore, if the MV of a movie is bigger than zero, it receives the treatment effect of being #1. On the other hand, if the MV is smaller than zero, it does not receive the treatment. Thus, *Number 1* dummy variable serves as the #1 indicator.

The reason that I included not only second-ranked movies, but also other rankings is that the rankings might be not that important while margin with the winner matters. For instance, #3 movies will still be similar with #1 ranked movies if the margin between the two is small. There will be no meaning of distinguishing #2 and #3 in this case. This definition of MV is used widely in applying the RD Design on estimating incumbency effect in elections. Especially, the assignment variable used in Hainmueller and Kern (2008) is similar with my MV in the setting.

Validity of RD design Testing the validity of applying the RD design is one of the important steps in the implementation process. It is testing the possibility of manipulation of the assignment variable (MV). If movies have perfect control over the opening weekend ranking, then the randomization process of the treatment effect near the threshold (MV=0) will fail. Intuitively, even though movies can exert some level of efforts to be the #1, there are more uncontrollable factors on the decision of rankings. η_i in equation (1) denotes this, and the presence of it is the key factor for randomization process near the threshold to work.

There are two ways for proving that there is no manipulation problem with the setting. One is the examination of the density of an assignment variable itself (McCrary (2008)). The density of the assignment variable should be continuous at the threshold. However, by construction in my model, the density of MV is symmetric on MV=0, so it is always continuous at MV=0. The other approach for testing the validity is to examine whether there is discontinuity in any pre-determined (baseline) variables at the RD threshold. (Lee (2008)). In other words, the existence of discontinuity should only be due to treatment effects, and not result from other factors. Figure 1-4 shows the graph of baseline covariates on the assignment variable, MV. In my setting, *Star Power, Director Power, Quality Rating* and *Open screen* are characteristics of a movie that are determined before the first weekend ranking. As in the figures, there are no notable discontinuity at MV=0 for all the 4 variables. Thus, I concluded that there is no perfect manipulation problem in my RD design setting that can impediment the randomization process of the #1 effect.

Graphical Presentation In the RD design, graphical presentation can be used for showing 56the transparency of its method. I followed the widely-used graphical presentation method, especially from Lee and Lemieux (2010). They suggest dividing the assignment variable into many bins, and the average value of the outcome variable can be computed for each bin and graphed against the mid-points of the bins. There are two issues to regard when presenting a graph in the RD design. One is choosing the bandwidth of the assignment variable, and the other is choosing the number of bin. A thorough decisions on the bandwidth and the number of bin should be made for a more precise graphical presentation, and are planned to be made in the near future. For now, I have set the bandwidth of 2,000,000, with MV ranging from -1,000,000 to 1,000,000. For the number of bin, I suggest one with 200 bins and the other with 100 bins.

Figure 5 and Figure 6 shows the graphical presentation of my data. Among the three performance variables, I only report the graph using *total audience*, while the discontinuity is clearly shown in the other two variables as well. In Figure 5, I divided the MV into 200 bins, while in Figure 6 the number of bin is 100. The figures also show the estimated 2nd order polynomial in margin of victory, allowing for a discontinuity at the 0 margin. There exists a clear discontinuity at MV=0 in both cases. Next, I suggest my model for identifying the discontinuity at the threshold and estimating the #1 effect.

Polynomial Regression For estimation, I follow the pooled polynomial regression of Lee and Lemieux (2010). I allow for a highly flexible functional form of MV on the performance. By estimating two separate regressions on each side of the cutoff point, I can identify the scale of discontinuity at the threshold as a treatment effect. Thus, τ will be the parameter that I am interested in, which measures the difference between right side (MV>0) intercept (α_r) and left side (MV<0) intercept (α_l). In the estimation process, the decision on the order of polynomial should be considered. In the next section where I report the estimated value for τ, I will show the results attained from various orders of polynomial.

$$
\begin{aligned}
&Performance_i = \alpha_l + \tau Number1_i \\
&+\beta_{l1}(MV_i - Number1_i MV_i) + \cdots + \beta_{ln}(MV_i^n - Number1_i MV_i^n) \\
&+\beta_{r1}\mathrm{Number1_i MV_i} + \cdots + \beta_{rn}\mathrm{Number1_i MV_i^{\mathit{n}}} + \mu_i \qquad (2)
\end{aligned}
$$

n: order of polynomial, α_r: $Right\ side\ intercept$, α_l: $Left\ side\ intercept$

$$\tau = \alpha_r - \alpha_l$$

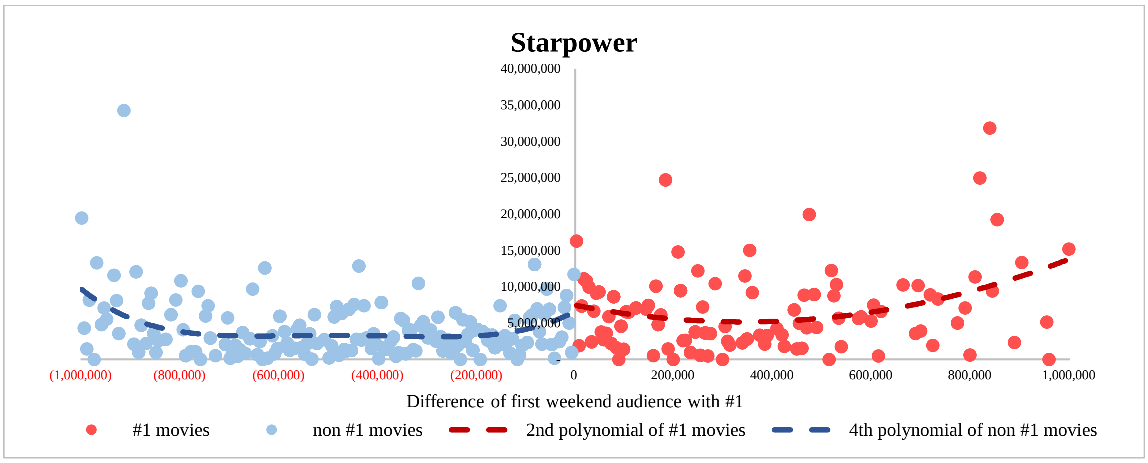


Figure 1
Validity of the RD design (Baseline covariate: Star Power)

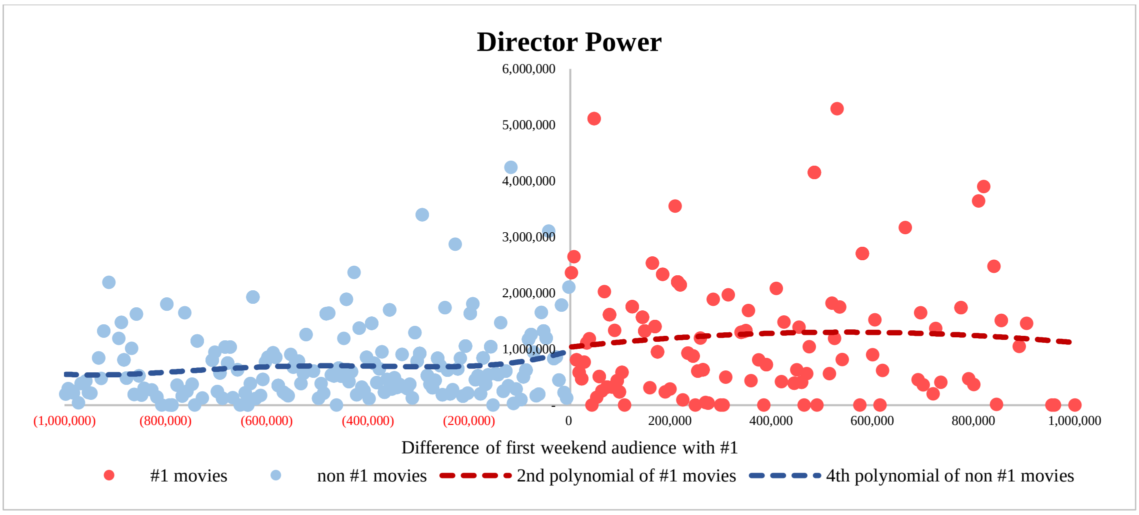


Figure 2
Validity of the RD design (Baseline covariate: Director Power)

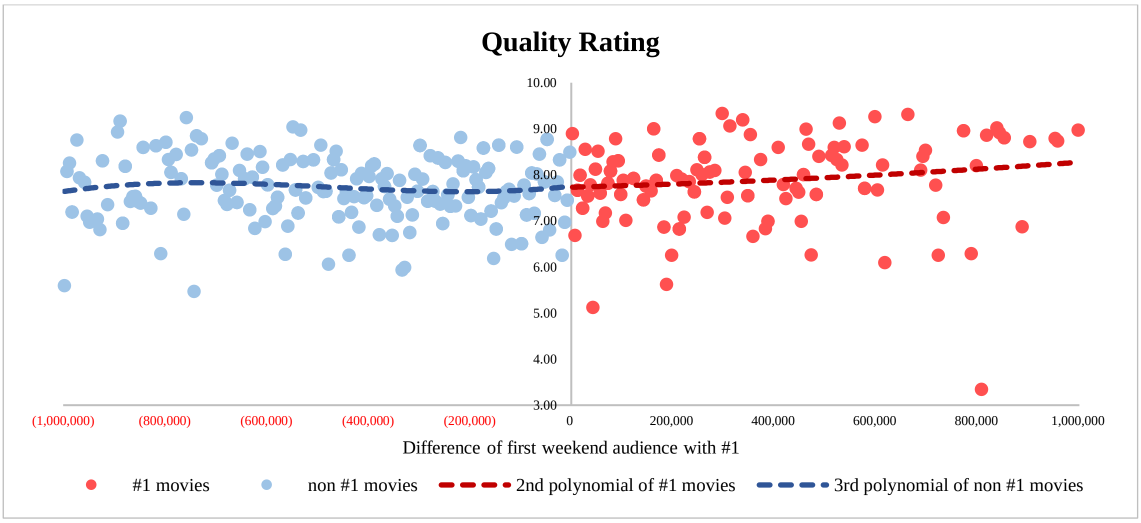


Figure 3
Validity of the RD design (Baseline covariate: Quality Rating)

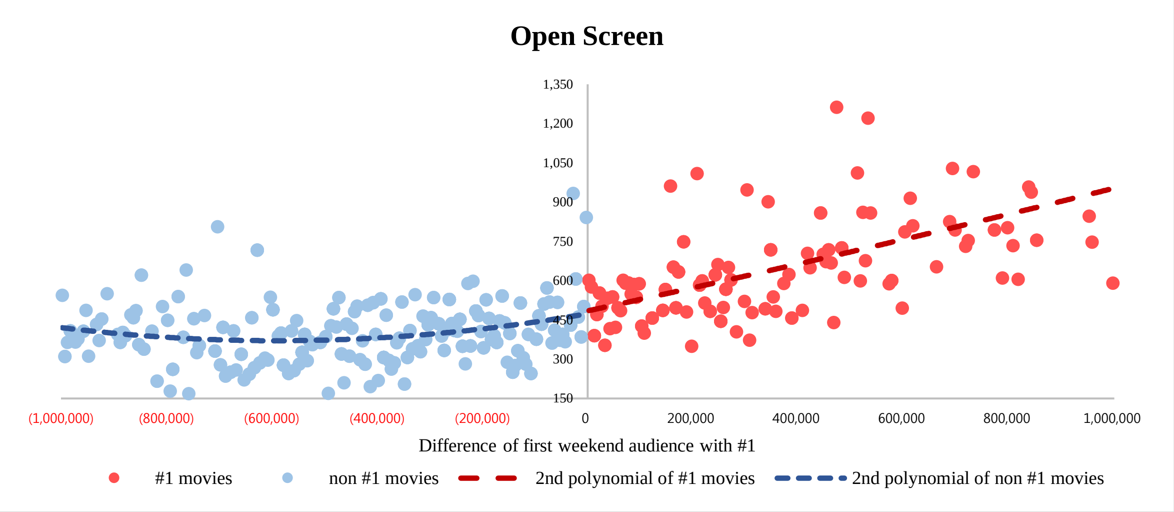
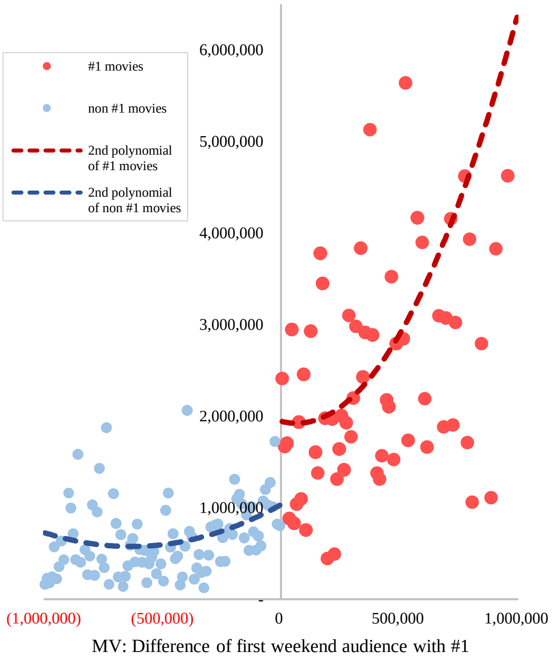


Figure 4
Validity of the RD design (Baseline covariate: Open Screen)



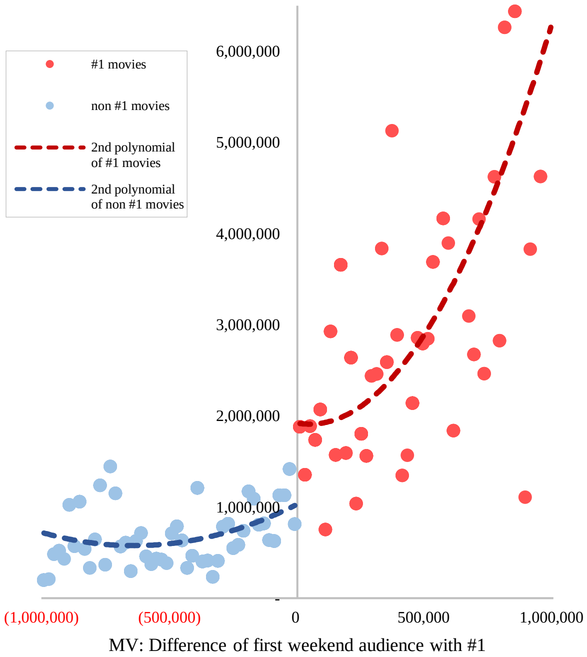


Figure 5
Graphical Presentation (Total audience),
200 bins of MV

Figure 6
Graphical Presentation (Total audience),
100 bins of MV

4. Results

For reporting the results, first I will show the results from benchmark regression where I ignore the omitted movie quality variables. Next, I will suggest the results from polynomial regressions based on the RD Design. Lastly, I will compare the two results and show the positive bias existing in the former model, while it is effectively controlled in the latter model. The interpretation of the coefficient will in made in the comparison part, 4.3.

4.1 Benchmark Regression Results

In Table 2, model (1) - (2) uses total audience, (3) - (4) daily audience average, and (5) - (6) uses 2nd weekend audience as a dependent variable. Overall, the key independent variable, *Number 1* is statistically significant (*p* value lower than 1%) and economically significant in all models. Control variables are also significant and allow for a sound interpretation. Being #1 during the opening weekend has an effect of drawing 1,178,952 audiences for *Total audience*. Controlling for weekly differences, the coefficient slightly increases to 1,298,703. The same interpretation can be made to other models with different dependent variables.

Table 2
Benchmark Regression Results

Model	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Dependent Variable	**Total aud**	**Total aud**	**Daily aud avg**	**Daily aud avg**	**2nd week aud**	**2nd week aud**
Number 1	1178952.2***	1298703.2***	42371.9***	43804.1***	239864.6***	255852.5***
	(153858.4)	(203889.3)	(5077.5)	(6058.9)	(30643.0)	(33321.5)
Star Power	640.0***	491.2**	21.49***	15.40**	119.1***	90.21**
	(164.0)	(217.7)	(5.172)	(6.612)	(29.26)	(36.44)
Director Power	2970.3***	3066.7***	77.00***	72.09***	391.9***	343.2**
	(781.6)	(885.9)	(23.51)	(26.87)	(133.9)	(145.7)
Quality Rating	488179.1***	549189.8***	14314.9***	16402.6***	90958.6***	104689.9***
	(55026.7)	(84063.6)	(1593.7)	(2455.9)	(9357.5)	(13911.2)
Open Screen	2467.9***	3109.6***	128.9***	157.2***	709.9***	805.2***
	(392.1)	(598.5)	(14.59)	(19.65)	(95.78)	(111.7)
Movie Characteristic	O	O	O	O	O	O
Week dummy	X	O	X	O	X	O
Sample size	694	694	694	694	694	694
R-square	0.628	0.806	0.716	0.862	0.675	0.853

Notes: ***, **, * significant at the 1%, 5% and 10% level. Robust standard errors in parentheses.

Next, I would like to report the results from the RD design estimates.

4.2 Polynomial Regression based on RD Design Results

Table 3
Polynomial Regression Results – Total audience

Total audience	(1)	(2)	(3)	(4)	(5)
Number 1	890875.5*** (160875.5)	736817.3*** (209178.2)	773779.9** (324832.7)	640687.9* (383752.3)	934526.9** (438825.4)
order of polynomial	1	2	4	5	6
N	694	694	694	694	694
R-sq	0.505	0.506	0.515	0.515	0.519

Notes: ***, **, * significant at the 1%, 5% and 10% level. Robust standard errors in parentheses.

Table 4
Polynomial Regression Results – Daily Average audience

Daily Average audience	(1)	(2)	(3)	(4)
Number 1	29458.8*** (4848.0)	24023.5*** (6297.7)	18256.5* (9765.4)	22758.8* (13171.6)
order of polynomial	1	2	4	6
N	694	694	694	694
R-sq	0.655	0.656	0.663	0.667

Notes: ***, **, * significant at the 1%, 5% and 10% level. Robust standard errors in parentheses.

Table 5
Polynomial Regression Results – 2nd week audience

2nd week audience	(1)	(2)	(3)	(4)
Number 1	169443.8*** (28544.2)	133525.0*** (37053.7)	93610.7** (47071.3)	110242.2* (57665.5)
order of polynomial	1	2	3	4
N	694	694	694	694
R-sq	0.625	0.627	0.630	0.632

Notes: ***, **, * significant at the 1%, 5% and 10% level. Robust standard errors in parentheses.

Table 3 - 5 shows the result of a polynomial regressions based on the RD design. Without any control variables, each table shows results from different dependent variables respectively. The models (1) – (4) in each table have different orders of polynomial. I reported only the models that are statistically significant.

From Table 3, the #1 effect has a range of 736,817 to 934,527 for total audiences, depending on the order of a polynomial. From Table 4, the #1 effect varies from 18,256 to 29,459 for daily average audiences. Lastly, Table 5 shows that the #1 effect has a range of 93,610 to 170,043 number of 2nd week audiences.

4.3 Comparison of benchmark regression and RD design regression

Lastly, to verify the positive bias, I would like to sum up all the results derived from the above regressions. Before getting started, I will show the #1 and non-#1 movies' mean value of performance for interpretation of the coefficients. Table 6 is a summary statistic divided by #1 movies and non-#1 movies.

Table 6
Summary statistic of #1 and non-#1

	Variable	Obs	Mean	Std. Dev.	Min	Max
Number 1	**Total audience**	186	3,106,337	2,616,762	339,476	14,300,000
	Daily avg audience	186	127,867	89,228	19,846	556,074
	2nd week audience	186	698,546	494,221	89,199	2,823,600
Non-Number 1	**Total audience**	508	698,432	859,205	85,827	7,805,641
	Daily avg audience	508	31,804	29,347	267	246,567
	2nd week audience	508	163,198	181,988	5,786	1,768,351

Table 7-9 shows the comparison between the results with the dependent variable being the total audience, daily average audience and 2nd week audience respectively. For the estimates of the polynomial regression, I have reported the minimum and maximum value of #1 coefficients that I have attained from different polynomial orders. Numbers in the interpretation row are derived by dividing the estimated #1 coefficients by the #1 movies' mean value of each dependent variables. I also show the average interpretation value of minimum and maximum of each model from (3) to (5).

Table 7 compares the results using total audience as a dependent variable. From the benchmark regression, model (1) – (2), being #1 accounts for 38 – 42% of the performance. However, the value drops to 26.9% with polynomial regression under RD design setting, in model (3). I include my control variables in model (4) and this leads to the coefficient dropping even more, to 17.4%. In model (5), as I include the week dummy variables, the #1 effect further falls to 15.27%.

Table 8 compares the results using daily average audience as a dependent variable. From the benchmark regression, model (1) – (2), being #1 accounts for 33 – 34% of the performance. The value drops to 18.7% with polynomial regression under RD design setting, in model (3). In model (4), the coefficient drops even more, to 11.3% and in model (5), to 8.0%.

Lastly, Table 9 shows the results using 2nd week audience as a dependent variable. Likewise, the #1 effect decreases under the RD design and more if I control for movie characteristics.

As I compare the results among outcome variables with different periods, I figure out that the causal effect of being #1 on the total audience is relatively bigger than the other two outcome variables. Daily average audience and 2nd week audience have similar levels of interpretation.

Overall, my models of benchmark regression and polynomial regression show consistency with different dependent variables. Comparing with model (1) – (2), the #1 effect derived from model (3) show a considerable decrease. In addition, controlling for movie characteristics using model (4), the estimate decreases furthermore. On average, the causal effect of being #1 during the first weekend is explains 15% of total audiences, 8% for daily average audiences and 11% for 2nd week audiences.

Table 7
Comparison of results – Total audience

Total audience	Benchmark Results		RD Design Results - Polynomial Regression					
	(1)	(2)	(3) min	(3) max	(4) min	(4) max	(5) min	(5) max
Number 1	1178952.2*** (153858.4)	1298703.2*** (203889.3)	736817.3*** (268535.2)	934526.9** (438825.4)	361373.0* (202011.6)	717117.8** (362212.7)	458316.1** (184475.1)	488765.9** (236802.8)
Movie Characteristic	O	O	X	X	O	O	O	O
Week dummy	X	O	X	X	X	X	O	O
order of polynomial	•	•	2	6	2	6	1	2
Sample size	694	694	694	694	694	694	694	694
R-square	0.628	0.806	0.506	0.519	0.687	0.694	0.872	0.875
Interpretation	38.0%	41.8%	23.7%	30.1%	11.6%	23.1%	14.8%	15.7%
			Model (3) avg: 26.9%		Model (4) avg: 17.4%		Model (5) avg: 15.27%	

Notes: ***, **, * significant at the 1%, 5% and 10% level. Robust standard errors in parentheses.

Table 8
Comparison of results – Daily average audience

Daily average audience	Benchmark Results		RD Design Results - Polynomial Regression					
	(1)	(2)	(3) min	(3) max	(4) min	(4) max	(5) min	(5) max
Number 1	42371.9*** (5077.5)	43804.1*** (6058.9)	18256.5* (10191.4)	29458.8*** (6097.2)	10725.2* (5702.9)	18152.0*** (4542.8)	9769.3** (4265.9)	10617.8* (5465.0)
Movie Characteristic	O	O	X	X	O	O	O	O
Week dummy	X	O	X	X	X	X	O	O
order of polynomial	•	•	4	1	2	1	1	2
Sample size	694	694	694	694	694	694	694	694
R-square	0.716	0.862	0.663	0.655	0.798	0.796	0.944	0.945
Interpretation	33.1%	34.3%	14.3%	23.0%	8.4%	14.2%	7.6%	8.3%
			Model (3) avg: 18.7%		Model (4) avg: 11.3%		Model (5) avg: 8.0%	

Notes: ***, **, * significant at the 1%, 5% and 10% level. Robust standard errors in parentheses.

Table 9

Comparison of results – 2nd weekend audience

2nd week audience	Benchmark Results		RD Design Results - Polynomial Regression				
	(1)	**(2)**	**(3) min**	**(3) max**	**(4) min**	**(4) max**	**(5)**
Number 1	239864.6***	255852.5***	93610.7**	169443.8***	63530.2*	111056.1***	77673.8***
	(30643.0)	(33321.5)	(51380.0)	(34620.0)	(37203.8)	(28187.7)	(29956.6)
Movie Characteristic	O	O	X	X	O	O	O
Week dummy	X	O	X	X	X	X	O
order of polynomial	•	•	3	1	2	1	1
Sample size	694	694	694	694	694	694	694
R-square	0.675	0.853	0.630	0.625	0.751	0.749	0.928
Interpretation	34.3%	36.6%	13.4%	24.3%	9.1%	15.9%	11.1%
			Model (3) avg: 18.9%		Model (4) avg: 12.5%		

Notes: ***, **, * significant at the 1%, 5% and 10% level. Robust standard errors in parentheses.

5. Conclusion

I plan to discuss the conclusion part later.

References

Anderson and Magruder, 2011, ‘Learning from the Crowd: Regression Discontinuity Estimates of the Effects of an Online Review Database’, *The Economic Journal*, 2012, 122 (563), pp957-989

Cabral and Natividad, 2016, ‘Box-Office Demand: The Importance of Being #1’, *Journal of Industrial Economics,* 64, pp277-294.

Chang et al, 2009, ‘Elaborating Movie Performance Forecast Through Psychological Variables: Focusing on the First Week Performance’, *Korean Journal of Journalism & Communication Studies*, 53 (4), pp346-371.

Einav, L., 2007, ‘Seasonality in the U.S. Motion Picture Industry,’ *Rand Journal of Economics*, 38, pp. 127-145.

Hainmueller and Kern, 2008, ‘Incumbency as a source of spillover effects in mixed electoral systems: Evidence from a regression-discontinuity design’. *Electoral Studies*, 27, pp213-227.

Lee, D.S., 2008, ‘Randomized experiments from non-random selection in U.S. house elections.’, *Journal of Econometrics,* 142 (2), pp675 – 697.

Lee and Lemieux, 2010, ‘Regression Discontinuity Designs in Economics’, *Journal of Economic Literature*, 48, pp281-355.

McCrary, 2008, ‘Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test.’ *Journal of Econometrics*, 142(2): 698–714.

Moul, C., 2007, ‘Measuring Word of Mouth’s Impact on Theatrical Movie Admissions,’ *Journal of Economics & Management Strategy*, 16, pp. 859-892.

Park and Jung, 2009, ‘The Determinants of Motion Picture Box Office Performance: Evidence from Movies Released in Korea, 2006-2008’, *Journal of Communication Science*, 9(4), pp243-276.

Appendix

Recall that the margin of victory (MV) was calculated by margin of the #1(#2/#3 and above) movie's first weekend audience and the #2(#1) movie in the same weekend. The latter movie did not have to be in its first weekend. Table A1-A3 show the results from samples with MVs calculated with the latter movies also in the same opening week. Only 363 movies have MVs calculated with movies that share same opening weeks. The estimates attained from the restricted samples are overall similar with the results from full sample. I am planning to increase my number of observations by adding 5 more years from 2004 to 2009, and using only the restricted samples for more precise estimates of the #1 effect.

Table A1

Comparison of results – Total audience (including same weekend opening samples)

Total audience	**Benchmark Results**		**RD Design Results - Polynomial Regression**										
	(1)	**(2)**	**(3-a) min**	**(3-a) max**	**(3-b) min**	**(3-b) max**	**(4-a) min**	**(4-a) max**	**(4-b) min**	**(4-b) max**	**(5-a) min**	**(5-a) max**	**(5-b)**
Number 1	1178952.2***	1298703.2***	736817.3***	934526.9**	703596.0**	821700.5***	361373.0*	717117.8**	570963.4**	674364.4***	458316.1**	488765.9**	434720.9**
	(153858.4)	(203889.3)	(268535.2)	(438825.4)	(306296.6)	(237408.0)	(202011.6)	(362212.7)	(240043.0)	(184695.5)	(184475.1)	(236802.8)	(183623.7)
Movie Characteristic	O	O	X	X	X	X	O	O	O	O	O	O	O
Week dummy	X	O	X	X	X	X	X	X	X	X	O	O	O
Order of polynomial	•	•	2	6	2	1	2	6	2	1	1	2	1
Sample size	694	694	694	694	363	363	694	694	363	363	694	694	363
R-square	0.628	0.806	0.506	0.519	0.531	0.530	0.687	0.694	0.724	0.724	0.872	0.875	0.876
Interpretation	38.0%	41.8%	23.7%	30.1%	22.7%	26.5%	11.6%	23.1%	18.4%	21.7%	14.8%	15.7%	14.0%
			(3-a) 26.9%		**(3-b)** 24.6%		**(4-a)** 17.4%		**(4-b)** 20.0%		**(5-a)** 15.3%		14.0%

Notes: ***, **, * significant at the 1%, 5% and 10% level. Robust standard errors in parentheses.

Table A2

Comparison of results – Daily average audience (including same weekend opening samples)

Daily average audience	Benchmark Results		RD Design Results - Polynomial Regression										
	(1)	(2)	(3-a) min	(3-a) max	(3-b) min	(3-b) max	(4-a) min	(4-a) max	(4-b) min	(4-b) max	(5-a) min	(5-a) max	(5-b)
Number 1	42371.9***	43804.1***	18256.5*	29458.8***	22568.4**	26971.6***	10725.2*	18152.0***	17127.6**	19591.9***	9769.3**	10617.8*	8471.7**
	(5077.5)	(6058.9)	(10191.4)	(6097.2)	(8838.6)	(6861.8)	(5702.9)	(4542.8)	(6640.3)	(5211.1)	(4265.9)	(5465.0)	(4204.6)
Movie Characteristic	O	O	X	X	X	X	O	O	O	O	O	O	O
Week dummy	X	O	X	X	X	X	X	X	X	X	O	O	O
Order of polynomial	•	•	4	1	2	1	2	1	2	1	1	2	1
Sample size	694	694	694	694	363	363	694	694	363	363	694	694	363
R-square	0.716	0.862	0.663	0.655	0.684	0.682	0.798	0.796	0.833	0.833	0.944	0.945	0.945
Interpretation	33.1%	34.3%	14.3%	23.0%	17.7%	21.1%	8.4%	14.2%	13.4%	15.3%	7.6%	8.3%	6.6%
			(3-a) 18.7%		**(3-b)** 19.4%		**(4-a)** 11.3%		**(4-b)** 14.35%		**(5-a)** 7.95%		6.6%

Notes: ***, **, * significant at the 1%, 5% and 10% level. Robust standard errors in parentheses.

Table A3

Comparison of results – 2nd week audience (including same weekend opening samples)

2nd week audience	Benchmark Results		RD Design Results - Polynomial Regression									
	(1)	(2)	(3-a) min	(3-a) max	(3-b) min	(3-b) max	(4-a) min	(4-a) max	(4-b) min	(4-b) max	(5-a) min	(5-b)
Number 1	239864.6***	255852.5***	93610.7**	169443.8***	115765.8**	174780.0**	63530.2*	111056.1***	97614.6**	141127.7***	77673.8***	76826.6***
	(30643.0)	(33321.5)	(51380.0)	(34620.0)	(52162.4)	(40580.8)	(37203.8)	(28187.7)	(43941.5)	(33816.0)	(29956.6)	(28640.8)
Movie Characteristic	O	O	X	X	X	X	O	O	O	O	O	O
Week dummy	X	O	X	X	X	X	X	X	X	X	O	O
Order of polynomial	•	•	3	1	2	1	2	1	2	1	1	1
Sample size	694	694	694	694	363	363	694	694	363	363	694	363
R-square	0.675	0.853	0.630	0.625	0.627	0.624	0.751	0.749	0.775	0.773	0.928	0.921
Interpretation	34.3%	36.6%	13.4%	24.3%	16.6%	25.0%	9.1%	15.9%	14.0%	20.2%	11.1%	11.0%
			(3-a) 18.9%		**(3-b)** 20.8%		**(4-a)** 12.5%		**(4-b)** 17.1%		**(5-a)** 11.1%	**(5-b)** 11.0%

Notes: ***, **, * significant at the 1%, 5% and 10% level. Robust standard errors in parentheses.