CESifo
AREA CONFERENCES 2021

Economics of Digitization

Munich, 19 – 20 November 2021

Quality Disclosures and Disappointment: Evidence from the Academy Awards

Michelangelo Rossi

CESifo

Quality Disclosures and Disappointment: Evidence from the Academy Awards*

Michelangelo Rossi†

July 22, 2021

Abstract

This paper studies the impact of quality disclosures on buyers' ratings using data from an online recommender system. Disclosures may alter expectations on sellers' quality and affect buyers' rating behavior. In particular, if buyers' utility depends on their expectations, a positive disclosure of quality such as an award may lead to buyers' disappointment with a negative influence on their ratings. I identify the disappointment effect in moviegoers' ratings originated from the rise in expectations due to movies' nominations for the Academy of Motion Picture Arts and Sciences awards. I control for the selection of moviegoers who watch and rate movies before or after nominations with a non-parametric matching technique. After nominations, ratings for nominated movies significantly drop relative to ratings for movies that were not nominated. Disappointment reduces the rating premium of nominated movies by more than five percent in the next thirty days after the nomination.

Keywords: Quality Disclosure, Expectation Formation, Reference Point, Disappointment
JEL codes: D82, D83, D84, D91

*I would like to thank Natalia Fabra and Matilde Pinto Machado for their patient guidance and advice. I also thank Alan Crawford, Juan José Dolado, Jesús Fernández-Huertas Moraga, Jaime Millán-Quijano, Alvaro Name-Correa, Jan Stuhler, and Kevin Ducbao Tran. Helpful feedback was received the internal seminar at Universidad Carlos III de Madrid, at the 19th International Industrial Organization Conference, the 7th AFREN Digital Economics Summer School, the 22nd ACM Conference on Economics and Computation (EC'21), and the 48th European Association Research Industrial Economics Conference. All errors are mine.

†Télécom Paris, Institut Polytechnique de Paris, France. E-mail: michelangelo.rossi@telecom-paris.fr.

1 Introduction

Certifications, and in general, third-party quality disclosures are often used in markets with asymmetries of information. In these circumstances, buyers are not perfectly informed about the seller's quality. Thus, a third-party certification may help to reduce the uncertainty on the buyers' side and increase their willingness to pay. Used cars' sellers may show the most recent inspections by the car's manufacturer to assure prospective buyers about the good state of the car. The positive effect of quality disclosure has already been studied in many contexts affected by asymmetric information.[1] When a certifier is credible, a third-party disclosure can effectively increase buyers' expectations and attract high-quality sellers to trade.

Still, altering buyers' expectations could have unexpected side-effects when buyers' utility depends on reference points induced by their expectations. With reference-dependent preferences, agents' reference points affect their utility throughout a "gain/loss" component which describes their perception of elation or disappointment. The role of agents' expectations to explain reference point formation received the attention by many scholars from a theoretical perspective (Kahneman and Tversky, 1979, Loomes and Sugden, 1982; Bell, 1985; Kőszegi and Rabin, 2006). Empirical evidence suggests that expectation-dependent preferences effectively describe agents' behavior in different settings (Camerer, Babcock, Loewenstein and Thaler, 1997; Crawford and Meng, 2011; Card and Dahl, 2011; Bartling, Brandes and Schunk, 2015; Backus, Blake, Masterov and Tadelis, 2017). Accordingly, disclosures by third-parties improving buyers' expectations on sellers' quality may increase the chances of buyers' disappointment and reduce the benefits of lower informational asymmetries.

This paper empirically identifies the disappointment effect due to quality disclosures estimating the causal impact of the nominations for the Academy of Motion Picture Arts and Sciences (AMPAS) awards on movie ratings. Nominations constitute the "shifter" of reference point about the movies' quality. The variations of ratings displayed on an online movie recommender system (MovieLens) provide a measure for the disappointment effect. I control for the selection of moviegoers who watch and rate movies before or after nominations with a non-parametric matching technique. After nominations, ratings for nominated movies significantly drop relative to ratings for not nominated movies with similar characteristics. In particular, the drop in ratings due to disappointment accounts for more than five percent of the rating premium received by the nomination.

This empirical exercise helps to shed light on the welfare impact of quality disclosures in a setting of asymmetry of information and expectation-dependent preferences. I do not question the positive impact of certifications and awards on sellers' performance. Still, quality disclosures also produce depressing effects on buyers' satisfaction (ratings) due to disappointment. This latter consideration

[1] Dranove and Jin (2010) present a detailed review of the power (and the limitations) of these tools to provide a credible signal for sellers' quality.

reduces the positive effects of certifications in terms of profits.

The mechanism presented here may broadly apply to all settings with asymmetries of information. Yet, the movie context presents specific advantages to study disappointment. The main advantage regards the absence of variations in prices charged by movie theaters before and after the Academy Awards nominations. Accordingly, by looking at differences in ratings, variations in prices do not influence the disappointment effect. In general, this is not the case for sellers who receive a certification. Given the increase in buyers' willingness to pay due to the certification, sellers may adjust and charge a higher price. Empirically, variations in prices are problematic since disappointment cannot be disentangled from a reduction in buyers' utility due to a higher price.[2] Moreover, movie quality is fixed over time, and ratings are not affected by changes in sellers' ongoing effort. Thus, the main challenge for identifying disappointment through the ratings' variation regards the selection of moviegoers who watch and rate the movie before or after the nomination. Due to the "quality disclosure", a different profile of moviegoers watches and rates movies: variations in ratings may depend on differences in tastes or preferences of moviegoers.

In this paper, first, I identify the impact of the AMPAS nominations on ratings of nominated movies with a difference-in-difference design (DiD). The results show a negative and significant drop in ratings for nominated movies after nomination relative to not nominated movies rated in the same period. Although this identification procedure allows to control for characteristics of moviegoers, it does not entirely disentangle selection from disappointment. Then, I present a second identification strategy to account for the selection of moviegoers. Here the response variable is the difference in ratings reported by the same individual for a couple of movies: a nominated movie and a not nominated movie that share several common features such as the genre. I show that, studying the variations of this difference, I can reduce the impact of selection on the ratings' variations over time and identify the disappointment effect. The results for both designs show a negative and significant drop in ratings for nominated movies after nominations.

This paper contributes to the literature regarding quality disclosures, reference-dependent preferences, and online reviews. The literature about quality disclosures has focused on the impact of these devices to improve sellers' performances and market outcomes. Jin and Leslie (2003) show that the introduction of restaurant hygiene report cards reduced food-related illnesses in Los Angeles with cleaner establishments attracting more consumers. Similarly, Chezum and Wimmer (1997) document a positive effect of certifications for racehorses in terms of higher prices and better racing performances. For what concerns digital platforms, Elfenbein et al. (2015), Hui, Saeedi, Shen and Sundaresan (2016), and Hui, Saeedi, Spagnolo and Tadelis (2018) study the role of certification

[2] Elfenbein, Fisman and McManus (2015) show that eBay sellers charge significantly higher prices after receiving the eBay's "top-rated seller" certification. Moreover, Li and Hitt (2010) investigate how changes in prices can bias consumer reviews when these reflect the difference between quality and prices.

programs in eBay and they report positive effects over seller's quality. Conversely, in an online platform for residential home services, Farronato, Fradkin, Larsen and Brynjolfsson (2020) observe that professionals showing their licensing status do not increase their number of transactions or charge higher prices. For what concerns the AMPAS awards, several articles (Donihue, Nelson, Waldman and Wheaton, 2001; McKenzie, 2012) document their positive impact on movies' box office; still, up to my knowledge, the effect of these awards on users' ratings has never been studied.

Few papers try to investigate the adverse effects of quality disclosures: Dranove, Kessler, McClellan and Satterthwaite (2003) show that after a hospital "Report Card" program was introduced in New York, hospitals started to avoid potential low grades declining to treat more "difficult" patients affected by severe diseases. Ho (2012) describes several flaws of the US restaurant hygiene system showing inflation in grades and a shift of inspection resources to resolve disputes in grading. Finally, Forbes, Lederman and Tombe (2015) investigate how a disclosure program for airline on-time performance distorted the incentives of airlines to manipulate arrival times. Yet, to my knowledge, this is the first paper to empirically show the side effects of quality disclosure in terms of perceived utility on the buyers' side. In particular, my results show that disappointment effect may be present even when the certification leads to better sellers' performances.

My work also contributes to the empirical literature about reference-dependent preferences. In his review, Barberis (2013) points out that few papers document the relevance of reference-based preferences outside the areas of finance and choice under uncertainty. My paper shows the importance of reference points in a new setting related to the introduction of quality disclosures. In the literature of reference point formation, the work by Backus et al. (2017) is the closest to mine. They study reference points in an online setting and focus on a hybrid format of the online auctions in eBay in which each buyer has the "Buy-It-Now" (BIN) option to buy at any moment the product at a specific price without taking part in the auction. They find that eBay's buyers who presented the highest bid and lost the auction since another buyer used the BIN option, have significantly higher chances to leave the platform. They interpret this behavior through the lenses of reference-dependent preferences and argue that buyers' disappointment is the main driver for their exit choice.

Finally, I also contribute to the growing literature about potential bias in online reviews. Several papers show that online reviews may be biased since reviewers may act strategically (Klein, Lambertz and Stahl, 2016, Mayzlin, Dover and Chevalier, 2014), or being influenced by previous reviews, social comparison and users' reciprocity (Talwar, Jurca and Faltings, 2007, Chen, Harper, Konstan and Li, 2010, Moe and Schweidel, 2012, Proserpio and Zervas, 2017, Proserpio, Xu and Zervas, 2018). Regarding movie reviews, the closest papers to mine are Lee, Hosanagar and Tan (2015) and Bondi and Stevens (2019). Lee et al. (2015) show that the behavior of online reviewers is influenced by prior ratings, especially when reviewers' friends have rated a movie. They observe that friends' ratings induce reviewers' herding behavior. Bondi and Stevens (2019) show that consumer heterogeneity affects social learning about movies' quality through ratings. In particular, expert users watch better movies, but they also have higher expectations lowering their ratings. Accordingly, my paper presents additional evidence about the role of "external" factors influencing reviewers' behavior. In contrast with the previous literature, I analyze the impact of expectations taking advantage of an exogenous change due to the AMPAS nominations.

The paper proceeds as follows. In Section 2, I provide some background context regarding the platform of movie recommendations MovieLens, and its rating process; then, I present the dataset. This is followed by a short theoretical framework in Section 3. I discuss my identification strategy in Section 4. Section 5 provides the empirical findings. Section 6 concludes. Additional tables, figures and extensions of the results are in Appendix.

2 Empirical Setting and Dataset

In this Section, I first present the movie recommendations platform, MovieLens, and explain the timing and functioning of its rating process. Then, I describe the MovieLens 20M Dataset and how I enriched it with information about nominations for the AMPAS Awards. Finally, I provide descriptive statistics concerning nominated and not nominated movies before and after nominations.

2.1 MovieLens

Run by GroupLens, a research group of the Department of Computer Science and Engineering at the University of Minnesota, MovieLens (http://www.movielens.org) is an online platform that provides movie recommendations for its users. Inscribed users are invited to rate movies. Then, following their preferences, the MovieLens algorithm creates a list of personalized recommended movies.

MovieLens was launched in 1997 and from that time it has become a visible recommendation system. According to the work by Chen et al. (2010), in April 2006, over 13 million user ratings of 9,043 movies were present in MovieLens. Although the rating process experienced some minor changes in the design of the platform, the main mechanism remained unaltered.

Here I briefly describe the current rating process.[3] Appendix Figure A.1 shows a snapshot of the MovieLens main page in which a group of movies is listed. If a user is interested in a movie, or a genre, she may type the desired query on the bar of top of the main page to get a more refined search. After clicking on a movie title, a user is redirected to the movie webpage (Appendix Figure A.2) where she has access to further information.Then, each user can rate a movie on a 0.5 to 5 star scale; and future recommendations are affected by the rating.

[3] Harper and Konstan (2015) present a precise description of the changes in the recommendation algorithm and in the rating process occurred during the years on the platform.

2.2 MovieLens 20M Dataset

The GroupLens group makes publicly available different types of datasets regarding MovieLens ratings. I use the "MovieLens 20M" Dataset generated in October 2016. It contains 20,000,263 ratings for 27,278 movies by 138,493 users displayed between January 1995 and March 2015. Each rating is linked with identification numbers of the movie and the user, and the date in which the rating is displayed on the platform. Not the entire universe of ratings from 1995 to 2005 is present in the dataset. Users were randomly selected among all users having rated at least twenty movies. No information is available for users apart from their identification number. Conversely, some movie characteristics (such as the movie genre or release year) are directly available. Moreover, I use additional information scraping other movies' characteristics with the corresponding link to the IMDb webpage. I have detailed information about the movies' director, main actors in the cast, movie's language, country of production, and release day.

2.3 AMPAS Academy Awards

Since 1929, the Academy of Motion Picture Arts and Sciences (AMPAS) has assigned awards for excellence in cinematic achievements. The awards are divided into several categories.[4] Still, some particularly relevant categories (the so-called "Big Five") usually receive the most of the public attention: Best Picture; Best Direction; Best Leading Actor and Actress; and Best Screenplay. Every year, the procedure for the assignment of the awards consists of two steps. First, a restricted group of movies, usually composed of five movies, is nominated for each category among those movies that qualify for the award. A movie must open in the previous calendar year in Los Angeles County to qualify for the awards in a given year. This rule differs for the "Best Foreign Language Film" award. Non-US movies are selected in each country, and the nominated movies for this award are chosen among this pre-selected pool.[5]

Since 2004, the nominated movies have been announced in mid-January. Before 2004, the results were released at the beginning of February. Then, six weeks after the announcement, the AMPAS awards are presented with a ceremony in one of the main theaters in Hollywood. To evaluate the impact of the nominations to the AMPAS awards, I consider all the nominated and awarded movies from 1995 to 2015, focusing on the nominations for the "Big Five" awards and the awards for the "Best Foreign Language Film", the "Best Documentary Feature", and the "Best Animated Feature Film". Accordingly, I select 544 nominated movies. Among them, 522 movies are matched with the

[4] A number of categories have been discontinued. From 2001, the same group of twenty-four categories has been used for the awards.

[5] The official AMPAS webpage (https://www.oscars.org/oscars/rules-eligibility) provides more information about the movies' selection process.

MovieLens dataset.[6] Moreover, for each year, I use the day of the nomination to establish if a rating for a nominated movie was displayed on the platform before or after the nomination.[7]

In order to do the same for movies that were not nominated, I associate each not nominated movie with a nomination date. To do that, I follow a similar criterion to the one used by the Academy to select movies that qualify for the awards. I consider the date of the first rating appearing on MovieLens for each movie; then, I select the first nomination ceremony after this date as the reference year of nomination. Finally, I remove all movies if their first rating on the platform is displayed more than two years after their year of production. In this way, I abstract from the ratings for old movies that cannot be compared with the most recent ones.

2.4 Descriptive Statistics

I conclude this Section with some relevant descriptive statistics about nominated and not nominated movies present in the dataset. I consider only movies whose first rating is displayed on the platform in the first two years after the year of production. First, I report movies' statistics regarding all ratings available. Then, I focus on the window of 120 days around the AMPAS nominations (60 days before and after), in line with the the empirical designs presented in Section 4.

Table 1 shows and compares some relevant statistics for nominated and not nominated movies using all ratings available for each movie. The first three rows compare the average ratings and show that nominated movies have higher ratings (0.62 stars), in line with the assumption that nominated movies have higher quality. The average ratings for not nominated movies seem to not vary before and after the nominations, whereas a slight decrease is observable for nominated movies. Nominated movies are also rated more frequently before and after nominations; they are rated for a longer period (Row 7) and they start to be rated before not nominated movies (Row 8). To visualize the difference in ratings between movies, Appendix Figure A.3 shows the distributions of ratings for nominated and not nominated movies considering only ratings displayed before nominations.

Nominated and not nominated movies have also different genres (Rows 9 - 13) and they match with a different pool of users. Rows 14 and 15 show the average ratings reported by MovieLens users who rate not nominated and nominated movies before and after nominations. Nominated movies seem to attract users with higher average ratings relative to not nominated movies. That can be explained by users differing in the probability to report a high rating; or by the fact that users are attracted by different movies. Certain users may only watch high-quality movies and thus they report high ratings, in line with Bondi and Stevens (2019).

[6] The unmatched nominated movies are documentaries (17) and foreign movies (5).

[7] The information about the nomination dates is extracted from the official AMPAS webpage and the associated Wikipedia entries.

Table 1: Summary Statistics: Not Nominated and Nominated Movies

	Not Nom. (1)	SD	Nom. (2)	SD	(1)-(2)	p-value
Average Ratings						
Total	3.12	0.58	3.73	0.23	-0.62	0.00
Before Nominations	3.11	0.69	3.84	0.34	-0.74	0.00
After Nominations	3.13	0.59	3.73	0.23	-0.60	0.00
Number of Ratings						
Total	955.29	3,246.14	5,048.08	8,493.16	-4,092.78	0.00
Before Nominations	172.47	873.88	360.00	622.40	-187.53	0.00
After Nominations	782.82	2,655.42	5,448.19	8,175.50	-4,665.37	0.00
Rating Period (date - date)						
Last rating - First rating	2,521.06	2,197.51	3,297.15	2,043.76	-776.10	0.00
First rating - Release	263.22	273.67	152.68	165.30	110.54	0.00
Genres (%)						
Action	14.34	-	6.60	-	-	-
Adventure	4.19	-	8.40	-	-	-
Comedy	26.45	-	15.20	-	-	-
Drama	27.41	-	41.80	-	-	-
Others	27.61	-	28.00	-	-	-
Average Ratings for Users						
Before Nominations	3.37	0.25	3.51	0.16	-0.14	0.00
After Nominations	3.43	0.21	3.56	0.09	-0.13	0.00
Number of Ratings by Users						
Before Nominations	1,185.59	886.93	909.48	535.97	276.10	0.00
After Nominations	990.66	507.19	750.93	389.90	239.73	0.00
Number of movies	10,817	-	507	-	-	-
Number of ratings	9,256,797		2,524,039	-	-	-

Note: In the first four sections, the table compares not nominated and nominated movies in terms of average ratings (measured on a 0.5 to 5 star scale), number of ratings, rating periods, and genres. In the last two sections, the table compares the characteristics of users who rate nominated and not nominated movies before and after the nomination. In particular, I study the average rating reported by users and the total number of ratings present on the platform for each user. All ratings are considered for all not nominated and nominated movies if their first rating is displayed on the platform in the first two years after the year of production. The last two columns show the differences between the averages and the $p-value$ of the difference.

Table 2: Not Nominated and Nominated Movies 120 Days around the AMPAS Nominations

	Not Nom.	SD	Nom.	SD	Δ (1)-(2)	p-value
Average Ratings						
from −60 to −30 days	3.18	0.77	3.84	0.42	-0.66	0.00
from −30 to −20 days	3.18	0.81	3.84	0.43	-0.66	0.00
from −20 to −10 days	3.17	0.80	3.84	0.49	-0.67	0.00
from −10 to 0 days	3.18	0.82	3.83	0.48	-0.65	0.00
from 0 to 10 days	3.19	0.81	3.84	0.49	-0.65	0.00
from 10 to 20 days	3.18	0.79	3.80	0.45	-0.62	0.00
from 20 to 30 days	3.18	0.82	3.82	0.49	-0.63	0.00
from 30 to 60 days	3.20	0.75	3.82	0.43	-0.62	0.00
Number of Ratings						
from −60 to −30 days	19.56	60.57	61.27	104.32	-41.71	0.00
from −30 to −20 days	6.66	18.17	27.72	42.38	-21.06	0.00
from −20 to −10 days	7.51	21.94	28.93	41.71	-21.42	0.00
from −10 to 0 days	6.45	16.67	29.16	38.29	-22.71	0.00
from 0 to 10 days	5.48	13.85	27.60	34.09	-22.12	0.00
from 10 to 20 days	5.68	15.13	29.48	35.37	-23.80	0.00
from 20 to 30 days	6.14	18.72	29.64	36.45	-23.50	0.00
from 30 to 60 days	18.55	58.52	93.68	111.77	-75.12	0.00
Number of movies	10,817	-	507	-	-	-
Number of ratings	619,460	-	150,288	-	-	-

Note: The table compares not nominated and nominated movies in terms of average ratings and number of ratings in the window of 120 days around the AMPAS nominations . All ratings are considered for all not nominated and nominated movies if their first rating is displayed on the platform in the first two years after the year of production. The last two columns show the differences between the averages and the $p-value$ of the difference.

The statistics reported in Table 2 focus on the window of 120 days around the AMPAS nominations. The first panel (Rows 1-7) shows the evolution of the average ratings for not nominated and nominated movies before and after the nominations. Not nominated movies present stable ratings around the nomination day. Conversely, the ratings of nominated movies slightly drop (0.05 stars) around 10 days after the nominations. The second panel (Rows 8-14) shows the evolution of the number of ratings displayed before and after the nominations. Nominated movies are rated almost five times more than not nominated movies. Still, the dynamics of the number of ratings is quite stable for both groups. Appendix Figures A.4 and A.5 show the dynamics of the arrival of movie ratings. The great majority of ratings occurs after nominations since movies are often released a few months before the AMPAS nominations and awards.

3 Theoretical Framework

In this Section, I present a conceptual framework to model the effect of the AMPAS nominations on users' reviewing process. This theory is developed to clarify the two main channels through which a shift in expectations may affect movie ratings: disappointment/elation and selection. Accordingly, I start with a formal description of the user's decision to watch and rate a movie keeping fixed users' expectations. Then, I study the impact of a change in expectations due to a shock, such as the AMPAS nominations, and I present the empirical predictions derived from this framework.

3.1 The Rating Process

Each movie is defined by $M+1$ parameters: a quality parameter $\theta_i \in \mathbb{R}$, and a vector $\boldsymbol{\mu}_i \in \mathbb{R}^M$ describing the movie "position" in the feature space with M dimensions (for instance: the movie's genres). Similarly, each user is defined by $M+1$ parameters: $v_j \in \mathbb{R}$ regarding the utility that user j receives by going to the theater (irrespectively of the movie watched), and a vector $\boldsymbol{\pi}_j \in \mathbb{R}^M$ describing the "position" of user j preferences in the feature space (the genres preferred by the user).

Before watching movie i, users do not observe the quality of a movie. Thus, the expected utility for user j watching movie i is defined as follows:

$$E(u_{ij}) = E(\theta_i) + v_j - d(\boldsymbol{\pi}_j, \boldsymbol{\mu}_i), \tag{3.1}$$

where the function $d(.)$ is a strictly increasing, continuous norm describing the distance between the position of movie i and the preferences of user j. Accordingly, it is possible to rewrite the expected utility as $E(u_{ij}) = E(\theta_i) + \alpha_{ij}$ with $\alpha_{ij} = v_j - d(\boldsymbol{\pi}_j, \boldsymbol{\mu}_i)$ representing the sum of the parameters related to the utility derived by going to the theater and the "match" value in terms of utility between user j and movie i.

After watching movie i, the quality is observed by user j. In line with Kőszegi and Rabin (2006), the expected utility $E(u_{ij})$ defines the reference point for user j about movie i. Thus, the ex-post utility for user j after watching movie i is the following:

$$\begin{aligned} u_{ij} &= \theta_i + \alpha_{ij} + \gamma(\theta_i + \alpha_{ij} - E(u_{ij})) \\ &= \theta_i + \alpha_{ij} + \gamma(\theta_i - E(\theta_i)), \end{aligned} \tag{3.2}$$

where $\gamma(.)$ represents the user's gain-loss utility factor as in Kőszegi and Rabin (2006). In the reference point formation literature, γ is a function of the gain/loss term $\theta_i - E(\theta_i)$ (Kahneman and Tversky, 1979; Loomes and Sugden (1982); Kőszegi and Rabin, 2006; Backus et al., 2017). For simplicity, I assume $\gamma(.)$ to be linear with the parameter γ multiplying the gain/loss term. It is possible to capture the disappointment effect in users' utility if $\gamma > 0$: the higher the expectations of user j regarding the quality of movie i, $E(\theta_i)$, the lower the user j gain-loss utility term.

For each movie i, the following timing describes the process through which users are informed about the movie, they decide to watch it and, in such a case, rate it:

1. Movie i is released and a unit measure of potential users associated with movie i is formed. Users are heterogeneous in α_{ij} and the distribution of α_{ij} among all potential users is $F(\alpha)$;

2. A signal s appears and a proportion $\lambda(s) \in (0,1)$ becomes aware of movie i. The distribution of α_{ij} among aware users after signal s is $F_s(\alpha)$;

3. Aware users form expectations about the quality of the movie $E(\theta_i|s)$ and watch movie i if $E(u_{ij}|s) = E(\theta_i|s) + \alpha_{ij} > 0$;

4. Finally, users who watch the movie always rate it reporting their ex-post utility. Accordingly, the rating of user j for movie i, r_{ij}, is equivalent to the ex-post utility: $r_{ij} = u_{ij}$.[8]

3.2 The Impact of Signals

Signals change the profile of users who are aware of movies ($F_s(\alpha)$ depends on s) and they affect the quality that all aware users expect by movies. Accordingly, we may interpret signal s as the result of an advertising campaign promoted by the movie producer, or of a quality disclosure such as the AMPAS nominations: they jointly impact the expectations about a movie and the potential audience becoming aware of the movie release. These two effects influence the profile of users who decide to watch the movie and, as a final result, the observed ratings. Specifically, the expected ratings of movie i given quality θ_i and signal s is the following:

$$E(r_{ij}|\theta_i, s) = \theta_i + \int_{\alpha > -E(\theta_i|s)} \alpha dF_s(\alpha) + \gamma(\theta_i - E(\theta_i|s)). \tag{3.3}$$

To have a better sense of the impact on ratings from a shift in the expected movie quality, I assume signals to change from s to s^+ such that $E(\theta_i|s) < E(\theta_i|s^+)$. The difference in expected ratings is the

[8]Users' rating decision may depend on their ex-post utility levels. In particular, Dellarocas and Wood (2008) and Hu, Pavlou and Zhang (2017) suggest that online users who rate products are more likely to be the ones who experienced more "extremely" positive or negative utility levels from the transactions. In line with this perspective, I discuss in the next Subsection how a change in expectations may affect ratings throughout the selection channel. In Appendix B, I investigate the selection into reviewing and show that the disappointment effect can be isolated also in this case.

following:

$$E(r_{ij}|\theta_i, s^+) - E(r_{ij}|\theta_i, s) = -\gamma(E(\theta_i|s^+) - E(\theta_i|s)) + \int_{\alpha > -E(\theta_i|s^+)} \alpha dF_{s^+}(\alpha) - \int_{\alpha > -E(\theta_i|s)} \alpha dF_s(\alpha).$$

I define the first term $(-\gamma(E(\theta_i|s^+) - E(\theta_i|s)))$ as the *disappointment effect* since, when $\gamma > 0$, it reflects the downward effect on ratings of an increase in expectations. Conversely, the second term $(\int_{\alpha > -E(\theta_i|s^+)} \alpha dF_{s^+}(\alpha) - \int_{\alpha > -E(\theta_i|s)} \alpha dF_s(\alpha))$ represents the *selection effect* on the profile of users who decide to watch movie i resulting from a shift in signal. While disappointment has always a downward effect on ratings (with $\gamma > 0$), the effect of selection is ambiguous. With $F_s(\alpha) = F_{s^+}(\alpha)$, selection drives ratings downward, but, if $F_s(\alpha) \neq F_{s^+}(\alpha)$, an upward effect on ratings is possible.[9]

The following observation recaps the results from the previous discussion.

Observation 1. *Assume the signal about movie's quality changes from s to s^+ such that $E(\theta_i|s) < E(\theta_i|s^+)$. The difference in movie ratings with different signals $E(r_{ij}|\theta_i, s^+) - E(r_{ij}|\theta_i, s)$ is composed by disappointment and selection. With $\gamma > 0$, disappointment moves downward users' ratings, whereas the effect of selection is ambiguous.*

3.3 Isolating the Disappointment Effect

According with Observation 1, the comparison of ratings reported by users with different signals cannot isolate the effect of disappointment from selection. Yet, this may be possible once we compare ratings posted by the same user for different movies. In particular, I denote the difference in ratings posted by user j about movie i and h as $\Delta^j_{ih} = r_{ij} - r_{hj}$. Recalling the previous definition of ratings as the ex-post users' utility, Δ^j_{ih} includes three different terms:

$$\begin{aligned} \Delta^j_{ih} &= (\theta_i - \theta_h) - (\alpha_{ij} - \alpha_{ih}) + \gamma(\theta_i - E(\theta_i|s_{ij}) - \theta_h + E(\theta_h|s_{hj})) \\ &= (\theta_i - \theta_h)(1+\gamma) - (\alpha_{ij} - \alpha_{ih}) - \gamma(E(\theta_i|s_{ij}) - E(\theta_h|s_{hj})) \\ &= (\theta_i - \theta_h)(1+\gamma) - (d(\boldsymbol{\pi}_j, \boldsymbol{\mu}_i) - d(\boldsymbol{\pi}_j, \boldsymbol{\mu}_h)) - \gamma(E(\theta_i|s_{ij}) - E(\theta_h|s_{hj})), \end{aligned} \tag{3.4}$$

where s_{ij} and s_{hj} are the signals received by user j before watching and rating movie i and movie j, respectively. The first term, $(\theta_i - \theta_h)(1+\gamma)$, is proportional to the difference in quality of the two movies and does not vary with signals; the second term, $(d(\boldsymbol{\pi}_j, \boldsymbol{\mu}_i) - d(\boldsymbol{\pi}_j, \boldsymbol{\mu}_h))$, regards the distances in terms of features of two movies relative to the preferences of user j; finally, the last term, $\gamma(E(\theta_i|s_{ij}) - E(\theta_h|s_{hj}))$ describes the difference in terms of expectations by user j.

[9] Assuming $F_s(\alpha) \neq F_{s^+}(\alpha)$ is in line with the statistics reported in Table 1, showing different average ratings by users who rate nominated movies before and after nominations. In Appendix B, I illustrate this point with a numerical example.

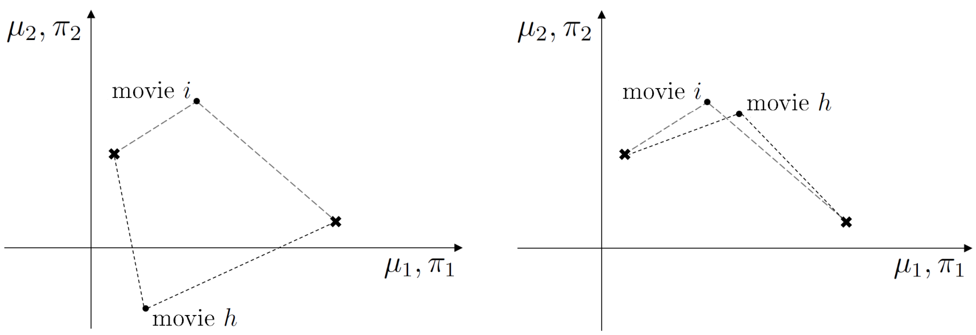


Figure 1: Variations in User Selection Comparing Different Movies

Studying the difference between ratings posted by the same user removes only a part of users' components of ratings (the utility parameter v_j). In particular, the match values between user j and the two movies do not disappear. Still, when movie i and movie h share a "similar" position in the feature space, users' components of ratings reduce as shown in Figure 1. Here, the two graphs show a two-dimensional feature space (for instance: the main and the secondary genre) in which the points denoted with a cross ($\times$) represent the preferences of two different users; whereas the points denoted with a dot ($\cdot$) represent the features of two different movies. When movie's features are not similar (like in the figure on the left), the difference $d(\boldsymbol{\pi}_j, \boldsymbol{\mu}_i) - d(\boldsymbol{\pi}_j, \boldsymbol{\mu}_h)$ is large since the distances between movie h and users are not a good predictor for their distances from movie i. Conversely, when movie's features are similar (like in the figure on the right), $d(\boldsymbol{\pi}_j, \boldsymbol{\mu}_i) - d(\boldsymbol{\pi}_j, \boldsymbol{\mu}_h)$ is small and the distances between movie h and users predict well their distances from movie i. More precisely, when movie i and movie h share the same features $\boldsymbol{\mu}_i = \boldsymbol{\mu}_h = \bar{\boldsymbol{\mu}}_{ih}$, users' components of ratings totally vanish from the difference Δ^j_{ih}.

Therefore, once we restrict the attention to movies sharing the same features, we can isolate the disappointment effect due to a shift in the expectations of movie quality. To do that, it is necessary to study a situation in which only one movie receives a shock in its signal. Therefore, we can study the expected Δ^j_{ih} when the signal for movie i changes from s_i to s_i^+, whereas the signal for movie h remains constant to s_h:

$$E(\Delta^j_{ih}|\theta_i, \theta_h, s_i^+, s_h) - E(\Delta^j_{ih}|\theta_i, \theta_h, s_i, s_h) = -\gamma(E(\theta_i|s_i^+) - E(\theta_i|s_i)). \tag{3.5}$$

To recap this result, I present the following observation.

Observation 2. *Consider two movies i and h sharing the same position in the feature space such that $\boldsymbol{\mu}_i = \boldsymbol{\mu}_h = \bar{\boldsymbol{\mu}}_{ih}$. Assume the signal about quality for movie i changes from s_i to s_i^+ such that $E(\theta_i|s) < E(\theta_i|s^+)$; and the the signal about quality for movie h remains constant to s_h. The difference in Δ^j_{ih} with different signals $E(\Delta^j_{ih}|\theta_i, \theta_h, s_i^+, s_h) - E(\Delta^j_{ih}|\theta_i, \theta_h, s_i, s_h)$ identifies the disappointment effect.*

In the following Section, I present two empirical designs to identify disappointment due to a change in users' expectations. Following the same approach used in the theoretical framework, I start studying variations in ratings before and after nominations.

I show supporting evidence to document the change in moviegoers' characteristics before and after nominations. To account for selection, I use not nominated movies as control group. Doing so, I assume not nominated movies to not be treated by the nomination shock; and I support this assumption by looking at their rating variations after nomination. This is in line with the theoretical result in Observation 2 where the signal of the not nominated movie h remains constant to s_h.

First, I develop a DiD strategy that cannot fully disentangle disappointment from selection. Then, I present a cluster analysis technique that allows to match similar movies in order to replicate the condition $\boldsymbol{\mu}_i = \boldsymbol{\mu}_h = \bar{\boldsymbol{\mu}}_{ih}$. Finally, I present a new design embedding pre-treatment matching and difference-in-difference to isolate the disappointment effect as it is reported in Equation 3.5.

4 Identification Strategy

In this Section, I propose two empirical strategies. In both designs, I assume that movies nominated for the AMPAS awards receive an upward shift in users' expectations about their quality. With the first strategy, I measure changes in ratings for nominated movies after the nomination controlling for the variation of ratings occurring to not nominated movies. In this way, following Observation 1, it is not possible to completely disentangle disappointment from selection, apart from controlling for observable characteristics of users and movies. Conversely, with the second strategy, I study variations in the differences of ratings reported by the same user for movies with similar features. Doing that, I remove the bias due to the selection effect following the assumptions presented in Section 3.

Users' expectations may be affected by a variety of marketing activities. Thus, in both strategies, I focus on a limited window of days around the AMPAS nominations.[10] In this way, I try to isolate the impact of AMPAS nomination events on users' expectations. As a direct consequence of this focus, I can only identify the short-run effect of an increase in expectations.

4.1 Difference-in-Difference

In this empirical design, I use movie ratings as a response variable. With a Difference-in-Difference (DiD) strategy, I study variations in ratings before and after the AMPAS nomination dates for nominated movies, controlling for the variation occurred to not nominated movies. Accordingly, the control group (not nominated movies) provides the counterfactual dynamics of ratings for nominated movies

[10] In the main text I always consider a window of 30 days before and after the nomination; in Appendix D, I repeat the main analysis considering a 60-day window before and after the nomination.

in case they did not receive a nomination. The main equation is:

$$r_{ij} = \alpha + \theta_i + v_j + \beta_1 Nom_i + \beta_2 T_{ij} + \beta_3 Nom_i \times T_{ij} + \boldsymbol{\delta X}_{ij} + \varepsilon_{ij}, \tag{4.1}$$

where r_{ij} is the rating for movie i by user j. θ_i and v_j are movie i and user j fixed effects, respectively. Nom_i takes value 1 if movie i is nominated, and 0 otherwise; T_{ij} takes value 1 if rating r_{ij} is displayed after the nomination date associated with movie i, and 0 otherwise. $\boldsymbol{X}_{ij}$ is a set of control variables regarding movie's and user's characteristics. I divide this set into two groups: 1) time-invariant movie variables (identified only when I do not use movie fixed effect): $US-Release_i$, $US-Production_i$ and $English-Language_i$ are three dummy variables taking value 1 if movie i is first released in US; if movie i is produced in US; or, if movie i's main language is English, respectively. $Director-Nominated_i$, $Director-Awarded_i$, $Stars-Nominated_i$, and $Stars-Awarded_i$ are four dummy variables taking value 1 if movie i's director has ever been nominated for AMPAS awards, or has ever won the AMPAS awards; and, if movie i's three main stars have ever been nominated for AMPAS awards, or have ever won the AMPAS awards, respectively.[11] 2) time-variant variables: $diff_{ij}$, the distance between the day of rating r_{ij} and the nomination day of movie i; $\bar{r}_{jt-1}$, the average rating by user j before rating movie i; n_{jt-1}, the number of posted ratings by user j before rating movie i. Controlling for these variables may be relevant to capture variations in users' features, and thus to isolate the disappointment component from selection.

Not nominated movies and nominated movies differ in ratings before nomination (see Tables 1 and 2 and Appendix Figure A.3). Accordingly, using movie fixed effects or controlling for movie characteristics (in $\boldsymbol{X}_{ij}$) is key in order to compare relatively similar movies and replicate the condition of a proper "random assignment" of the shock (in this case the AMPAS nominations). Yet, looking at ratings before nominations, the dynamics of the treated (nominated) and the control (not nominated) groups are similar. To document these trends, I perform two event-studies separately for nominated and not nominated movies. I regress ratings r_{ij} over a full set of dummy variables for each group of five days around nominations from 60 days before to 30 days after:

$$r_{ij} = \alpha + \sum_{t=-60}^{30} \delta_\tau \mathbb{1}(\tau = t) + \varepsilon_{ij}. \tag{4.2}$$

Appendix Figures C.1 and C.2 show the estimated coefficients δ_τ for nominated and not nominated movies, respectively. The coefficients show a parallel, slightly negative trend before nominations for both groups (confirming the evidence of Table 2). Moreover, not nominated movies do not seem to receive a positive shift in their ratings after nomination (at least not of the same magnitude relative to nominated movies). This is in line with the previous assumption according to which not nominated

[11] The identities of the director and of the three main stars of movies are provided by the scraped data from IMDb.

movies are not treated. If this was not the case, the parameter β_3 would capture users' disappointment for nominated movies, together with users' surprise or elation for not nominated movies.[12]

I conclude this analysis with a third event study design to check the parallel trends between treated and control groups. Extending the identification presented in Equation 4.1, I substitute the indicator T_{ij} with multiple dummy variables for each group of five days around nominations:

$$r_{ij} = \alpha + \beta_1 Nom_i + \sum_{t=-60}^{30} \delta_\tau \mathbb{1}(\tau = t) + \sum_{t=-60}^{30} \beta_\tau Nom_i \times \mathbb{1}(\tau = t) + \varepsilon_{ij} \tag{4.3}$$

Figure 2 reports the coefficients β_τ associated with the combinations between the variable Nom_i and the time dummy variables in Equation 4.3. No trend in the period before the nomination can be detected. This suggests again that nominated and not nominated movies follow comparable paths at least before the shock. In Appendix C, I present two additional event-study graphs to show evidence in favor of the parallel-trend assumption. Appendix Figure C.3 shows the coefficients β_τ once I add all controls and fixed effects to Equation 4.3. In Appendix Figure C.4, I repeat the same analysis of Figure 2 for a longer span of days after the nomination. Finally, I provide additional evidence in favor of the parallel trend assumption supporting the DiD design using a placebo test. In particular, I study the same specification expressed by Equation 4.1 shifting back the window of time by 30 days. In this way, I compare the dynamics of nominated and not nominated movies starting 60 days before the nomination and imposing a placebo shock 30 days before the actual day of AMPAS nominations. To do that, I introduce a new indicator variable T_{ij}^{-30} taking value 1 if rating r_{ij} is displayed in the 30 days before the nomination date associated with movie i, and 0 otherwise (if r_{ij} is displayed between 60 and 30 days before the nomination). Appendix Table C.1 shows the results of the placebo test. I use different control and fixed effects, and, in all specifications, the parameters for $Nom_i \times T_{ij}^{-30}$ are always much smaller and never significant at five percent level.

This DiD design provides a credible way to measure the change in ratings for nominated movies after the AMPAS nominations. If moviegoers were randomly assigned to movies, the coefficient β_3 in Equation 4.1 could identify the disappointment generated by the quality disclosure. Yet, moviegoers are not randomly matched with movies. Using not nominated movies as a control group captures selection related to seasonality (moviegoers during the Winter holidays are different from those going to theaters in March).

[12]Using the notation from the previous theoretical framework, this could be equivalent to compare the dynamics of ratings for a movie i that receives a change in signal from s_i to s_i^+ with $E(\theta_i|s_i) < E(\theta_i|s_i^+)$ (the movie is nominated), with a movie h receiving a change in signal from s_h to s_h^- with $E(\theta_h|s_h) > E(\theta_h|s_h^-)$ (the movie is not nominated). Not being nominated does not seem to positively affect ratings after nomination, and we may conclude that the impact on expectations of signal s_h^- is relativel small: $E(\theta_h|s_h) \sim E(\theta_h|s_h^-)$. Yet, even with $E(\theta_h|s_h) > E(\theta_h|s_h^-)$, the parameter β_3 can still be interpreted as the empirical estimate of the disappointment/elation coefficient γ.

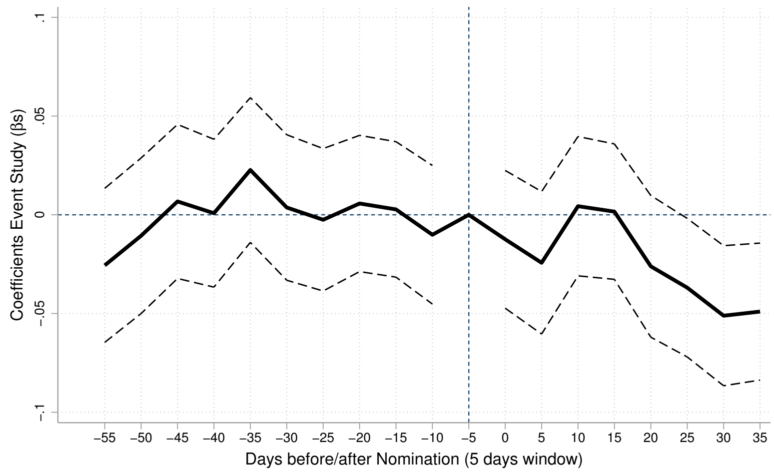


Figure 2: The Event Study Graph around the AMPAS Nominations

Note: The figure shows the dynamics of the parameters β_τ of Equation 4.3 around the AMPAS Nominations starting 60 days before until 30 days after. 95% confidence intervals with robust s.e. are displayed.

Moreover, selection on observation may be invoked after controlling for movie and user fixed effects: these variables absorb all fixed characteristics of users and movies that could affect ratings. To capture the match values between users and movies, I further control for the average ratings and the number of reviews of a user before rating a movie.

However, this strategy does not consider all available information regarding the movie/user matches before and after nominations. With the following design, I consider these features of the dataset connecting the theoretical framework in Section 3 with the conditional independence assumption necessary to account for the selection effect.

4.2 Difference-in-Difference with Movie Matching

The response variable for this second identification strategy is the difference in ratings reported by the same user. In particular, for each nominated movie i, I construct the difference in ratings between i and each not nominated movie h rated by the same user j: $\Delta_{ih}^{j} = r_{ij} - r_{hj}$.[13]

Studying variations of Δ_{ih}^{j} over time (over user j) for the same couple of movies absorbs all fixed characteristics of moviegoers (ν_j) and the differential of movies' quality (θ_i and θ_h). To this extent, the approach is similar to controlling for movie and user fixed effects in Equation 4.1. Yet, the time-variant quality of matches between users and movies is still not captured and selection in terms of match-value may drive the results: moviegoers with no interest in a specific genre could watch and

[13]As before, only ratings displayed in a short window of days around the nomination are used. Moreover, only users who watch at least one nominated movie during the period of analysis are considered.

rate a movie after its nomination, attracted by the expected high quality.

To address this selection channel, I match nominated movies with not nominated movies (rated by the same users) sharing similar features, in line with Observation 2 in Section 3. To do so, I use information about movies genres and users' profile who watched the movies before nomination. I perform a cluster analysis using two unsupervised learning algorithms described in the next Subsection.[14] These algorithms aim at finding patterns in datasets and select categories. Accordingly, with a non-parametric matching among movies, I try to select nominated and not nominated movies sharing similar characteristics. Restricting on couples of movies in the same cluster is equivalent to removing (or at least reducing) the selection effect due to changes in matches' quality.

4.2.1 Cluster Analysis: k-mean and k-median Algorithms

Cluster analysis is an exploratory data-analysis technique used to find patterns and select categories for multidimensional datasets. In the last decades, several algorithms and methods have been proposed to group data. Here I use two different algorithms that have been successful in finding clusters in multiple contexts: k-mean and k-median algorithms.

I apply these clustering techniques to group movies in the following ways. First, I group movies for different genres and I perform the k-mean and the k-median clustering algorithms for each genre separately. I do this since genres appear to be an extremely relevant movie's feature and I want to separate movies with different features. Then, for each user j, I consider the proportion of movies rated by j (before nominations) belonging to a specific genre g over the total number of movies rated by j (before nominations). These ratios p_j^g are related to the preferences of user j in terms of genres. If user j only watches and rates "action" movies, then $p_j^{Action} = 1$, with all the other parameters equal to zero. After having computed p_j^g for each user and each genre, I derive for each movie i the average proportion of ratings $\bar{p}_i^g$ using all p_j^g of users who rated movie i. Accordingly, different values of $\bar{p}_i^g$ reveal which types of users rate movie i and a movie is characterized by twenty variables since the dataset divides movies into twenty genres.

A possible approach could be to use Δ_{ih}^j only for movies with the same values for all twenty $\bar{p}_i^g$. However, this would dramatically shrink the number of observations (with a similar curse of dimensionality occurring with exact matching techniques (Abadie and Imbens, 2006). Therefore, I use k-mean and the k-median clustering to select different clusters of movies (that share the same genre) using the values of $\bar{p}_i^g$ for each movie. This approach shares a similar intuition with the nearest neighbor covariate matching techniques: here clustering algorithms use different notions of distance to select movies with similar characteristics. Moreover, the main assumption required to identify

[14] MovieLens datasets have been often used as a laboratory to test recent techniques measuring similarity among movies. For a list of studies regarding cluster analyses performed on the MovieLens datasets see the webpage: https://grouplens.org/publications/.

disappointment regards the power of this technique to select movies with features such that moviegoers matching value is equivalent for the two movies: $d(\boldsymbol{\pi}_j, \boldsymbol{\mu}_i) = d(\boldsymbol{\pi}_j, \boldsymbol{\mu}_h)$. This assumption is again analogous to matching techniques: we require the differential in distances $d(\boldsymbol{\pi}_j, \boldsymbol{\mu}_i) - d(\boldsymbol{\pi}_j, \boldsymbol{\mu}_h)$ being as-good-as random after selecting movies in the same clusters.

To have a sense of the clustering results, Appendix Figures C.5 and C.6 show scatter plots that display "drama" and "comedy" movies (the two more frequent genres) over the variables $\bar{p}_i^{Drama}$ and $\bar{p}_i^{Comedy}$. In both cases, the k-mean algorithm splits the dataset into two clusters with comparable results. Movies are grouped depending on the high or low values of $\bar{p}_i^{Drama}$ and $\bar{p}_i^{Comedy}$. Appendix Figure C.7 shows the same scatter plots for the eight most present genres. In the majority of the cases, two clusters are chosen.

4.2.2 Difference-in-Difference with Movie Matching: Regression

After having grouped movies in different clusters, it is possible to analyze the dynamics of Δ_{ih}^j around the AMPAS nomination dates in order to identify the disappointment effect. In the following regression, once I restrict on movies h sharing the same genre and clusters with movie i, the coefficient β multiplying T_{ij} represents the impact of disappointment:

$$\Delta_{ih}^j = \alpha + \lambda_{ih} + \beta T_{ij} + \boldsymbol{\delta} \boldsymbol{X}_{ih}^j + \varepsilon_{ih}^j. \tag{4.4}$$

Here, λ_{ih} represents the movie i/movie h combination fixed effect; T_{ij} is, as before, an indicator variable taking value 1 if r_{ij}, the rating for the nominated movie i, is displayed after the nomination date, and 0 otherwise. Finally, $\boldsymbol{X}_{ih}^j$ is a set of the following control variables for movie i and h varying over time (all fixed characteristics are not identified by the presence of the movie-combination fixed effect) : $\bar{r}_{jt-1(h)}$ and $\bar{r}_{jt-1(i)}$ are the average ratings by user j before rating movie h and i, respectively; $\bar{r}_{ht-i(j)}$ and $\bar{r}_{it-i(j)}$ are the average ratings for movie h and i before the rating by user j, respectively. $n_{jt-1(h)}$ and $n_{jt-1(i)}$ are the number of ratings by user j before rating movie h and i, respectively. Finally, $n_{ht-1(j)}$ and $n_{it-1(j)}$ are the number of ratings for movie h and i before the rating by user j, respectively.

This new identification design exploits the potential fixed effects enabled by the dataset; still, it continues to rely on the relationship between nominated and (a subsample of) not nominated movies to identify the parameter of interest. For this reason, it is relevant to provide evidence about the absence of pre-trends in the dynamics of Δ_{ih}^j before nominations. In particular, it is possible to extend the identification presented in Equation 4.4 with multiple dummy variables for each group of five days around the nominations as it is proposed for the DiD design (see Subsection 4.1):

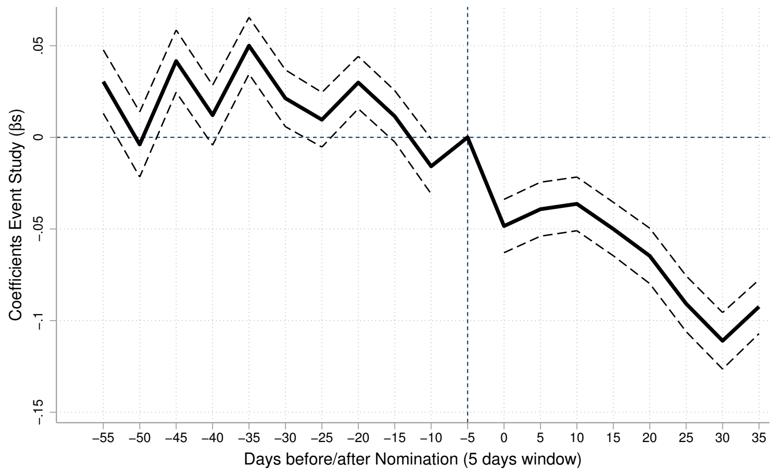


Figure 3: The Event-Study Graph for Δ^j_{ih} around the AMPAS Nominations

Note: The figure shows the dynamics of the parameters β_τ of Equation 4.5 around the AMPAS Nominations starting 60 days before until 30 days after. 95% confidence intervals with s.e. clustered at movie combination level are displayed.

$$\Delta^j_{ih} = \lambda_{ih} + \sum_{t=-60}^{30} \beta_\tau \mathbb{1}(\tau = t) + \varepsilon^j_{ih}. \tag{4.5}$$

Figure 3 reports the coefficients β_τ associated with the combinations between the dummy variables in Equation 4.5 when only movies with the same genre and belonging to the same k-mean and k-median clusters are compared. No trends can be detected in the period before the nomination. Still, a slight anticipation effect can be observed.

A possible interpretation for these dynamics can be given by dividing the time window around nominations into three periods. In the first period (until the window $-15/-10$), users are not aware of the identity of potential candidates for the nominations. Thus, they do not get disappointed by movies. Then, after window $-15/-10$, the identity of the potential nominated movies starts to become clear and the disappointment effect starts to affect ratings. Crucially, this is the period in which the Golden Globes, the second most important awards for movies and TV shows, are assigned. The correlation between movies nominated and awarded with the Golden Globe and nominated movies for the AMPAS awards is very high and this could ignite a shift in users' expectations. Finally, after window $10/15$, users' ratings are affected by the official nominations with a further drop.

Comparing this graph with the previous ones in Figure 2, it is possible to notice that the drop in ratings here is much more significant and of greater magnitude. Therefore, assuming that the new design is able to isolate disappointment, it is possible to conclude that the selection effect positively affects ratings after nomination. In Section 3, I present a framework in which this is possible and it may be due to a change in the distribution of the users' utility parameters before and after nominations.

In Appendix C, I present two additional event-study graphs to show evidence in favor of the parallel-trend assumption. Appendix Figure C.8 shows the coefficients β_τ once I add all controls and fixed effects to Equation 4.3. In Appendix Figure C.9, I repeat the same analysis of Figure 2 for a longer span of days after the nomination.

5 Results

In this Section, I show the results for the two identification strategies described before. Following the structure in Section 4, first I describe the results for the DiD design; then, I move to the results of the second strategy adding the movie matching with the cluster algorithms to the DiD approach.

5.1 Difference-in-Difference

The DiD strategy can credibly show whether the AMPAS nominations affect movie ratings, but it is silent regarding the source of this change. Table 3 shows the main results presenting nine different specifications and restricting on a window of 30 days before and after the nominations. A wider window (60 days before and after) is reported in Appendix Table D.1.

Different fixed effects are added in Columns 1-5; whereas, in Column 6 I add the following time-invariant controls: $US-Release_i$, $US-Production_i$ and $English-Language_i$; $Director-Nominated_i$, $Director-Awarded_i$, $Stars-Nominated_i$, and $Stars-Awarded_i$. In Columns 7-8, movie and user fixed effects are present and only time-varying controls can be identified.

The results show that the main parameter of interest (the interaction $Nom_i \times T_{ij}$) is negative, significant and stable across different specifications. Accordingly, after the AMPAS nominations, nominated movies get significantly lower ratings relative to not nominated ones.[15]

Although statistically significant, the effect is relatively small if we look at the average ratings for nominated movies (3.73 in Table 1). Yet, once this effect is compared with the rating premium for nominated movies (the coefficient for the parameter Nom_i in Table 3), its magnitude and economic significance look greater: the drop in ratings for nominated movies in the first 30 days after the AMPAS nomination accounts for five percent of the premium of being a nominated movie. Similar results are present with a wider window of time around the nominations (60 days), as it is shown in Appendix Table D.1.

[15]To account for users' selection, I have also studied specifications with an extensive set of variables related to users' histories such as the proportion of watched movies belonging to specific genres, and users' average ratings regarding movies belonging to specific genres. The negative and significant impact of the AMPAS nominations on nominated movies' ratings is robust to all specifications (available upon request).

Table 3: Difference-in-Difference: Regressing r_{ij} as in Equation 4.1

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Nom_i	0.613***	0.573***	0.596***	0.596***	0.571***	0.394***	0	0
	(0.0239)	(0.0255)	(0.0259)	(0.0256)	(0.0263)	(0.0262)	(.)	(.)
T_{ij}	-0.00361	-0.00256	-0.000249	-0.000271	-0.0294***	-0.0271***	-0.0347***	0.00900
	(0.00492)	(0.00484)	(0.00469)	(0.00471)	(0.00661)	(0.00650)	(0.00485)	(0.00825)
$Nom_i \times T_{ij}$	-0.0225**	-0.0225**	-0.0274***	-0.0261***	-0.0282***	-0.0282***	-0.0170**	-0.0180**
	(0.0103)	(0.0101)	(0.0101)	(0.00978)	(0.0109)	(0.0105)	(0.00819)	(0.00818)
$diff_{ij} \times 100$								-0.141***
								(0.0235)
$\bar{r}_{jt-1}$								0.00690
								(0.0140)
$n_{jt-1} \times 1000$								-0.103***
								(0.00916)
Constant	3.371***	3.379***	3.373***	3.373***	3.389***	3.508***	3.508***	3.501***
	(0.0120)	(0.0115)	(0.0113)	(0.0111)	(0.0105)	(0.0303)	(0.00201)	(0.0505)
$\boldsymbol{X}_{ij}$						✓		
Genre FE		✓	✓	✓	✓	✓		
Year(Award) FE			✓	✓	✓	✓		
Clusters FE				✓	✓	✓		
User FE					✓	✓	✓	✓
Movie FE							✓	✓
R^2	0.0544	0.0622	0.0647	0.0649	0.289	0.299	0.418	0.419
Observations	390,213	390,212	390,212	389,856	385,842	385,842	384,894	384,159

Note: In Column 1 I present a specification with no controls or fixed effects. From Column 2 to 5 I add fixed effects referring to the genre (Genre FE), the year of the award (Year (Award) FE), the k-mean and k-median clusters (Clusters FE), and users (User FE) respectively. In Column 6, I add all time-invariant controls grouped in $\boldsymbol{X}_{ij}$. In Columns 7 and 8, I add user and movie fixed effects (Movie FE) together with time-invariant controls. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. I restrict on a time window of 30 days before and after the Academy Nominations. Cluster standard errors at movie level are in parentheses. After adding movie fixed effects, the parameter regarding the variable Nom_i (and all time-invariant controls) cannot be identified due to multicollinearity with the fixed effects.

Table 4: Difference-in-Difference: Regressing rating r_{ij} as in Equation 4.1 for Different Nominated Movies

	(1)	(2)	(3)
$Nom_i^1 \times T_{ij}$	-0.00711		
	(0.0118)		
$Nom_i^{>1} \times T_{ij}$	-0.0252**		
	(0.0102)		
$Nom_i^{award} \times T_{ij}$		-0.0105	-0.0283**
		(0.0119)	(0.0115)
$diff_{ij} \times 100$	-0.141***	-0.141***	-0.101***
	(0.0236)	(0.0236)	(0.0121)
$\bar{r}_{jt-1}$	0.00693	0.00693	0.0200
	(0.0140)	(0.0140)	(0.0128)
$n_{jt-1} \times 1000$	-0.103***	-0.102***	-0.105***
	(0.00915)	(0.00915)	(0.00794)
Constant	3.501***	3.501***	3.483***
	(0.0505)	(0.0505)	(0.0463)
User FE	✓	✓	✓
Movie FE	✓	✓	✓
R^2	0.419	0.419	0.409
Observations	384,159	384,159	575,650

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. In Columns 1 and 2 I restrict on a time window of 30 days before and after the Academy Nominations. In Column 3, I add further 30 days after the nomination to capture the impact of the awards. Cluster standard errors at movie level are in parentheses. All three specifications report the full set of controls and fixed effects as in Column 8 in Table 3. Column 1 shows results for nominated movies with one nomination only, and nominated movies with more than one nomination. Columns 2 and 3 show results for nominated movies that win at least one AMPAS award.

To study the heterogeneity of the nomination effect, I report in Table 4 three different specifications. In Column 1, I divide nominated movies in two groups: movies that are nominated only for one award (Nom_i^1); and movies that are nominated for more than one award ($Nom_i^{>1}$). Thus, I repeat the same analysis as in Equation 4.1 substituting the dummy variable Nom_i with a variable taking three values: not nominated; nominated for one award, and nominated for more than one award.

The coefficient of interest for movies nominated only for one award $Nom_i^1 \times T_{ij}$ is statistically insignificant, but still negative. Conversely, the coefficient of interest for movies nominated for more than one award $Nom_i^{>1} \times T_{ij}$ is negative and statistically significant. Moreover, comparing the magnitude of this coefficient with the one obtained in Column 8 in Table 3 (with the same specification), it is possible to observe that the coefficient $Nom_i^{>1} \times T_{ij}$ is greater than the one about all nominated movies. In Columns 2 and 3, I restrict the attention to nominated movies receiving at least one AMPAS award. The award ceremonies usually occur around 30-35 days after nominations. Once I restrict my analysis to the first 30 days after nomination (when awards are not yet assigned), the coefficient $Nom_i^{award} \times T_{ij}$ is negative, but not significant (Column 2). In Column 3, I extend the study to 60 days after nominations (so roughly 30 days after awards). Doing so, the parameter turns significant with greater negative magnitude.

The results in Tables 3 and 4 are in line with the empirical predictions proposed in Section 3. In particular, Table 3 shows that the increase in expectations due to the AMPAS nominations leads to a significant decrease in ratings. Furthermore, the heterogeneous effects reported in Table 4 show that those movies whose expectations receive a greater positive shock (those with a higher number of nominations) experience a more negative drop in ratings after the nominations. In the next Section, I provide further evidence in line with these two observations; and I show that the disappointment effect is the main driving force for the resulting drop in ratings.

5.2 Difference-in-Difference with Movie Matching

With the second design, I can identify the role of disappointment in the drop of movie ratings due to the AMPAS nominations. To do so, I study variations in the difference Δ_{ih}^{j} before and after the nomination as in Equation 4.4. Table 5 shows the main results presenting seven different specifications and restricting on a window of 30 days before and after nominations. A wider window (60 days before and after) is used in Appendix Table D.2.

Columns 1-4 report the specification without adding controls apart from the movie-combination fixed effects. Column 1 does not restrict on movies with similar characteristics and the coefficient of interest may capture disappointment and selection due to the change in the matching between movies and users. Columns 2-4 show results varying the type of not nominated movies compared with the nominated ones: in Column 2, I consider movies with the same genre; in Columns 3 and 4, with the same genre and belonging to the same clusters (using the k-mean and k-median algorithms). Accordingly, these estimates should further remove selection and identify the disappointment effect alone. In the remaining columns, I report results adding further controls for movies and users characteristics: $\bar{r}_{jt-1(h)}$, $\bar{r}_{jt-1(i)}$, $\bar{r}_{ht-i(j)}$, $\bar{r}_{it-i(j)}$, $n_{jt-1(h)}$, $n_{jt-1(i)}$, $n_{ht-1(j)}$, and $n_{it-1(j)}$. All other controls used in the previous design cannot be considered here since all movie-specific variables vanish by the presence of the movie-combination fixed effects. All specifications show negative and significant results for the coefficient of interest. Therefore, disappointment is the major driving force in the drop of ratings after nominations: when the selection channel is reduced (if not totally removed), the coefficients keep being negative and significant. Comparing these results with the ones obtained with the DiD design, it is possible to claim that selection seems to shift upward ratings after nominations.

Table 5: Difference-in-Difference with Movie Matching: Regressing Δ_{ih}^{j} as in Equation 4.4

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
T_{ij}	-0.0595***	-0.0580***	-0.0609***	-0.0612***	-0.0627***	-0.0618***	-0.0332***
	(0.00127)	(0.00278)	(0.00316)	(0.00328)	(0.00330)	(0.00330)	(0.00434)
$\bar{r}_{ht-1(j)}$					0.192***	0.183***	0.204***
					(0.0149)	(0.0149)	(0.0149)
$\bar{r}_{jt-1(h)}$					-0.334***	-0.341***	-0.357***
					(0.0227)	(0.0226)	(0.0228)
$\bar{r}_{it-1(j)}$						-0.200***	-0.176***
						(0.0127)	(0.0129)
$\bar{r}_{jt-1(i)}$						0.119***	0.118***
						(0.0128)	(0.0131)
$n_{jt-1(h)} \times 1000$							0.0500***
							(0.00449)
$n_{jt-1(i)} \times 1000$							0.201***
							(0.0131)
$n_{ht-1(j)} \times 1000$							-0.148***
							(0.0138)
$n_{it-1(j)} \times 1000$							-0.00971**
							(0.00435)
Constant	0.607***	0.581***	0.558***	0.552***	1.211***	1.549***	1.413***
	(0.000621)	(0.00135)	(0.00152)	(0.00158)	(0.103)	(0.105)	(0.107)
Movie-Combination FE	✓	✓	✓	✓	✓	✓	✓
All Movies	✓						
Movies with Same Genre		✓	✓	✓	✓	✓	✓
Movies in Same Cluster (k-mean)			✓	✓	✓	✓	✓
Movies in Same Cluster (k-median)				✓	✓	✓	✓
R^2	0.291	0.283	0.282	0.289	0.289	0.290	0.292
Observations	3,764,044	765,015	556,088	505,489	501,069	500,636	500,636

Note: In Column 1 I present a specification with movie-combination fixed effect, but without restricting to a subsample of movies with similar features. From Column 2 to 4 I restrict to movies with the same genre, and movies belonging to the same genre and and the same k-mean and k-median clusters, respectively. In Columns 6 and 7, I add time-invariant controls to the specification in Column 5. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. I restrict on a time window of 30 days before and after the Academy Nominations. Cluster standard errors at movie combination level are in parentheses.

Table 6: Difference-in-Difference with Movie Matching: Regressing Δ^{j}_{ih} as in Equation 4.4 for Different Nominated Movies

	(Nom^{1}_{i}) (1)	$(Nom^{>1}_{i})$ (2)	(Nom^{award}_{i}) (3)	(Nom^{award}_{i}) (4)
T_{ij}	-0.0101	-0.0406***	-0.0340***	-0.0440***
	(0.00736)	(0.00561)	(0.0120)	(0.00992)
$\bar{r}_{ht-1(j)}$	0.176***	0.221***	0.216***	0.159***
	(0.0240)	(0.0191)	(0.0414)	(0.0326)
$\bar{r}_{jt-1(h)}$	-0.440***	-0.280***	-0.496***	0.0693**
	(0.0338)	(0.0307)	(0.0881)	(0.0335)
$\bar{r}_{it-1(j)}$	-0.141***	-0.197***	-0.259***	-0.219***
	(0.0221)	(0.0159)	(0.0298)	(0.0229)
$\bar{r}_{jt-1(i)}$	0.0925***	0.132***	0.116***	0.114***
	(0.0220)	(0.0163)	(0.0301)	(0.0235)
$n_{jt-1(h)} \times 1000$	0.0646***	0.0428***	0.0141	0.0644***
	(0.00699)	(0.00588)	(0.0147)	(0.00973)
$n_{jt-1(i)} \times 1000$	0.181***	0.212***	0.220***	0.209***
	(0.0213)	(0.0167)	(0.0307)	(0.0248)
$n_{ht-1(j)} \times 1000$	-0.323***	-0.110***	-0.170***	-0.116***
	(0.0381)	(0.0154)	(0.0262)	(0.0146)
$n_{it-1(j)} \times 1000$	-0.0370***	0.00358	0.0475***	-0.00823
	(0.00714)	(0.00552)	(0.0149)	(0.00966)
Constant	1.669***	1.164***	2.619***	0.297
	(0.159)	(0.144)	(0.400)	(0.183)
Movie-Combination FE	✓	✓	✓	✓
Movies with Same Genre	✓	✓	✓	✓
Movies in Same Cluster (k-mean)	✓	✓	✓	✓
Movies in Same Cluster (k-median)	✓	✓	✓	✓
R^2	0.295	0.277	0.258	0.238
Observations	193,332	307,304	66,264	115,777

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. In Columns 1, 2, and 3 I restrict on a time window of 30 days before and after the Academy Nominations. In Column 4, I add further 30 days after the nomination to capture the impact of the awards. Cluster standard errors at movie combination level are in parentheses. All four specifications report the full set of controls and fixed effects as in Column 7 in Table 5. Column 1 show results for nominated movies with one nomination only. Column 2 show results considering nominated movies with more than one nomination. Columns 3 and 4 show results for nominated movies that win at least one AMPAS award.

Similar results are present with a wider window of time around the nominations (60 days), as it is shown in Appendix Table D.2. Moreover, as it is shown in Figure 3, the disappointment effect is characterized by a short anticipation since the drop in ratings starts five-ten days in advance relative to the nomination dates. This may be due to the release of relevant information about the potential candidates for the AMPAS awards a few days before the actual nominations.[16] Furthermore, the lag between the effect observed in Figure 3 and the one in Figure 2 about the DiD design suggests again that the selection effect should not significantly drive the drop in ratings and it may actually have a positive effect. A further difference between the results of the two designs regards the magnitude of the estimated effect. In the previous Subsection, the coefficient of interest results to be negative and significant. The short-run drop in ratings due to nominations is, on average, 0.02 stars. As reported in the previous Subsection, it accounts for almost five percent of the rating premium for nominated movies. With this second strategy, the disappointment effect driven by nominations almost triplicates, reaching 0.06 stars, on average. This implies that, in the first 30 days after nominations, the rating premium for nominated movies may drop by more than ten percent because of disappointment. This result may be of importance for movie studios (and in general, for sellers with a product portfolio) producing movies with different qualities and likelihoods to receive a quality certification such as the AMPAS awards.

I conclude this Section commenting on the results about different types of nominated movies shown in Table 6. Similarly to the analysis of the previous design, Column 1 and 2 report the results of the regression in Equation 4.5 selecting nominated movies that received only one, and more than one nominations, respectively. Differently, I select movies that receive the AMPAS awards in Column 3 and 4. In Column 3 I consider only 30 days before and after nominations - when movies are not yet awarded; whereas in Column 4 I study 60 days after nominations. The results of this heterogeneity analysis are similar to the ones reported for the DiD design. Yet, here the difference in coefficients is less significant with respect to the variation reported in Table 4.[17] However, in line with the previous results, the magnitude of the variation of Δ^{j}_{ih} is greater (more negative) for those movies that are nominated for more than one award. Accordingly, it is possible to claim that, when the increase in expectations is greater, the greater is the drop in ratings. Finally, as in the previous design, ratings for awarded movies seem to drop after nomination but only considering a longer span after the nomination dates. In Appendix E , I elaborate on this last result and I propose a new methodology to identify the disappointment effect associated with the AMPAS awards. Here I use nominated movies that are not awarded as a control group for nominated movies receiving an award. Both groups received the previous shocks on expectations by the nominations, but only those receiving at least one award are treated by a new shock that trigger further disappointment. The estimates show that also the AMPAS awards negatively impact ratings for awarded movies in line with the theoretical framework.

[16]As supporting evidence for this phenomenon, it is worth recalling that there is a flourishing betting activity over AMPAS awards and nominations.

[17]The not nominated movies used to compute the difference Δ^{j}_{ih} differ across columns making the comparison more difficult.

6 Conclusion

This paper shows empirical evidence supporting the presence of reference points in agents' utility. In particular, I identify the disappointment effect in users' utility related to the increase in expectations due to a quality disclosure. I analyze movie ratings displayed by users on an online recommender system (MovieLens). The Academy of Motion Picture Arts and Sciences awards' nomination is the main quality disclosure event that shifts (upward) users' expectations about movies. The results show that ratings for nominated movies drop after nominations, and disappointment due to the rise of expectations is the main driver for such a drop.

I propose two identification strategies: first, I use a DiD design to show that nominations have a negative and significant impact on ratings of nominated movies. To do that, I restrict my analysis over a short window of time around the AMPAS nominations and I compare the variations in ratings for nominated and not nominated movies before and after the nominations. This strategy relies on a parallel trend assumption regarding the evolution of ratings for the two groups of movies (nominated and not nominated). I provide evidence in favor of this assumption by studying rating variations for both groups before the shift in expectations. Results show that nominated movies' ratings experience a drop right after nomination. The effect is equivalent to five percent of the nomination premium regarding ratings to not nominated movies. This drop may be due to the disappointment faced by users after an increase in expectations. Still, it may also originate from a change in the characteristics of users who watch and rate movies after the nominations. The second identification aims to disentangle disappointment from selection, studying the difference in ratings by the same user between nominated and not nominated movies. I use unsupervised learning techniques (k-mean and k-median algorithms) to cluster movies with similar characteristics. In this way, comparing movies in the same cluster, the users' selection effect is highly reduced, and variations in the rating difference before and after the nomination can capture the disappointment effect. This analysis confirms the drop in ratings of nominated movies after the nominations and suggests that users' disappointment is a relevant driver.

The implications of these results are relevant for the design of third-party quality disclosures in many contexts with informational asymmetries. Certifications and quality disclosures are effective tools to reduce buyers' uncertainty; still, the resulting shift in expectations may partially backfire creating disappointment side-effects in users. With this paper, I show evidence regarding the existence and relevance of disappointment effects after a quality disclosure announcement. This finding is of special interest for sellers of experience goods. They should be aware that quality disclosures such as awards may generate unrealistic expectations and foster disappointment on the buyers' side. Buyers' dissatisfaction may lead to lower ratings and reduced future profits.[18] Therefore the management of

[18] Although the relationship between movie ratings and box-office is unclear, online reputation has proven to affect sellers' profits in many contexts (Chevalier and Mayzlin, 2006, Anderson and Magruder, 2012).

buyers' expectations by sellers can be critical in order to maximize the positive effects of certification and ensure buyers' learning induced by ratings to converge to the "true" sellers' quality. New research avenues may be explored in terms of applications for reference-based preferences in different settings, and regarding the optimal implementation and timing of quality disclosures.

References

Abadie, Alberto, and Guido W Imbens (2006) 'Large sample properties of matching estimators for average treatment effects.' *Econometrica* 74(1), 235–267

Anderson, Michael, and Jeremy Magruder (2012) 'Learning from the crowd: Regression discontinuity estimates of the effects of an online review database.' *The Economic Journal* 122(563), 957–989

Backus, Matthew, Thomas Blake, Dimitriy V Masterov, and Steven Tadelis (2017) 'Expectation, disappointment, and exit: Reference point formation in a marketplace.' Technical Report, National Bureau of Economic Research

Barberis, Nicholas C (2013) 'Thirty years of prospect theory in economics: A review and assessment.' *Journal of Economic Perspectives* 27(1), 173–96

Bartling, Björn, Leif Brandes, and Daniel Schunk (2015) 'Expectations as reference points: Field evidence from professional soccer.' *Management Science* 61(11), 2646–2661

Bell, David E (1985) 'Disappointment in decision making under uncertainty.' *Operations Research* 33(1), 1–27

Bondi, Tommaso, and Ryan Stevens (2019) 'The good, the bad and the picky: Consumer heterogeneity and the reversal of movie ratings'

Camerer, Colin, Linda Babcock, George Loewenstein, and Richard Thaler (1997) 'Labor supply of new york city cabdrivers: One day at a time.' *The Quarterly Journal of Economics* 112(2), 407–441

Card, David, and Gordon B Dahl (2011) 'Family violence and football: The effect of unexpected emotional cues on violent behavior.' *The Quarterly Journal of Economics* 126(1), 103–143

Chen, Yan, F Maxwell Harper, Joseph Konstan, and Sherry Xin Li (2010) 'Social comparisons and contributions to online communities: A field experiment on movielens.' *American Economic Review* 100(4), 1358–98

Chevalier, Judith A, and Dina Mayzlin (2006) 'The effect of word of mouth on sales: Online book reviews.' *Journal of Marketing Research* 43(3), 345–354

Chezum, Brian, and Brad Wimmer (1997) ‘Roses or lemons: Adverse selection in the market for thoroughbred yearlings.’ *Review of Economics and Statistics* 79(3), 521–526

Crawford, Vincent P, and Juanjuan Meng (2011) ‘New york city cab drivers’ labor supply revisited: Reference-dependent preferences with rational-expectations targets for hours and income.’ *American Economic Review* 101(5), 1912–32

Dellarocas, Chrysanthos, and Charles A Wood (2008) ‘The sound of silence in online feedback: Estimating trading risks in the presence of reporting bias.’ *Management Science* 54(3), 460–476

Donihue, Michael R, Randy A Nelson, Donald M Waldman, and Calbraith Wheaton (2001) ‘What’s an oscar worth?’ *Economic Inquiry* 39(1), 1–6

Dranove, David, and Ginger Zhe Jin (2010) ‘Quality disclosure and certification: Theory and practice.’ *Journal of Economic Literature* 48(4), 935–63

Dranove, David, Daniel Kessler, Mark McClellan, and Mark Satterthwaite (2003) ‘Is more information better? the effects of “report cards” on health care rroviders.’ *Journal of Political Economy* 111(3), 555–588

Elfenbein, Daniel W, Raymond Fisman, and Brian McManus (2015) ‘Market structure, reputation, and the value of quality certification.’ *American Economic Journal: Microeconomics* 7(4), 83–108

Farronato, Chiara, Andrey Fradkin, Bradley Larsen, and Erik Brynjolfsson (2020) ‘Consumer protection in an online world: An analysis of occupational licensing.’ Technical Report, National Bureau of Economic Research

Forbes, Silke J, Mara Lederman, and Trevor Tombe (2015) ‘Quality disclosure programs and internal organizational practices: Evidence from airline flight delays.’ *American Economic Journal: Microeconomics* 7(2), 1–26

Harper, F Maxwell, and Joseph A Konstan (2015) ‘The movielens datasets: History and context.’ *ACM Transactions on Interactive Intelligent Systems (tiis)* 5(4), 1–19

Ho, Daniel E (2012) ‘Fudging the nudge: Information disclosure and restaurant grading.’ *Yale Law Journal* 122, 574

Hu, Nan, Paul A Pavlou, and Jie Jennifer Zhang (2017) ‘On self-selection biases in online product reviews.’ *MIS Quarterly* 41(2), 449–471

Hui, Xiang, Maryam Saeedi, Giancarlo Spagnolo, and Steven Tadelis (2018) ‘Certification, reputation and entry: An empirical analysis.’ Technical Report, National Bureau of Economic Research

Hui, Xiang, Maryam Saeedi, Zeqian Shen, and Neel Sundaresan (2016) 'Reputation and regulations: Evidence from ebay.' *Management Science* 62(12), 3604–3616

Jin, Ginger Zhe, and Phillip Leslie (2003) 'The effect of information on product quality: Evidence from restaurant hygiene grade cards.' *The Quarterly Journal of Economics* 118(2), 409–451

Kahneman, Daniel, and Amos Tversky (1979) 'Prospect theory: An analysis of decision under risk.' *Econometrica* 47(2), 363–391

Klein, Tobias J, Christian Lambertz, and Konrad O Stahl (2016) 'Market transparency, adverse selection, and moral hazard.' *Journal of Political Economy* 124(6), 1677–1713

Kőszegi, Botond, and Matthew Rabin (2006) 'A model of reference-dependent preferences.' *The Quarterly Journal of Economics* 121(4), 1133–1165

Lafky, Jonathan (2014) 'Why do people rate? theory and evidence on online ratings.' *Games and Economic Behavior* 87, 554–570

Lee, Young-Jin, Kartik Hosanagar, and Yong Tan (2015) 'Do i follow my friends or the crowd? information cascades in online movie ratings.' *Management Science* 61(9), 2241–2258

Li, Xinxin, and Lorin M Hitt (2010) 'Price effects in online product reviews: An analytical model and empirical analysis.' *MIS Quarterly* pp. 809–831

Loomes, Graham, and Robert Sugden (1982) 'Regret theory: An alternative theory of rational choice under uncertainty.' *The Economic Journal* 92(368), 805–824

Mayzlin, Dina, Yaniv Dover, and Judith Chevalier (2014) 'Promotional reviews: An empirical investigation of online review manipulation.' *American Economic Review* 104(8), 2421–55

McKenzie, Jordi (2012) 'The economics of movies: A literature survey.' *Journal of Economic Surveys* 26(1), 42–70

Moe, Wendy W, and David A Schweidel (2012) 'Online product opinions: Incidence, evaluation, and evolution.' *Marketing Science* 31(3), 372–386

Proserpio, Davide, and Georgios Zervas (2017) 'Online reputation management: Estimating the impact of management responses on consumer reviews.' *Marketing Science* 36(5), 645–665

Proserpio, Davide, Wendy Xu, and Georgios Zervas (2018) 'You get what you give: Theory and evidence of reciprocity in the sharing economy.' *Quantitative Marketing and Economics* 16(4), 371–407

Talwar, Arjun, Radu Jurca, and Boi Faltings (2007) ‘Understanding user behavior in online feedback reporting.’ In ‘Proceedings of the 8th ACM conference on Electronic commerce’ pp. 134–142

Online Appendix

"Quality Disclosures and Disappointment: Evidence from the Academy Awards"

A Appendix: Empirical Setting and Dataset

Figure A.1: Snapshot of MovieLens.org: Main Page

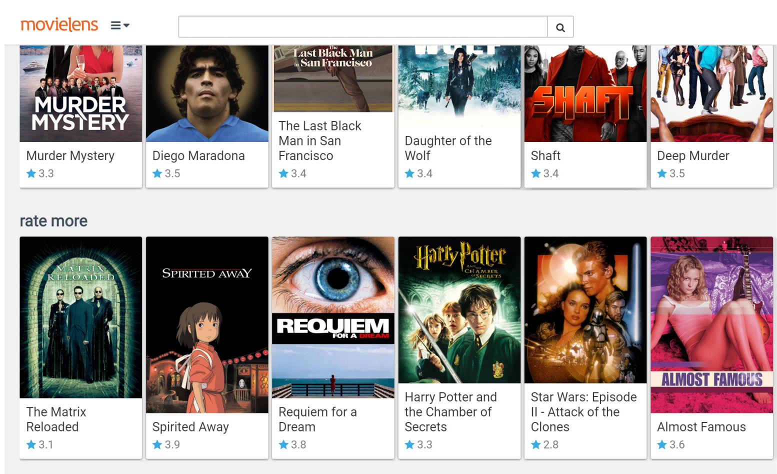
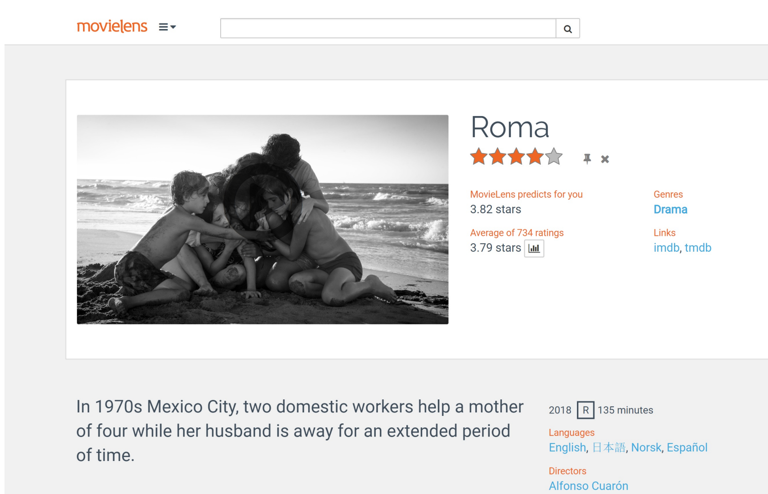


Figure A.2: Snapshot of MovieLens.org: Movie Page



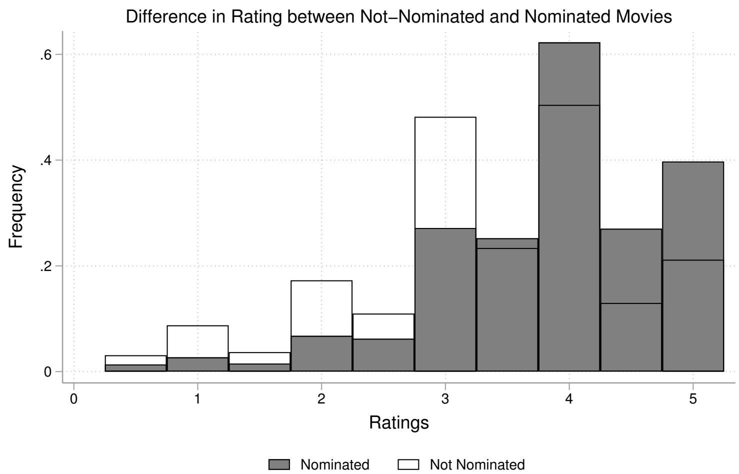


Figure A.3: Difference in Ratings between Not Nominated and Nominated Movies

Note: The figure shows the distributions of ratings for not nominated and nominated movies considering only pre-nomination ratings. Nominated movies (in gray) have higher ratings than not nominated movies.

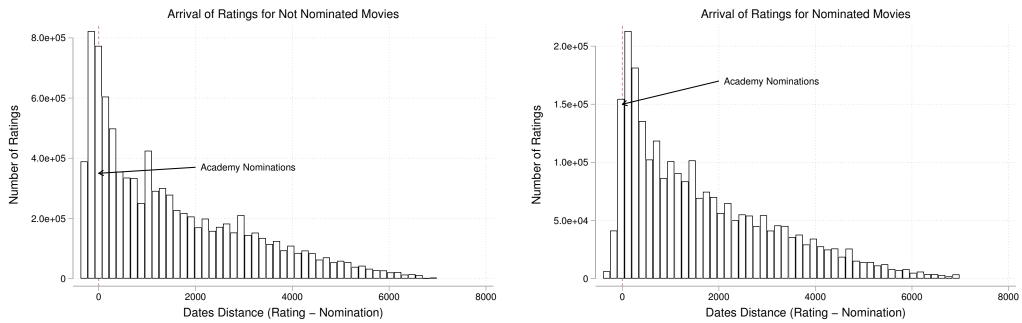


Figure A.4: Arrival of Ratings for Not Nominated and Nominated Movies (All)

Note: The two figures show the amount of ratings that are displayed over time for not nominated and nominated movies. On the x-axis, time is measured in terms of days of distance from the nomination dates.

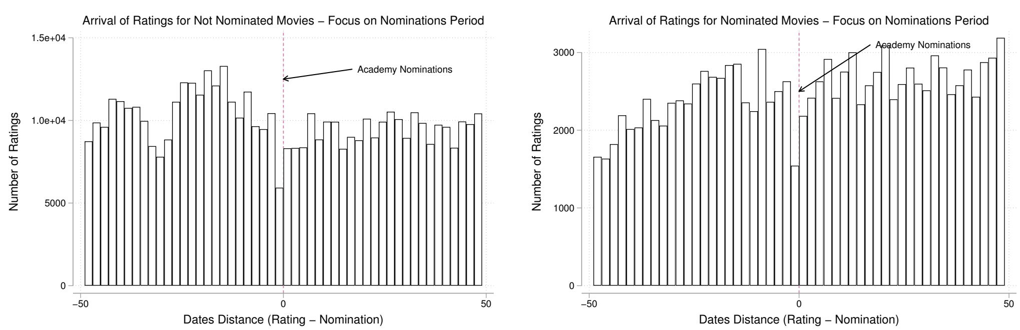


Figure A.5: Arrival of Ratings for Not Nominated and Nominated Movies (Nominations Period)

Note: The two figures show the amount of ratings that are displayed over time for not nominated and nominated movies. On the x-axis, time is measured in terms of days of distance from the nomination dates.

B Appendix: Theoretical Framework

B.1 Users' Selection into Reviewing

In Section 3, all users who watch a movie report their ex-post utility with a rating. This is equivalent to assume that, after watching a movie, a proportion of users are randomly drawn and rate the movie. A few papers (Dellarocas and Wood, 2008, Lafky, 2014, Hu et al., 2017) challenged this assumption: not all users have the same propensity to rate and, when users' experience is "extremely" positive or negative, they could be more willing to leave a review.

Here I show that the disappointment effect from a shift in expectations can still be isolated from the selection effect after allowing for users' selection into reviewing in line with Dellarocas and Wood (2008). To do that, I assume that users rate a movie only when their ex-post utility u_{ij} is strictly greater or lower than a threshold level $\Psi > 0$. Therefore, the expected ratings of movie i given quality θ_i and signal s becomes:

$$E(r_{ij}|\theta_i, s, \Psi) = \theta_i + \int_{\{\alpha > -E(\theta_i|s)\} \cap \{|(1-\gamma)\theta_i + \alpha - \gamma E(\theta_i|s)| > \Psi\}} \alpha dF_s(\alpha) + \gamma(\theta_i - E(\theta_i|s)).$$

The following difference in expected ratings captures the impact on ratings from a shift in the expected movie quality:

$$\begin{aligned} E(r_{ij}|\theta_i, s^+, \Psi) - E(r_{ij}|\theta_i, s, \Psi) = & -\gamma(E(\theta_i|s^+) - E(\theta_i|s)) \\ & + \int_{\{\alpha > -E(\theta_i|s^+)\} \cap \{|(1-\gamma)\theta_i + \alpha - \gamma E(\theta_i|s^+)| > \Psi\}} \alpha dF_{s^+}(\alpha) \\ & - \int_{\{\alpha > -E(\theta_i|s)\} \cap \{|(1-\gamma)\theta_i + \alpha - \gamma E(\theta_i|s)| > \Psi\}} \alpha dF_s(\alpha). \end{aligned}$$

Users' disappointment is mixed with two types of selection effects: users who decide to watch a movie are selected among the aware users; and, users who decide to rate are selected among those who watch the movie. Yet, these two types of selection effects completely vanish studying variations of Δ_{ih}^j about movies i and h with similar features ($\boldsymbol{\mu}_i = \boldsymbol{\mu}_h = \bar{\boldsymbol{\mu}}_{ih}$):

$$E(\Delta_{ih}^j|\theta_i, \theta_h, s_i^+, s_h) - E(\Delta_{ih}^j|\theta_i, \theta_h, s_i, s_h) = -\gamma(E(\theta_i|s_i^+) - E(\theta_i|s_i)).$$

This is because the same user j rate movie i and h (so the value of v_j is the same in r_{ij} and r_{hj}), and $\boldsymbol{\mu}_i = \boldsymbol{\mu}_h = \bar{\boldsymbol{\mu}}_{ih}$ and the realizations of $d(\boldsymbol{\pi}_j, \boldsymbol{\mu}_i)$ and $d(\boldsymbol{\pi}_j, \boldsymbol{\mu}_h)$ cancel out.

Accordingly, the strategy proposed to identify disappointment is robust to assuming two different types of users' selection. Allowing only certain users to rate the movie they watch does not bias the results as long as we correctly match movies with similar features.

B.2 Potential Upward Impact on Ratings from the Selection Effect

Assume that $E(\theta_i|s) = -1$ and $E(\theta_i|s^+) = 0$. With $F_s(\alpha) = F_{s^+}(\alpha) = \Phi(0,1)$, the disappointment effect equals to $-\gamma < 0$, and the selection effect equals to $(0.798 - 1.525) < 0$. Therefore, both disappointment and selection effects depress ratings after a change in signal. This may not be the case when the change in signals affects the distribution of α_{ij} among users who are aware of the movie. For instance, before receiving a nomination (with signal s), the distribution of users who are aware of movie i may be quite concentrated around the mean: $F_s(\alpha) = \Phi(0,1)$. Conversely, after the nomination (with signal s^+), the aware users have more dispersed levels of α_{ij} and $F_{s^+}(\alpha) = \Phi(0,4)$. In this case, users' selection effect equals to $(1.595 - 1.525) > 0$ and it has an upward impact on ratings.

C Appendix: Identification Strategy

C.1 Difference-in-Difference

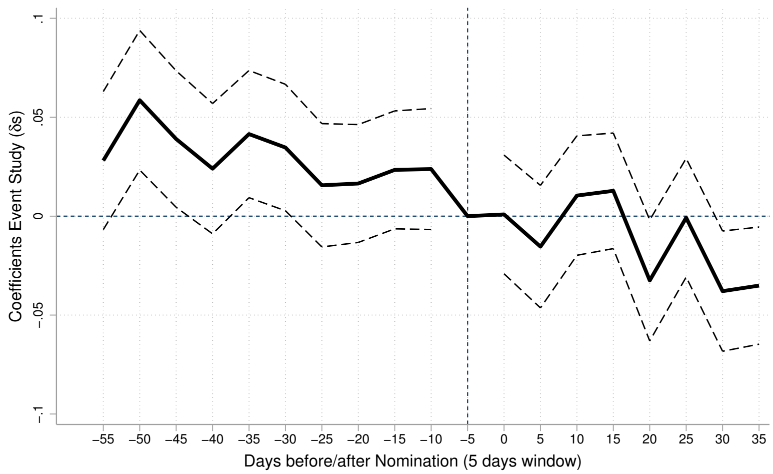


Figure C.1: The Event Study Graphs around the AMPAS Nominations for Nominated Movies

Note: The figure shows the dynamics of the parameters δ_τ of Equation 4.2 around the AMPAS Nominations starting 60 days before until 30 days after for Nominated Movies. 95% confidence intervals with robust s.e. are displayed.

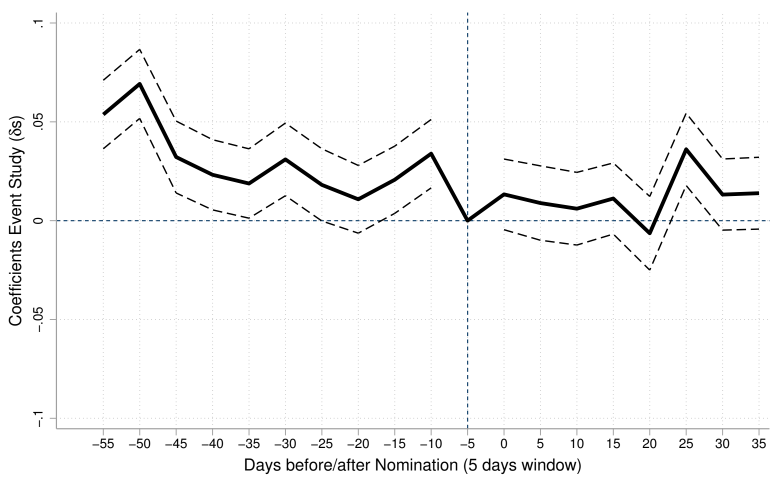


Figure C.2: The Event Study Graphs around the AMPAS Nominations for Not Nominated Movies

Note: The figure shows the dynamics of the parameters δ_τ of Equation 4.2 around the AMPAS Nominations starting 60 days before until 30 days after for Not Nominated Movies. 95% confidence intervals with robust s.e. are displayed.

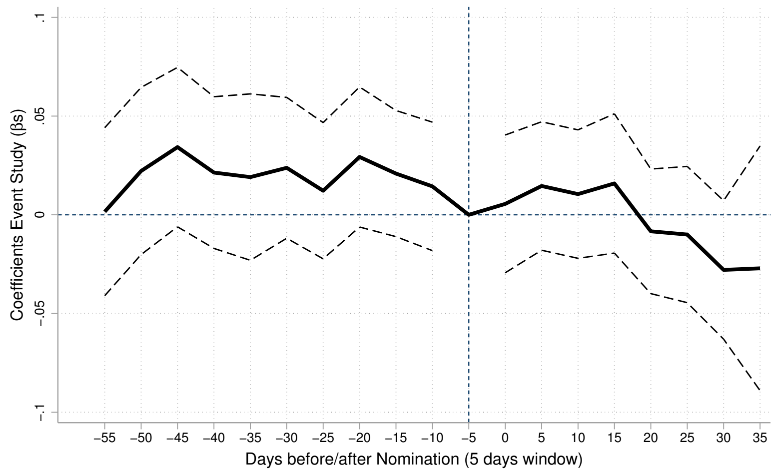


Figure C.3: The Event Study Graph around the AMPAS Nominations (full controls)

Note: The figure shows the dynamics of the parameters β_τ of Equation 4.3 adding all controls as in the column (8) of Table 3 around the AMPAS Nominations starting 60 days before until 30 days after. 95% confidence intervals with s.e. clustered at movie level are displayed.

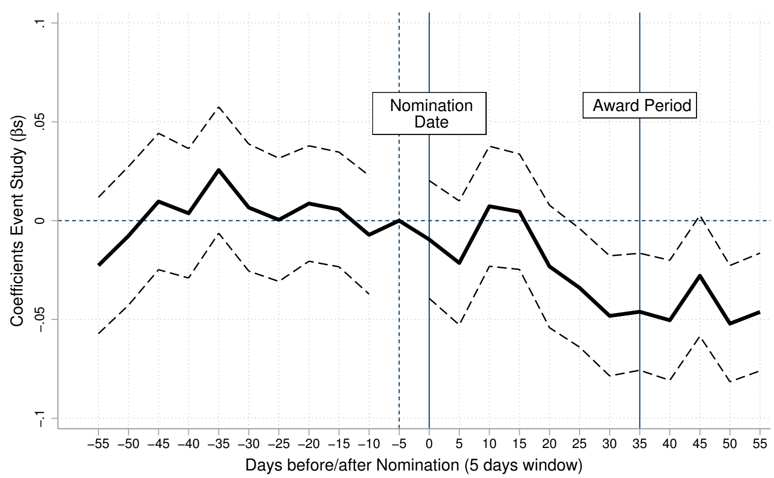


Figure C.4: The Event Study Graph around the AMPAS Nominations (longer span)

Note: The figure shows the dynamics of the parameters β_τ of Equation 4.3 around the AMPAS Nominations starting 60 days before until 60 days after. 95% confidence intervals with robust s.e. are displayed.

Table C.1: Placebo Difference-in-Difference: Regressing r_{ij} as in Equation 4.1 Anticipating the Nominations by 30 Days

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Nom_i	0.614***	0.576***	0.603***	0.604***	0.596***	0.408***	0	0
	(0.0288)	(0.0299)	(0.0300)	(0.0289)	(0.0274)	(0.0279)	(.)	(.)
T_{ij}^{-30}	-0.0190***	-0.0180***	-0.0145**	-0.0146**	-0.0112	-0.0135*	-0.0304***	0.00789
	(0.00651)	(0.00642)	(0.00579)	(0.00576)	(0.00796)	(0.00767)	(0.00509)	(0.00846)
$Nom_i \times T_{ij}^{-30}$	-0.00351	-0.00477	-0.00852	-0.00976	-0.0226	-0.0231*	-0.0140	-0.0158*
	(0.0141)	(0.0143)	(0.0143)	(0.0141)	(0.0138)	(0.0128)	(0.00944)	(0.00949)
$diff_{ij} \times 100$								-0.137***
								(0.0255)
$\bar{r}_{jt-1}$								0.00418
								(0.0133)
$n_{jt-1} \times 1000$								-0.0840***
								(0.00940)
Constant	3.392***	3.398***	3.392***	3.392***	3.391***	3.529***	3.503***	3.458***
	(0.0132)	(0.0127)	(0.0120)	(0.0118)	(0.0116)	(0.0327)	(0.00237)	(0.0488)
$\boldsymbol{X}_{ij}$						✓		
Genre FE		✓	✓	✓	✓	✓		
Year(Award) FE			✓	✓	✓	✓		
Clusters FE				✓	✓	✓		
User FE					✓	✓	✓	✓
Movie FE							✓	✓
R^2	0.0490	0.0577	0.0603	0.0606	0.285	0.296	0.420	0.420
Observations	376,133	376,133	376,133	376,133	372,214	372,214	370,776	369,897

Note: In Column 1 I present a specification with no controls or fixed effects. From Column 2 to 5 I add fixed effects referring to the genre (Genre FE), the year of the award (Year (Award) FE), the k-mean and k-median clusters (Clusters FE), and users (User FE) respectively. In Column 6, I add all time-invariant controls grouped in $\boldsymbol{X}_{ij}$. In Columns 7 and 8, I add user and movie fixed effects (Movie FE) together with time-invariant controls. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. I restrict on a time window of 60 days before the Academy Nominations. Cluster standard errors at movie level are in parentheses. After adding movie fixed effects, the parameter regarding the variable Nom_i (and all time-invariant controls) cannot be identified due to multicollinearity with the fixed effects.

C.1.1 Cluster Analysis: k-mean and k-median Algorithms

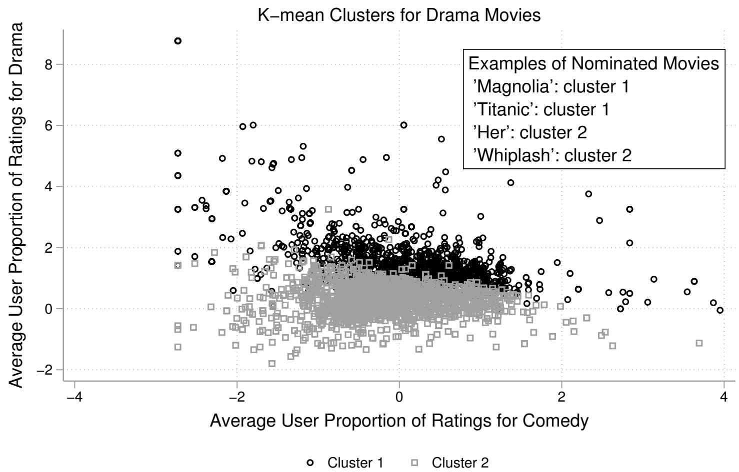


Figure C.5: K-mean Movie Clusters for Genre "Drama"

Note: The figure shows the scatter plot of "drama" movies over the variables $\bar{p}_i^{Drama}$ and $\bar{p}_i^{Comedy}$ (standardized). Each dot correspond to a movie and the movies are divided in two clusters following the k-mean algorithm.

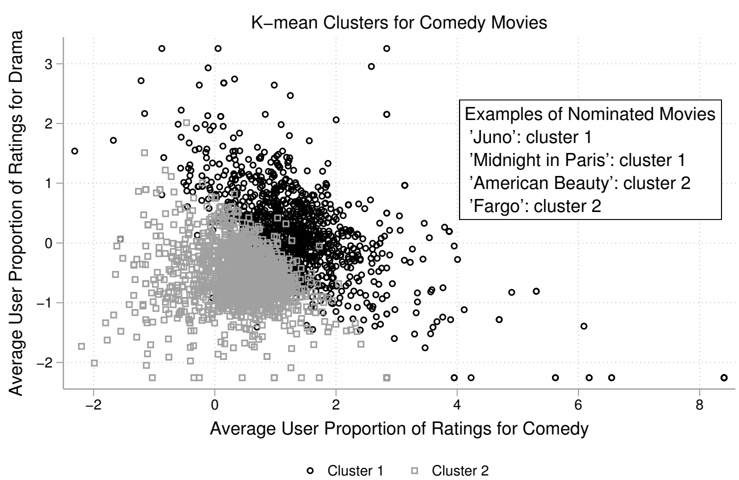


Figure C.6: K-mean Movie Clusters for Genre "Comedy"

Note: The figure shows the scatter plot of "comedy" movies over the variables $\bar{p}_i^{Drama}$ and $\bar{p}_i^{Comedy}$ (standardized). Each dot correspond to a movie and the movies are divided in two clusters following the k-mean algorithm.

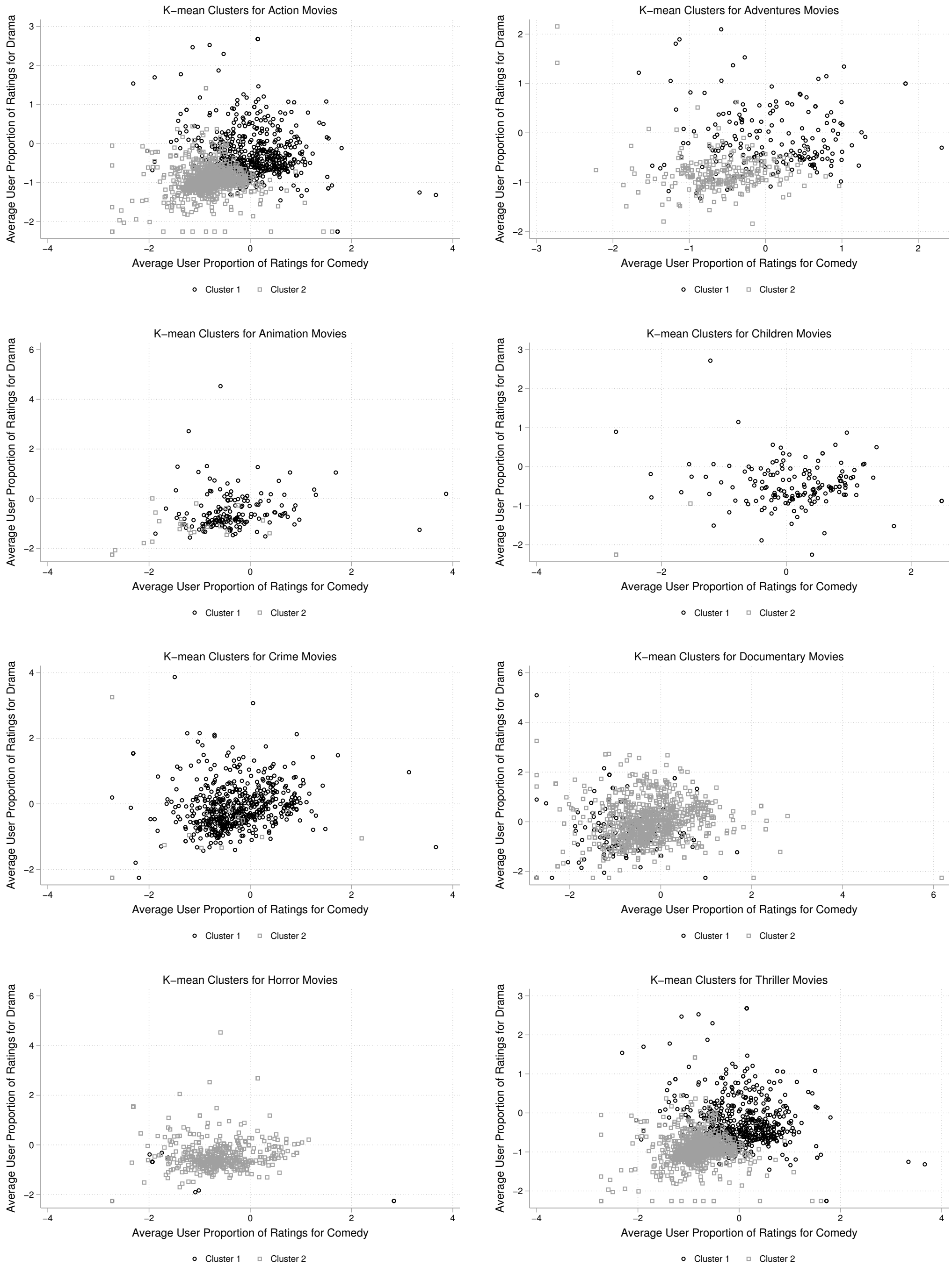


Figure C.7: K-mean Movie Clusters for the Eight Most Present Genres

Note: The figures show scatter plots of movies belonging to different genres over the variables $\bar{p}_i^{Drama}$ and $\bar{p}_i^{Comedy}$ (standardized). Each dot correspond to a movie and the movies are divided in two clusters following the k-mean algorithm.

C.2 Difference-in-Difference with Movie Matching

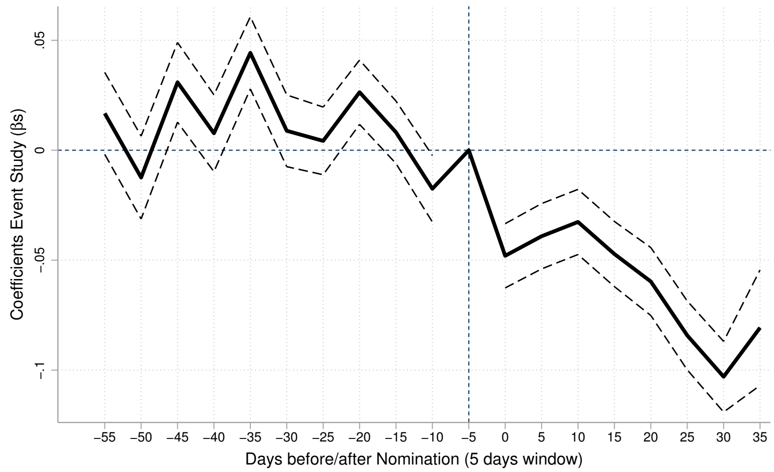


Figure C.8: The Event-Study Graph for Δ^{j}_{ih} around the AMPAS Nominations

Note: The figure shows the dynamics of the parameters β_τ of Equation 4.5 adding all controls as in the column (7) of Table 5 around the AMPAS Nominations starting 60 days before until 30 days after. 95% confidence intervals with s.e. clustered at movie combination level are displayed.

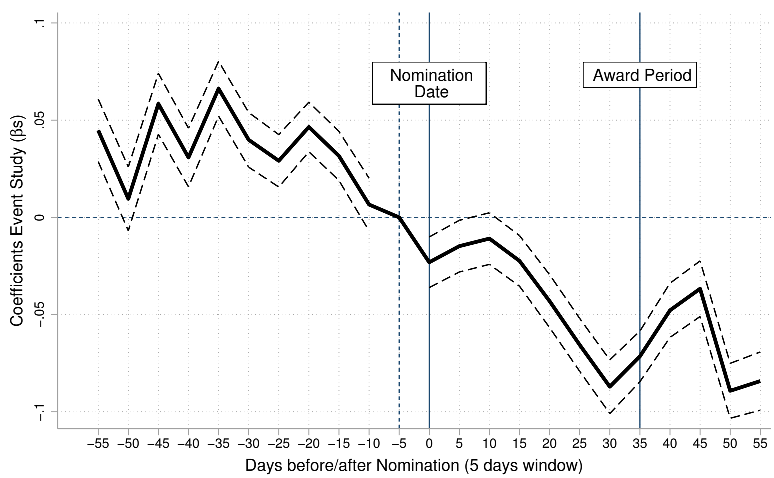


Figure C.9: The Event-Study Graph for Δ^{j}_{ih} around the AMPAS Nominations

Note: The figure shows the dynamics of the parameters β_τ of Equation 4.5 around the AMPAS Nominations starting 60 days before until 30 days after. 95% confidence intervals with s.e. clustered at movie combination level are displayed.

D Appendix: Results

Table D.1: Difference-in-Difference: Regressing r_{ij} as in Equation 4.1 with Extended Window of Time

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Nom_i	0.612***	0.573***	0.607***	0.606***	0.583***	0.408***	0	0
	(0.0252)	(0.0266)	(0.0266)	(0.0259)	(0.0261)	(0.0260)	(.)	(.)
T_{ij}	0.00334	0.00452	0.00646	0.00634	-0.0276***	-0.0282***	-0.0431***	0.00426
	(0.00585)	(0.00561)	(0.00503)	(0.00503)	(0.00645)	(0.00628)	(0.00381)	(0.00584)
$Nom_i \times T_{ij}$	-0.0284*	-0.0283*	-0.0391***	-0.0422***	-0.0553***	-0.0548***	-0.0327***	-0.0350***
	(0.0148)	(0.0145)	(0.0137)	(0.0126)	(0.0124)	(0.0118)	(0.00751)	(0.00767)
$diff_{ij} \times 100$								-0.0808***
								(0.00913)
$\bar{r}_{jt-1}$								0.0255**
								(0.0116)
$n_{jt-1} \times 1000$								-0.0939***
								(0.00688)
Constant	3.382***	3.389***	3.382***	3.383***	3.402***	3.522***	3.522***	3.442***
	(0.0123)	(0.0117)	(0.0113)	(0.0111)	(0.0106)	(0.0315)	(0.00167)	(0.0420)
$\boldsymbol{X}_{ij}$						✓		
Genre FE		✓	✓	✓	✓	✓		
Year(Award) FE			✓	✓	✓	✓		
Clusters FE				✓	✓	✓		
User FE					✓	✓	✓	✓
Movie FE							✓	✓
R^2	0.0515	0.0592	0.0628	0.0620	0.278	0.288	0.403	0.403
Observations	769,748	769,748	769,748	767,813	761,769	761,769	762,635	760,243

Note: In Column 1 I present a specification with no controls or fixed effects. From Column 2 to 5 I add fixed effects referring to the genre (Genre FE), the year of the award (Year (Award) FE), the k-mean and k-median clusters (Clusters FE), and users (User FE) respectively. In Column 6, I add all time-invariant controls grouped in $\boldsymbol{X}_{ij}$. In Columns 7 and 8, I add user and movie fixed effects (Movie FE) together with time-invariant controls. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. I restrict on a time window of 60 days before and after the Academy Nominations. Cluster standard errors at movie level are in parentheses. After adding movie fixed effects, the parameter regarding the variable Nom_i (and all time-invariant controls) cannot be identified due to multicollinearity with the fixed effects.

Table D.2: Difference-in-Difference with Movie Matching: Regressing Δ^{j}_{ih} as in Equation 4.4 with Extended Window of Time

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
T_{ij}	-0.0824***	-0.0785***	-0.0772***	-0.0765***	-0.0757***	-0.0748***	-0.0469***
	(0.000977)	(0.00218)	(0.00247)	(0.00254)	(0.00257)	(0.00257)	(0.00333)
$\bar{r}_{ht-1(j)}$					0.198***	0.190***	0.213***
					(0.0113)	(0.0113)	(0.0114)
$\bar{r}_{jt-1(h)}$					-0.0611***	-0.0627***	-0.0845***
					(0.0124)	(0.0124)	(0.0125)
$\bar{r}_{it-1(j)}$						-0.189***	-0.167***
						(0.00929)	(0.00940)
$\bar{r}_{jt-1(i)}$						0.119***	0.116***
						(0.00940)	(0.00958)
$n_{jt-1(h)} \times 1000$							0.0539***
							(0.00317)
$n_{jt-1(h)} \times 1000$							0.201***
							(0.0105)
$n_{ht-1(j)} \times 1000$							-0.0941***
							(0.00657)
$n_{it-1(j)} \times 1000$							-0.0204***
							(0.00305)
Constant	0.621***	0.589***	0.569***	0.563***	0.124**	0.394***	0.272***
	(0.000508)	(0.00112)	(0.00125)	(0.00129)	(0.0625)	(0.0636)	(0.0648)
Movie-Combination FE	✓	✓	✓	✓	✓	✓	✓
All Movies	✓						
Movies with Same Genre		✓	✓	✓	✓	✓	✓
Movies in Same Cluster (k-mean)			✓	✓	✓	✓	✓
Movies in Same Cluster (k-median)				✓	✓	✓	✓
R^2	0.261	0.252	0.249	0.256	0.256	0.257	0.259
Observations	7,050,692	1,407,841	1,014,364	914,226	905,011	903,986	903,986

Note: In Column 1 I present a specification with movie-combination fixed effect, but without restricting to a subsample of movies with similar features. From Column 2 to 4 I restrict to movies with the same genre, and movies belonging to the same genre and and the same k-mean and k-median clusters, respectively. In Columns 6 to 7, I add time-invariant controls to the specification in Column 5. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. I restrict on a time window of 60 days before and after the Academy Nominations. Cluster standard errors at movie combination level are in parentheses.

E Appendix: Extension - Disappointment and AMPAS Awards

Until now, I exploit the AMPAS nomination dates as the shock in users' expectations. However, the AMPAS award dates may also trigger disappointment by increasing users' expectations. To show this, I repeat the event study analysis in Equation 4.5 for two categories of nominated movies: those who are awarded, and those that are not. Still, now I consider the dynamics of Δ_{ih}^{j} around the AMPAS award dates. Appendix Figures E.1 and E.2 show the estimated coefficients β_τ for awarded and not awarded (but nominated) movies. In both figures, a similar negative trend is observable because of the disappointment effect caused by nominations. AMPAS awards are usually assigned 35 days after the nominations. Thus, the drop in Δ_{ih}^{j} before the award dates is contemporaneous to the nominations. Then, after the award dates, the paths of Δ_{ih}^{j} for the two categories differ. Δ_{ih}^{j} continues to drop for awarded movies after the awards.

Conversely, the evolution of Δ_{ih}^{j} is stable after the awards for those nominated movies that do not receive any award. This discrepancy is again in line with the different signals that movies received. After the AMPAS award is assigned to a movie, users' expectations rise again and they trigger further disappointment. Yet, this is only true for awarded movies.

After observing the two pre-award parallel dynamics of Δ_{ih}^{j} in the figures above, a new DiD design can be proposed to capture the disappointment related to the AMPAS awards. This is possible using the dynamics for not awarded nominated movies as a counterfactual. The main equation is:

$$\Delta_{ih}^{j} = \alpha + \lambda_{ih} + \beta_1 T_{ij}^{award} + \beta_2 Award_i \times T_{ij}^{award} + \boldsymbol{\delta X}_{ih}^{j} + \varepsilon_{ih}^{j}. \tag{E.1}$$

The only novelties relative to Equation 4.4 are the indicator T_{ij}^{award}, taking value 1 if r_{ij} is displayed after the award date, and 0 otherwise; and the interaction between the dummy variable $Award_i$, selecting nominated movies receiving at least one AMPAS award, and T_{ij}^{award}. Here the coefficient β_2 is supposed to capture the disappointment effect for awarded movies together with the elation for the not awarded nominated movies. To corroborate the presence of pre-award parallel trends among nominated movies, Figure E.3 displays the estimates of the coefficients β_τ for the following regression:

$$\Delta_{ih}^{j} = \alpha + \lambda_{ih} \sum_{t=-60}^{30} \delta_\tau \mathbb{1}(\tau = t) + \sum_{t=-60}^{30} \beta_\tau Award_i \times \mathbb{1}(\tau = t) + \varepsilon_{ij}. \tag{E.2}$$

Appendix Table E.1 shows the main results presenting seven different specifications and restricting on a window of 30 days before and after the awards (as for the previous design). A wider window (60 days before and after) is used in Appendix Table E.2.

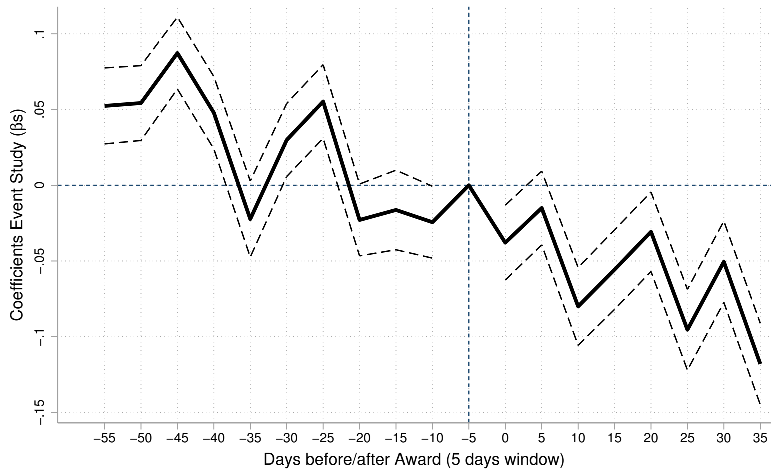


Figure E.1: The Event-Study Graph for Δ^j_{ih} around the AMPAS Awards for Awarded Movies

Note: The figure shows the dynamics of the parameters β_τ of Equation 4.5 around the AMPAS Awards starting 60 days before until 30 days after. The analysis regards only nominated movies that receive at least an award. 95% confidence intervals with s.e. clustered at movie combination level are displayed.

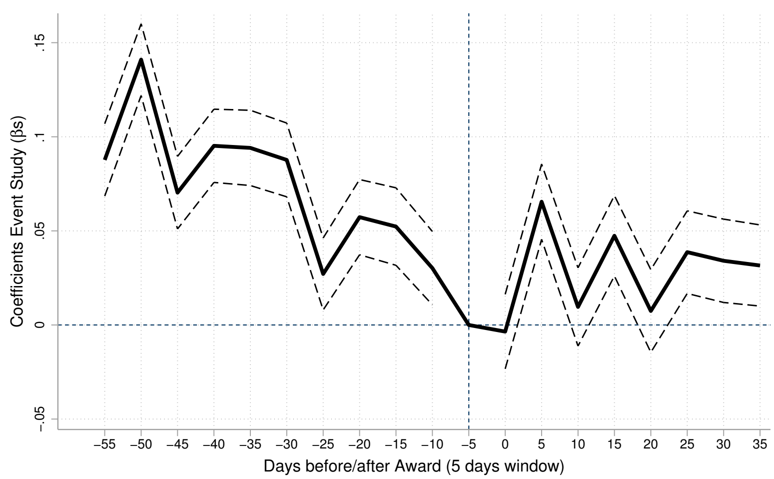


Figure E.2: The Event-Study Graph for Δ^j_{ih} around the AMPAS Awards for Not Awarded Movies

Note: The figure shows the dynamics of the parameters β_τ of Equation 4.5 around the AMPAS Awards starting 60 days before until 30 days after. The analysis regards only nominated movies that do not receive any award. 95% confidence intervals with s.e. clustered at movie combination level are displayed.

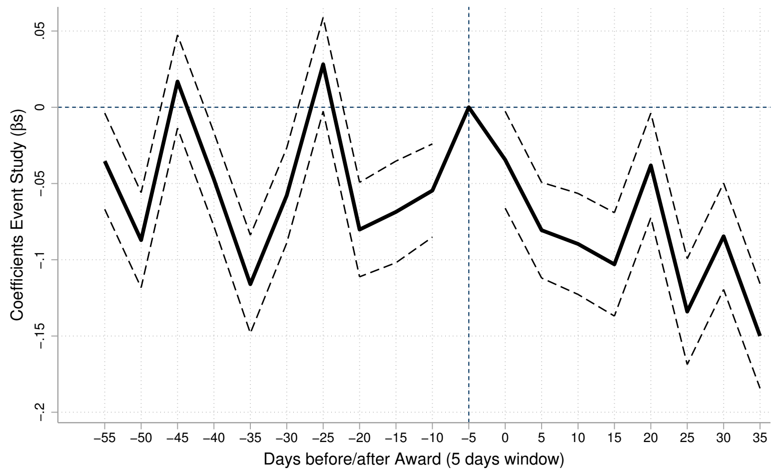


Figure E.3: The Event-Study Graph for Δ^{j}_{ih} around the AMPAS Awards

Note: The figure shows the dynamics of the parameters β_τ of Equation E.2 around the AMPAS Awards starting 60 days before until 30 days after. 95% confidence intervals with s.e. clustered at movie combination level are displayed.

All specifications show negative and significant results for the coefficient of the interaction $Award_i \times T^{award}_{ij}$. Thus, AMPAS awards seem to increase users' expectations and form a disappointment effect depressing the ratings for awarded movies after awards. Similar results are present with a wider window of time around the awards (60 days), as it is shown in Appendix Table E.2.

The magnitude of this effect is indeed similar (if not slightly smaller) relative to the disappointment caused by nominations. Yet, the comparison between these two forms of disappointment has to take into account that this new effect adds to the previous dynamics that were already affected by the nominations. Accordingly, the disappointment related to the AMPAS nominations and awards depresses the ratings for awarded nominated movies and accounts for more than fifteen percent of the rating premium for award movies.

Table E.1: Difference-in-Difference with Movie Matching: Regressing Δ_{ih}^{j} as in Equation E.1

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
T_{ij}^{award}	-0.0226***	-0.0183***	-0.00222	0.000924	-0.00463	-0.00316	0.0251***
	(0.00173)	(0.00389)	(0.00457)	(0.00476)	(0.00482)	(0.00482)	(0.00572)
$Award_i \times T_{ij}^{award}$	-0.0191***	-0.0227***	-0.0385***	-0.0513***	-0.0524***	-0.0545***	-0.0354***
	(0.00270)	(0.00608)	(0.00707)	(0.00738)	(0.00741)	(0.00740)	(0.00755)
$\bar{r}_{ht-1(j)}$					0.218***	0.211***	0.241***
					(0.0193)	(0.0193)	(0.0195)
$\bar{r}_{jt-1(h)}$					-0.408***	-0.449***	-0.574***
					(0.0564)	(0.0564)	(0.0580)
$\bar{r}_{it-1(j)}$						-0.179***	-0.160***
						(0.0129)	(0.0131)
$\bar{r}_{jt-1(i)}$						0.111***	0.106***
						(0.0131)	(0.0133)
$n_{jt-1(h)} \times 1000$							0.0552***
							(0.00533)
$n_{jt-1(i)} \times 1000$							0.214***
							(0.0140)
$n_{ht-1(j)} \times 1000$							-0.216***
							(0.0224)
$n_{it-1(j)} \times 1000$							-0.0273***
							(0.00507)
Constant	0.550***	0.519***	0.493***	0.486***	1.344***	1.765***	2.136***
	(0.000610)	(0.00134)	(0.00155)	(0.00160)	(0.230)	(0.231)	(0.241)
Movie-Combination FE	✓	✓	✓	✓	✓	✓	✓
All Movies	✓						
Movies with Same Genre		✓	✓	✓	✓	✓	✓
Movies in Same Cluster (k-mean)			✓	✓	✓	✓	✓
Movies in Same Cluster (k-median)				✓	✓	✓	✓
R^2	0.284	0.273	0.272	0.281	0.281	0.282	0.283
Observations	3,502,181	685,433	489,739	439,454	437,582	437,067	437,067

Note: In Column 1 I present a specification with movie-combination fixed effect, but without restricting to a subsample of movies with similar features. From Column 2 to 4 I restrict to movies with the same genre, and movies belonging to the same genre and and the same k-mean and k-median clusters, respectively. In Columns 6 and 7, I add time-invariant controls to the specification in Column 5. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. I restrict on a time window of 30 days before and after the Academy Awards. Cluster standard errors at movie combination level are in parentheses.

Table E.2: Difference-in-Difference with Movie Matching: Regressing Δ^j_{ih} as in Equation E.1 with Extended Window of Time

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
T^{award}_{ij}	-0.0470***	-0.0483***	-0.0351***	-0.0330***	-0.0357***	-0.0335***	0.000581
	(0.00132)	(0.00294)	(0.00341)	(0.00353)	(0.00356)	(0.00356)	(0.00409)
$Award_i \times T^{award}_{ij}$	-0.0146***	-0.0199***	-0.0345***	-0.0442***	-0.0459***	-0.0474***	-0.0253***
	(0.00211)	(0.00471)	(0.00546)	(0.00564)	(0.00568)	(0.00567)	(0.00575)
$\bar{r}_{ht-1(j)}$					0.212***	0.204***	0.228***
					(0.0128)	(0.0128)	(0.0129)
$\bar{r}_{jt-1(h)}$					-0.182***	-0.190***	-0.230***
					(0.0202)	(0.0201)	(0.0202)
$\bar{r}_{it-1(j)}$						-0.216***	-0.193***
						(0.00979)	(0.00992)
$\bar{r}_{jt-1(i)}$						0.134***	0.132***
						(0.00992)	(0.0101)
$n_{jt-1(h)} \times 1000$							0.0559***
							(0.00344)
$n_{jt-1(i)} \times 1000$							0.207***
							(0.0108)
$n_{ht-1(j)} \times 1000$							-0.151***
							(0.00778)
$n_{it-1(j)} \times 1000$							-0.0161***
							(0.00328)
Constant	0.573***	0.546***	0.521***	0.515***	0.500***	0.842***	0.811***
	(0.000390)	(0.000840)	(0.000956)	(0.000973)	(0.0906)	(0.0915)	(0.0932)
Movie-Combination FE	✓	✓	✓	✓	✓	✓	✓
All Movies	✓						
Movies with Same Genre		✓	✓	✓	✓	✓	✓
Movies in Same Cluster (k-mean)			✓	✓	✓	✓	✓
Movies in Same Cluster (k-median)				✓	✓	✓	✓
R^2	0.263	0.254	0.251	0.259	0.258	0.260	0.262
Observations	6,284,457	1,249,205	891,850	805,300	799,944	799,161	799,161

Note: In Column 1 I present a specification with movie-combination fixed effect, but without restricting to a subsample of movies with similar features. From Column 2 to 4 I restrict to movies with the same genre, and movies belonging to the same genre and and the same k-mean and k-median clusters, respectively. In Columns 6 and 7, I add time-invariant controls to the specification in Column 5. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. I restrict on a time window of 60 days before and after the Academy Awards. Cluster standard errors at movie combination level are in parentheses.