

A Service of

ZBШ

Leibniz-Informationszentrum Wirtschaft Leibniz Information Centre for Economics

Rossi, Michelangelo; Schleef, Felix

Working Paper Quality Disclosures and Disappointment: Evidence from the Academy Nominations

CESifo Working Paper, No. 11573

Provided in Cooperation with: Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Rossi, Michelangelo; Schleef, Felix (2024) : Quality Disclosures and Disappointment: Evidence from the Academy Nominations, CESifo Working Paper, No. 11573, CESifo GmbH, Munich

This Version is available at: https://hdl.handle.net/10419/312083

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



WWW.ECONSTOR.EU



Quality Disclosures and Disappointment: Evidence from the Academy Nominations

Michelangelo Rossi, Felix Schleef



Impressum:

CESifo Working Papers ISSN 2364-1428 (electronic version) Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute Poschingerstr. 5, 81679 Munich, Germany Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de Editor: Clemens Fuest https://www.cesifo.org/en/wp An electronic version of the paper may be downloaded • from the SSRN website: www.SSRN.com

- from the RePEc website: <u>www.RePEc.org</u>
- from the CESifo website: <u>https://www.cesifo.org/en/wp</u>

Quality Disclosures and Disappointment: Evidence from the Academy Nominations

Abstract

This study examines the unintended consequences of quality disclosures, focusing on how Academy Award nominations impact consumer satisfaction in the movie industry. Awards, certifications, and nominations typically signal high quality and increase consumer expectations. Yet, if the experience falls short of the expectation, they may also lead to disappointment. Using a novel dataset from Movie-Lens, we analyze user ratings for movies surrounding Academy Award nominations from 1995 to 2019. We employ a difference-in-differences framework and a recommendation-based matching approach to isolate the "disappointment effect" showing a significant decline in ratings for nominated movies post-nomination. This effect accounts for more than 7% of the pre-nomination rating gap between nominated and non-nominated films and is most pronounced among less experienced users. Our findings are validated with data from IMDb, where the effect is even more pronounced, likely due to groupthink and public review dynamics. By isolating disappointment from selection and supply side responses, this study offers insights into the dual-edged nature of quality signals and their implications for consumer welfare.

JEL-Codes: D820, D830, D910.

Keywords: quality disclosure, movie industry, expectation formation, disappointment.

Michelangelo Rossi Télécom Paris, Institut Polytechnique de ENSAE, Institut Polytechnique de Paris / France michelangelo.rossi@telecom-paris.fr

Felix Schleef Paris / France felix.schleef@ensae.fr

December 16, 2024

We would like to thank Heski Bar-Isaac, Tommaso Bondi, Rahul Deb, Xavier D'Haultfoeuille, Natalia Fabra, Roxana Fernandez, Duarte Goncalvez, Matilde Machado, Leonardo Madio, Pedro Rey Biel, Kevin Ducbao Tran, Mark J. Tremblay, and Thibaud Verge. Helpful feedback was received at the internal seminars at Universidad Carlos III de Madrid, CREST, University of Toronto, and University of Nevada; the 19th International Industrial Organization Conference, the 22nd ACM Conference on Economics and Computation (EC'21), and the 50th European Association Research Industrial Economics Conference. All errors are ours.

1 Introduction

Third-party quality disclosures are often used in markets where consumers are imperfectly informed about sellers' quality. In the hospitality industry, hotels and restaurants advertise with Michelin stars, and in the film industry, movies are promoted based on awards like the Academy Awards (colloquially known as "Oscars"). Used car sellers might show prospective buyers certificates of the car's last inspections. The positive effect of quality disclosure has already been studied in many contexts (Chen and Xie, 2005; Hilger, Rafert and Villas-Boas, 2011; Friberg and Grönqvist, 2012). When a third-party is credible, its certification or awards can effectively signal sellers' quality and attract buyers towards high-quality products.

Yet, the effect of certifications and awards on consumers' satisfaction depends on several factors (Li, Deng, Manchanda and De Reyck, 2024). Awarded sellers may have incentives to change either the quality of the product they offer or its price (Farronato and Zervas, 2022). Such adjustments may affect the consumer experience and mitigate the positive demand shock generated by the certification itself.

Even when quality and prices are held constant, quality disclosures may affect consumer satisfaction through different mechanisms. A positive certification may attract a different type of consumers (Li and Hitt, 2008). For example, a movie that receives an Oscar nomination may draw more inexperienced viewers. As a result, changes in the audience composition may lead to shifts in overall satisfaction due to the differing preferences of the consumers attracted by the award (Bondi, 2023; Lagios and Méon, 2023). In addition, quality disclosures have the potential to alter how consumers experience the product (Kovács and Sharkey, 2014). Awards often raise expectations, which may increase the likelihood of disappointment if the actual experience does not match the raised expectations. Lastly, consumers may exhibit a conformity bias, deriving more enjoyment from a positively certified product (Gergaud, Storchmann and Verardi, 2015), or experience a reverse-bandwagon effect, where the increased popularity due to an award diminishes their satisfaction (Leibenstein, 1950).

In this paper, we isolate one of these channels from the others. We provide empirical evidence on how certifications can lead to consumers' disappointment, shedding light on the unintended downside of positive quality signals in the movie industry. We examine the causal impact of Academy of Motion Picture Arts and Sciences (AMPAS) award nominations on movie ratings. Using data from the MovieLens platform, we focus on an 80-day window around AMPAS nominations (40 days before and after) from 1995 to 2019.

We start by presenting a difference-in-differences (DiD) estimate of the effect of nominations on ratings, comparing ratings for nominated and non-nominated movies before and after the AMPAS nominations. This analysis shows that nominated films have higher ratings compared to non-nominated movies, but they expe-

rience a significant drop in their ratings relative to non-nominated films after the nominations. This approach captures both selection effects and disappointment, as it does not fully disentangle shifts in consumer composition from the effect of unmet expectations.

To isolate disappointment, we propose a novel recommendation-based matching approach. Our method involves two steps: first, we train a recommender system using only ratings posted prior to the AMPAS nominations, generating a vector representation of consumer preferences. We then use this vector representation of tastes in a propensity score matching strategy, enabling us to account for taste-based selection. By comparing the ratings of these matched users posted before and after the nominations, we isolate the disappointment effect. Our results show that users who rate a nominated movie after its nomination give significantly lower ratings than users with similar tastes who rated the movie before. Specifically, this short-run disappointment effect accounts for more than 7% of the pre-nominations difference in ratings between nominated and non-nominated films. This effect is primarily driven by users who have posted relatively few ratings on the platform. This suggests that quality disclosures, such as AMPAS nominations, mainly influence (and disappoint) less experienced users, who are more likely to rely on awards as quality signals. As a placebo analysis, we repeat the matching approach for non-nominated movies and find no evidence of a disappointment effect, further supporting the robustness of our findings.

In this context, the quality of a film is fixed after its release, and prices remain relatively stable: nominated films are no more expensive to watch than non-nominated films. Therefore, aside from potential selection effects, our analysis should primarily capture changes in how consumers experience movies following an award nomination. Unlike platforms such as IMDb, Rotten Tomatoes, or Letterboxd, MovieLens does not make user ratings visible to others; they are used solely to refine the platform's recommendation algorithm. As a result, MovieLens users are less likely to incorporate biases related to groupthink or social signaling into their ratings. To further explore the relationship between disappointment and groupthink, and to provide external validity to our findings, we use data from IMDb, one of the most popular online movie databases and forums. We scrape all reviews posted for the nominated movies in our sample, along with all reviews from users who rated these films. While this dataset does not allow us to apply the same recommendation-based matching approach, we control for several observable user characteristics, helping to reduce potential confounders related to selection. The significant drop in ratings after nominations is confirmed, and its magnitude is approximately three times larger than the effect observed on MovieLens. This suggests that our findings extend to other platforms and that our analysis may underestimate the true extent of disappointment. Platforms like IMDb attract a larger share of inexperienced users, who are more prone to disappointment than those on MovieLens. Moreover, groupthink and the reverse-bandwagon effect may

further magnify the disappointment in contexts where reviews are publicly displayed and may also function as social signals.

Our analysis sheds light on the welfare consequences of quality disclosures. Most of the literature has focused on their positive impact on sellers' performance and market outcomes (Dranove and Jin, 2010). In digital contexts, Elfenbein, Fisman and McManus (2015), and Hui, Saeedi, Shen and Sundaresan (2016) find certification programs on eBay improve seller quality, while Farronato, Fradkin, Larsen and Brynjolfsson (2024) report that licensing status in an online platform for home services does not increase transactions or prices. Regarding the AMPAS awards, several studies document their positive effect on box office performance (McKenzie, 2023).

Research on how quality disclosures influence consumer experience is still relatively limited. Kovács and Sharkey (2014) and Lagios and Méon (2023) examine the impact of book awards, while Gergaud et al. (2015) and Li et al. (2024) study the effect of Michelin stars on restaurants. However, no study has yet examined how awards influence consumer ratings in the movie industry. Moreover, previous research has documented a decline in ratings for books and restaurants post-award, but these studies have not fully disentangled the disappointment effect from user selection or potential strategic responses by sellers. Our study seeks to credibly isolate the disappointment effect, which is essential for accurately assessing the true welfare impact of quality disclosures and, more broadly, for understanding how expectations shape consumers' satisfaction.

Empirical evidence of reference-based preferences is scarce outside the realms of finance and choice under uncertainty (Barberis, 2013). The work closest to ours is by Rey-Biel, Gneezy, Gneezy, Lauga and Llull (2024), who document expectations-dependent preferences in a theater setting using field experiment data. We build on this by demonstrating how reference points interact with quality disclosures. Similarly, Backus, Blake, Masterov and Tadelis (2022) examine the role of expectations in an online context, finding that eBay bidders who lose to a "Buy-It-Now" option are more likely to leave the platform, attributing this behavior to disappointment.

We also contribute to the growing body of literature on potential biases in online reviews. Several studies highlight that user reviewing behavior does not necessarily reflect only product quality. Reviews can be directly manipulated by sellers (Luca and Zervas, 2016; He, Hollenbeck and Proserpio, 2022), reviewers may act strategically (Moe and Schweidel, 2012; Nosko and Tadelis, 2015) or be influenced by social comparison (De Langhe, Fernbach and Lichtenstein, 2016; Wang, Zhang and Hann, 2018; Sunder, Kim and Yorkston, 2019) and user reciprocity (Proserpio, Xu and Zervas, 2018; Fradkin, Grewal and Holtz, 2021). In the context of movie reviews, Lee, Hosanagar and Tan (2015) demonstrates that online reviewers' behavior is influenced by prior ratings, particularly when their friends have rated a movie. Our paper provides additional evidence on the role of "external" factors influencing reviewers' behavior, as well as the impact of users' expectations.

Finally, our matching approach contributes to the emerging literature on the use of recommendation algorithms for causal inference with applications to marketing and economics. Specifically, we build on the literature on latent factor recommender systems to compute user embeddings (Billsus, Pazzani et al., 1998; Sarwar, Karypis, Konstan and Riedl, 2000; Adomavicius and Tuzhilin, 2005; Koren, Bell and Volinsky, 2009). Previous research has acknowledged the potential of recommendation algorithms for addressing identification problems for panel data (Athey and Imbens, 2019; Dell, 2024). However, to the best of our knowledge, we are the first to combine embeddings derived from a recommendation algorithm with a matching strategy. By treating the user embeddings generated by the algorithm as representations of individual preferences, we are able to address the selection concerns created by the quality disclosure.

The paper proceeds as follows. In Section 2, we provide background on MovieLens followed by a description of the dataset. Section 3 outlines the theoretical framework. In Section 4, we present the initial difference-in-differences (DiD) analysis, comparing the ratings of nominated and non-nominated movies before and after the AMPAS nominations. Section 5 introduces the recommendation-based matching approach and presents the main results. Section 6 reports the results using IMDb data. We provide some managerial implications of our findings in Section 7. Finally, Section 8 concludes.

2 Empirical Setting and Dataset

In this Section, we introduce the movie recommendation platform, MovieLens, and outline the timing and functionality of its rating process. We then describe the "MovieLens 25M" dataset and explain how we enhance it with information about AMPAS nominations and awards. Finally, we present descriptive statistics for nominated and non-nominated movies, comparing these statistics before and after the nomination announcements, along with an analysis of the types of users who watch and rate nominated movies before and after the nominations.

2.1 MovieLens and the 25M Dataset

MovieLens is an online movie recommendation platform managed by GroupLens, a research group in the Department of Computer Science and Engineering at the University of Minnesota. Registered users rate

movielens ≡•				Q			27 ☆▼ 🛔▼
top picks 🔅							
LION	O BROTHER,	SHÜHUNG		RESERVOIR DOGS			THE CUCKOOST NEST
The Lion King ★ 4.9	O Brother, Where Art Thou? ★ 4.8	The Shining	There Will Be Blood ★ 4.6	Reservoir Dogs ★ 4.6	The Grand Budapest Hotel ★ 4.6	Life of Brian	One Flew Over the Cuckoo's Nest ★ 4.6
recent releases							
Angra	HEAVIER	TIMESTALKER	FAMILY	DOMINIQUE	PIECE PIECE	LIVE	LONELY PLANET
Anora	Heavier Trip	Timestalker	Chosen Family	Dominique	Piece by Piece	We Live in Time	Lonely Planet

Figure 1. Snapshot of MovieLens.org: Main Page



Figure 2. Snapshot of MovieLens.org: Movie Page

movies, and these ratings inform a personalized list of recommended films generated by the MovieLens algorithm.

Launched in 1997, MovieLens has become a widely used recommendation system. Its data has supported research in computer science, economics, and management, contributing to numerous studies on recommender systems and user rating behavior (Chen, Harper, Konstan and Li, 2010; Aridor, Gonçalves, Kluver, Kong and Konstan, 2022; Bondi, Rossi and Stevens, 2024). Although the platform's design has evolved, the core rating mechanism has remained largely unchanged.¹

The current rating process works as follows: Figure 1 shows a snapshot of the MovieLens homepage, where a selection of movies is listed. Users can search for specific films or genres via the search bar. Clicking a movie title redirects the user to the film's detailed page (Figure 2), where further information is available. Users rate movies on a 0.5 to 5-star scale. Unlike platforms like IMDb, Rotten Tomatoes, or Letterboxd, where user ratings are publicly visible, MovieLens keeps these ratings private, using them exclusively to enhance the recommendation algorithm for each individual user.

GroupLens offers several public datasets of MovieLens ratings. For our analysis, we use the "MovieLens 25M" dataset, released in December 2019. It includes 25 million ratings for 62,423 movies, given by 162,541 users between January 1995 and November 2019. Each rating is linked to both the movie's and the user's ID, along with the date of the rating. The dataset represents a random sample of users who rated at least 20 movies. While no user-specific details are provided beyond the ID, the dataset includes some movie attributes (e.g., genre, release year, IMDb ID). Additionally, we enriched the dataset by scraping IMDb for more details such as the director, lead actors, language, country of production, and release date.

2.2 AMPAS Nominations and Awards

Since 1929, the Academy of Motion Picture Arts and Sciences (AMPAS) has awarded excellence in cinematic achievements across various categories.² However, the so-called "Big Five" categories typically receive the most public attention: Best Picture, Best Director, Best Leading Actor, Best Leading Actress, and Best Screenplay.

The award process involves two main steps each year. First, a set of nominated movies, often five per category, is selected from those that qualify. To qualify, a movie must have premiered in Los Angeles County in the previous calendar year.³ Since 2004, nominations have been announced in mid-January, whereas prior

¹Harper and Konstan (2015) provides a detailed account of the changes in the recommendation algorithm and rating process over time.

²Some categories have been discontinued, but since 2001, the awards have consistently featured twenty-four categories.

³More details on the selection process can be found on the official AMPAS website (https://www.oscars.org/oscars/

to 2004, the announcement took place in early February. The awards ceremony follows six weeks after the nominations, held at a major theater in Hollywood.⁴

To assess the impact of AMPAS nominations, we analyze all nominated and awarded movies from 1995 to 2019, focusing on the "Big Five" categories and the "Best Animated Feature Film" award. This yields 428 nominated films, of which 427 are present in the MovieLens dataset.⁵ We exclude two animated films that only received reviews after their nomination and none prior. For each year, we use the nomination date to distinguish whether a movie's ratings on MovieLens occurred before or after its nomination.⁶ To be able to use non-nominated movies for our analysis, we assign each a reference nomination date: we use the date of the first rating shown on MovieLens to select the first nomination ceremony after that rating as the reference year. Finally, we exclude any movies whose first rating on the platform appears more than two years after their production year, ensuring older films are not compared with more recent ones.

2.3 Descriptive Statistics

Before analyzing the impact of nominations and isolating the disappointment effect in user ratings, we first present some summary statistics. These include a comparison of ratings for nominated and non-nominated movies, and how these ratings change over time, focusing on the period before and after the nominations. Finally, we examine the profiles of users who rate movies before and after the nominations.

Figure 3 shows the relative frequency with which nominated and non-nominated movies receive each rating level. We observe that a large proportion of ratings for both categories fall between 3 and 4. However, nominated movies have a 5-10 percentage point higher share of ratings at 4, 4.5, and 5 compared to non-nominated movies, which are more likely to be rated 3 or below.

For both groups, the ratings, rather than exhibiting a j-shape, follow a distribution that is roughly normal but with a slight rightward skew. The absence of j-shaped distributions in our dataset supports the notion that *extremity* bias, where users often rate products only when they are either highly satisfied or dissatisfied (Dellarocas and Narayan, 2006; Hu, Pavlou and Zhang, 2009, 2017), is less likely to occur on information

rules-eligibility).

⁴In this analysis, we focus on AMPAS nominations rather than the awards themselves, as we show that users' expectations are influenced by the nominations, causing a negative shift in their ratings. Concentrating on the awards would be more problematic, as it would involve studying users who have already been impacted by a disclosure of the film's perceived quality.

⁵The missing film is the animated movie *Ernest & Celestine*.

⁶Nomination dates are sourced from the official AMPAS website and related Wikipedia entries.



Figure 3. Difference in Ratings Between Nominated and Non-Nominated Movies

Note: The figure shows the distributions of ratings for nominated and non-nominated movies. Nominated movies (in gray) have higher ratings than non-nominated movies.

	Mean	SD	Min	Max
Average Ratings				
All	3.704	.2278	2.753	4.284
Before Nom.	3.841	.3268	2.286	4.833
After Nom.	3.697	.2285	2.747	4.284
Number of Ratings				
All	8193	10721.7	59	57309
Before Nom.	374.3	643.9	1	4986
After Nom.	7819	10392.7	53	57303
Average Ratings (80-days window)				
All	3.826	.31	2.732	4.75
Before Nom.	3.832	.3429	2.286	4.833
After Nom.	3.823	.3242	2.5	4.857
Number of Ratings (80-days window)				
All	264	282.9	2	1832
Before Nom.	126.5	155.3	0	1032
After Nom.	137.4	140.8	2	927

Table 1. Summary Statistics: Nominated Movies

Note: The table presents information on all 426 movies nominated for AMPAS awards included in our sample. We report the average ratings and the number of ratings both before and after the nominations. The first two panels include all available ratings without any time restrictions, while the second two panels focus on ratings within a 80-day window surrounding the nominations.

platforms with a large scale of ratings, such as MovieLens (Schoenmueller, Netzer and Stahl, 2020).

Movie ratings are not static; they can vary over time, particularly in response to significant events like AMPAS nominations. In Table 1 we show that nominated films tend to receive slightly higher average ratings prior to their nomination, suggesting that initial audience perceptions may be more favorable during this period. This trend may indicate a sense of disappointment among viewers, but it could be influenced by the characteristics of users who choose to rate movies before or after the nominations. The role of the nominations is more pronounced in terms of audience engagement, as evidenced by the increase in the number of ratings following the nominations. Table A1 in Appendix presents the same statistics for non-nominated movies. Notably, the average rating for all non-nominated films is lower than that of their nominated counterparts. Before nominations, the average ratings for these movies remain relatively stable, showing little fluctuation after the nominations. When examining the number of ratings, there is an observable increase in engagement following the nomination announcements, though this increase is not as pronounced as that seen with nominated films. Figure A1 in Appendix examines the daily arrival of ratings around the nomination period for both nominated and non-nominated films. For nominated movies, we observe a positive time trend after nominations, characterized by an increase in the number of ratings, especially around weekends. In contrast, we do not see similar positive trends for non-nominated movies. Instead, we observe a noticeable influx of ratings in the weeks leading up to the nominations.

2.4 User Selection Before and After the Nominations

As the number of ratings varies over time, the set of users who post ratings for nominated films may change before and after the AMPAS nominations.

To explore the self-selection of users into watching and rating nominated movies, we focus on four dimensions that might describe users' characteristics and the determinants of their rating behavior, namely their average rating, their experience on the platform, the types of movies they rate, and their rating stringency. For each user *i*, we define the average user rating \bar{r}_i^j before watching movie *j* as the average of the ratings that user *i* posted before rating the movie *j*. We also define the total number of ratings user *i* posted before rating movie *j* as N_i^j . These two variables provide insight into the users' experience at the time they rated the movies and the types of movies they rated. In line with the analysis of Bondi et al. (2024), we measure the stringency of a user relative to others. To do so, we calculate a rating stringency as $s_{ij} = \bar{r}_j - r_{ij}$, where \bar{r}_j is the average rating for movie *j* across all users, and r_{ij} is the rating given by user *i* for that movie. We define the stringency of user *i* in rating movie *j*, denoted as \bar{s}_i^j , as the moving average of the stringency of each user over their previous 10 ratings. Positive values of \bar{s}_i^j indicate lower than average ratings, suggesting

	Mean	SD	Min	Max
Average User Ratings				
All	3.451	.2469	1.207	5
Before Nom.	3.45	.2925	1.038	5
After Nom.	3.463	.2958	1.14	5
Number of User Ratings				
All	752.9	696.468	2	11354.3
Before Nom.	725.3	716.4	0	14959
After Nom.	764.4	1032.24	1	25034
Average User Stringency				
All	.0915	.3189	-2.204	3.161
Before Nom.	.0943	.3644	-2.204	3.204
After Nom.	.077	.3799	-2.855	3.161
Average Share of User Ratings for Current-Year Releases				
All	.052	.0374	0	.6667
Before Nom.	.0585	.0463	0	1
After Nom.	.0451	.0408	0	.6667

Table 2. User Selection Before and After the Nominations: 80-days window

Note: The table presents descriptive statistics of users who rated movies nominated for AMPAS Awards included in our sample. We only included users who submitted a rating within an 80-day window surrounding the nominations. We report the average user rating, the average number of ratings per user, the average rating stringency (a measure of how strictly users rate, based on how their ratings compare to the average rating of the movie at the time), and the average share of ratings for movies produced in the same year as the rating. The stringency of a user is defined as a moving average of the stringency of their past five ratings. These statistics are presented separately for users who rated the movies before and after the nominations.

greater stringency. Finally, we compute the proportion of movies that user *i* rated before watching movie *j* that were released in the current year of the rating. This share provides information about users' interest in recent movies and their likelihood to watch and rate movies that have been or are expected to be nominated for AMPAS awards.

In Table 2 we present a summary of these variables for users who rated nominated movies within an 80-day window surrounding the nominations.⁷ In terms of average user ratings, we observe a slight increase after nominations, suggesting that the quality signal provided by AMPAS nominations may attract users who are more likely to give higher ratings on average. This could be because these users are generally more positive in their ratings, or because they tend to select higher quality movies. Additionally, the total number of ratings a user has posted before reviewing a nominated movie is higher after the nominations. We also observe a decrease in user stringency after the nominations: those users who watch and rate nominated movies after the nominations are less critical and generally post higher-than-average ratings. Finally, the

⁷Table A2 in Appendix presents a similar analysis, but considering the full history of ratings.

share of ratings for current year releases decreases after the nominations, suggesting that users who rate nominated films after the nominations are generally less inclined to rate recent films.

All of this evidence points to a shift in the profile of users before and after the nominations. Those who rate nominated movies before they receive AMPAS nominations rate fewer movies, tend to be more stringent, and are more active in rating recent releases. To understand how this selection may affect our identification strategy, in the next Section we present a conceptual framework in which we model users self-selecting into movie viewing.

3 Conceptual Framework

To guide the empirical analysis of how users experience disappointment following AMPAS nominations, we propose a conceptual framework. This framework clarifies our assumptions regarding the movie rating process and the mechanisms through which quality signals influence user ratings. It also helps illustrate how different empirical strategies can isolate the distinct effects of nominations on user evaluations.

3.1 The Decision to Watch: User Utility, Quality Signals, and Ratings

Each movie is defined by a quality parameter $\theta_j \in \mathbb{R}$ and a vector $\mu_j \in \mathbb{R}^M$, which describes the movie's "position" in an *M*-dimensional feature space (e.g., the movie's genres, actors, length, etc.). Similarly, each user is characterized by $v_i \in \mathbb{R}$, representing the utility derived from going to the theater (regardless of the movie watched), and a vector $\pi_i \in \mathbb{R}^M$, which describes user *i*'s preferences in the feature space (e.g., their likes and dislikes for specific genres, actors, screenwriters, etc.).

We assume that users are aware of their own position in the feature space as well as the positions of the movies. However, before watching movie j, users do not perfectly observe its quality. Thus, the expected utility for user i when watching movie j is defined as follows:

$$E(u_{ij}) = E(\theta_j) + v_i - d(\pi_i, \mu_j), \qquad (1)$$

where the function d(.) is a strictly increasing, continuous norm that describes the distance between the position of movie *j* and the preferences of user *i*. Accordingly, we can rewrite the expected utility as $E(u_{ij}) = E(\theta_j) + \alpha_{ij}$, where $\alpha_{ij} = v_i - d(\pi_i, \mu_j)$ represents the sum of all observed parameters: the utility from going to the theater, and the utility "match" value between user *i* and movie *j*.

After watching movie i, user j observes its quality. Ex-ante expectations may enter into the ex-post

utility of users for various reasons. We believe that MovieLens users are unlikely to incorporate biases related to groupthink or social signaling into their ratings: MovieLens does not make user ratings visible to others. Yet, the expected utility $E(u_{ij})$ may serve as a reference point and affect user *i* ex-post satisfaction. Thus, we define the ex-post utility for user *i* after watching movie *j* as follows:

$$u_{ij} = \theta_j + \alpha_{ij} + \gamma(\theta_j + \alpha_{ij} - E(u_{ij})) = \theta_j + \alpha_{ij} + \gamma(\theta_j - E(\theta_j)),$$
(2)

where $\gamma(.)$ represents the user's gain-loss utility factor as in Kőszegi and Rabin (2006). We assume $\gamma(.)$ to be linear with the parameter $\gamma > 0$ multiplying the gain/loss term. Consequently, the higher user *i*'s expectation about the quality of movie *j*, $E(\theta_j)$, the lower the gain-loss utility term, and the lower user *i*'s ex-post utility.

For each movie j, 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 *j* is released and a unit measure of potential users associated with movie *j* is formed;
- 2. A signal *s* appears and a proportion of users becomes aware of movie *j*. The distribution of α_{ij} among aware users after signal *s* at time *t* is $F_{s,t}(\alpha)$;
- 3. Aware users form expectations about the quality of the movie $E(\theta_j|s)$ and watch movie *j* if $E(u_{ij}|s) = E(\theta_j|s) + \alpha_{ij} > 0$;
- 4. Finally, users who rate movie *j* are randomly drawn from the set of users who watched it, and they report their ex-post utility in their ratings: $r_{ij} = u_{ij}$.⁸

3.2 The Impact of the Nominations for the Academy Awards

A positive quality signal like the AMPAS nominations alters the profile of users aware of the movie (i.e., $F_{s,t}(\alpha)$ depends on *s*) and influences the quality expectations of all users aware of the movie. These two effects impact the ratings before and after the nominations. Specifically, the expected rating of movie *i*, given its quality θ_i and signal *s*, is as follows:

$$E(r_{ij}|\theta_j,s) = \theta_j + \int_{\alpha > -E(\theta_j|s)} \alpha, dF_{s,t}(\alpha) + \gamma(\theta_j - E(\theta_j|s)).$$
(3)

⁸Users' decision to rate may depend on their ex-post utility levels. Specifically, users who rate products online may be more likely to have experienced either extremely positive or extremely negative utility levels from the transaction (Dellarocas and Narayan, 2006; Hu et al., 2017; Schoenmueller et al., 2020). In the next subsections, we show how changes in quality signals affect aggregate ratings and discuss the role of this assumption (no selection into reviewing).

To better understand how a shift in expected movie quality affects ratings, let us assume the signal changes from s to s⁺, where $E(\theta_j|s) < E(\theta_j|s^+)$. The difference in expected ratings for the same movie j but different signals is given by:

$$E(r_{ij}|\theta_j, s^+) - E(r_{ij}|\theta_j, s) = -\gamma(E(\theta_j|s^+) - E(\theta_j|s)) + \int_{\alpha > -E(\theta_j|s^+)} \alpha, dF_{s^+,t}(\alpha) - \int_{\alpha > -E(\theta_j|s)} \alpha, dF_{s,t}(\alpha).$$
(4)

The first term, $-\gamma(E(\theta_j|s^+) - E(\theta_j|s))$, represents the *disappointment effect*, as it captures the negative impact on ratings due to increased expectations when $\gamma > 0$. This effect is present across all users and its magnitude depends on the size of γ and the variation of quality expectations due to the change in signals.

The second term, $\int_{\alpha > -E(\theta_j|s^+)} \alpha, dF_{s^+,t}(\alpha) - \int_{\alpha > -E(\theta_j|s)} \alpha, dF_{s,t}(\alpha)$, represents the *selection effect*. It reflects the change in the profile of users who decide to watch movie *i* due to the shift in the signal. While disappointment always drives ratings downward (for $\gamma > 0$), the effect of selection is more ambiguous. If $F_{s,t}(\alpha) = F_{s^+,t}(\alpha)$, selection will reduce the average ratings. However, if $F_{s,t}(\alpha) \neq F_{s^+,t}(\alpha)$, the selection effect could potentially drive ratings upward. The sign of the selection effect becomes even more challenging to predict if we assume some degree of selection into reviewing. In fact, if different types of users choose to rate a movie depending on the signal they receive and the ex-post utility they experience (Dellarocas and Narayan, 2006), it becomes much more likely that $F_{s,t}(\alpha) \neq F_{s^+,t}(\alpha)$, further complicating the prediction of the selection effect. However, the selection into reviewing does not directly affect the disappointment effect, which primarily depends on the gap between users' expectations and the actual quality of the movie, regardless of the profile of users who post ratings.

3.3 Isolating the Disappointment Effect

A comparison of average ratings for the same movie under different quality signals does not allow us to isolate disappointment from selection, as shown in Equation 4. Nevertheless, this approach offers valuable insights into the overall effect of quality signals on user satisfaction. In Section 4, we begin by exploring the empirical counterpart of this analysis with a DiD approach.

To isolate the disappointment effect from selection, we must observe the user-movie features α_{ij} and hold them constant when comparing ratings with and without nominations:

$$E(r_{ij}|\theta_j,\alpha_{ij},s^+) - E(r_{ij}|\theta_j,\alpha_{ij},s) = -\gamma \left(E(\theta_j|\alpha_{ij},s^+) - E(\theta_j|\alpha_{ij},s) \right).$$
(5)

By comparing users with identical α_{ij} values, we can fully isolate disappointment. The empirical analog of this method, based on recommendation-based matching, is presented in Section 5.

In this framework, γ is assumed to be constant across all users. Therefore, restricting the analysis to users who can be successfully matched based on observable characteristics does not affect the estimation of the disappointment effect. However, this assumption might be overly restrictive, as different users may experience varying degrees of disappointment. In such cases, the estimated size of the disappointment effect should be interpreted with caution, as it would only be valid for the subset of users for whom matching is feasible.

4 Comparing Ratings for Nominated and Non-Nominated Movies

To assess how a shift in expected movie quality influences users' satisfaction, we can compare the ratings of users who rated nominated movies before and after the nomination date. This approach mirrors the idea in Equation 4 from the previous Section, where we aim to compare ratings of the same movies under different quality signals. In practice, however, comparing average ratings before and after nomination simultaneously captures both the impact of the AMPAS nomination as a quality signal and any temporal differences.⁹ Specifically, we compare ratings posted around January and February, before the nominations, with those posted later in February and March, after the nominations.

As our theoretical framework shows, the type of users who is aware of a movie might change over time. To account for the fact that fans of the actor, screenwriter, or genre might be among the first to watch a movie, we implement a DiD design, comparing ratings for nominated and non-nominated movies before and after the nomination date. For non-nominated films, we assign a "nomination date" using the date of the first rating on MovieLens to identify the first nomination ceremony after that rating as the reference year. To capture the impact of the AMPAS nominations on nominated movies, we use the following DiD specification:

$$r_{jt} = \gamma_1 + \gamma_2 Nominated_j + \gamma_3 post_t^{Nom} + \delta Nominated_j \times post_t^{Nom} + \beta \mathbf{X}_{jt} + \varepsilon_{it},$$
(6)

where r_{jt} is the rating of movie *j* on day *t*. Nominated_j equals 1 if movie *j* has received at least one AMPAS nomination; and post^{Nom} equals 1 for all ratings posted after the nominations. The set of controls \mathbf{X}_{jt} includes genre, year, movie, and day fixed effects, where the day fixed effects capture all variations in ratings that posted on specific days relative to the nomination date.

⁹Recall that the distribution of α_{ij} , $F_{s,t}(\alpha)$, is a function of signal and time.

Table 3 shows four specifications for the DiD estimates in Equation 6. The sample includes MovieLens ratings from 1995 to 2019, posted within an 80-day window around the AMPAS nomination date, excluding ratings posted after the AMPAS award ceremony. We cluster standard errors at the movie level.

In Column (1), we do not use any fixed effects, while we add fixed effects for the genre and the production year of the movies in Column (2). In Column (3), we add movie fixed effects that absorb all timeinvariant movie characteristics. Finally, Column (4) presents the results with movie and day fixed effects, accounting for seasonality and time-varying confounders that affect all movies uniformly. The estimates of the coefficient δ in Equation 6 are negative and significant across all specifications. The drop in ratings posted for nominated movies after the nominations is approximately 0.02 stars. This drop may seem small compared to the average rating (around 3.5); however, in Columns (1) and (2), we observe that the difference in ratings between nominated and non-nominated movies (before nominations) is less than 0.6 stars. Thus, the short-term negative effect linked to the quality signal from nominations is approximately 5% of the "rating premium" experienced by nominated movies, which may reflect either the higher quality of these movies or the characteristics of users who rate them.

The coefficient δ captures the effect of the AMPAS nominations on nominated movies under the assumption that non-nominated movies provide a valid counterfactual group for the evolution of ratings after nomination. To check for potential pre-trends between treated and control groups, we illustrate the evolution of r_{jt} over time with an event-study approach. We specify the following lead-lag model, in which r_{jt} is regressed over the interaction between the dummy *Nominated*_j and a set of dummy variables for each 5-day window around the AMPAS nomination dates. The model controls for movie and day fixed effects:

$$r_{jt} = \gamma_j + \rho_t + \sum_{\tau = -40 \text{ days}}^{40 \text{ days}} \delta_{\tau} Nominated_j \times 1(t = \tau) + \varepsilon_{jt}.$$
(7)

We present the estimates of Equation 7 in Figure 4, plotting the estimated δ_{τ} from 40 days before to 40 days after the AMPAS nominations, excluding ratings posted after the AMPAS award ceremony. The coefficients for days prior to the nominations are close to zero and show no clear trend, suggesting that the evolution of ratings for nominated and non-nominated movies was similar prior to the nominations. This finding supports the parallel trend assumption, which is essential for our analysis. In contrast, after the AMPAS nominations, particularly around 15 days after, the estimated δ_{τ} turns negative, indicating a relative decline in ratings for nominated movies compared to non-nominated ones.

The observed lag in the ratings decline may be due to the time it takes for users who have watched the movies after the nominations to post their ratings. Similarly, ratings posted on the same day or a few days

	(1)	(2)	(3)	(4)
Nominated _j	0.590***	0.575***	0.000	0.000
	(0.023)	(0.024)	(.)	(.)
$post_t^{Nom}$	0.007	0.006	-0.005	0.000
	(0.005)	(0.005)	(0.003)	(.)
Nominated $_i \times post_t^{Nom}$	-0.034***	-0.037***	-0.019***	-0.018**
v -	(0.011)	(0.011)	(0.007)	(0.007)
Genre FEs		\checkmark		
Year FEs		\checkmark		
Movie FEs			\checkmark	\checkmark
Date FEs				\checkmark
R^2	0.05	0.06	0.20	0.21
Ν	497,266	497,266	497,266	497,266
Mean Dep. Var.	3.503	3.503	3.503	3.503

Table 3. Difference-in-Differences: Comparing Nominated and Non-Nominated Movies



Figure 4. Event study: Comparing Nominated and Non-Nominated Movies

Notes: The sample includes MovieLens ratings from 1995 to 2019, posted within an 80-day window around the AMPAS nomination date, excluding ratings posted after the AMPAS award ceremony. MovieLens ratings are regressed on movie fixed effects and on interactions between a dummy variable (indicating whether a movie is nominated for an AMPAS award) and a set of five-day interval dummies centered on the AMPAS nomination date. The graphs display the estimated coefficients of these interactions, with the coefficient for the five-day interval of the nomination date normalized to zero. Standard errors, clustered by movie, are reported at the 10% level.

after the nomination are likely to be from users who watched the movies before the AMPAS nomination dates, and thus may not reflect the influence of the nominations. To account for this lag, Table B1 in Appendix presents the same specification as in Column (4) of Table 3, excluding ratings posted 5, 10, and 15 days after the nominations. When we omit these ratings (as they may not fully reflect treatment effects), the coefficient δ doubles in magnitude, consistent with the lag observed in the event study.

The parallel trends in Figure 4 suggest that the ratings for nominated and non-nominated films evolved similarly prior to the AMPAS nominations. However, not all non-nominated films may provide an appropriate counterfactual for nominated films, which tend to be of higher quality and generally receive higher ratings (as discussed in Section 2.3). To examine whether our results are sensitive to the selection of non-nominated films, Table B4 in Appendix presents three alternative specifications that select the pool of non-nominated films. We restrict the analysis to movies with at least one professional review on the Metacritic website,¹⁰ those with an average Metacritic score above 50 (out of 100), and those with an average IMDb rating above 5. By refining the control group to exclude non-nominated films. All specifications confirm a significant negative effect of nominations on the ratings of nominated films, suggesting that the observed ratings drop is not highly sensitive to the selection criteria for non-nominated films in the control group.

In Table B3 in Appendix, we examine how the results vary across different groups of nominated movies to identify which groups drive the observed effects. Our findings indicate that the negative effect of nominations is particularly pronounced for movies that receive more than one Oscar nomination, regardless of whether they ultimately win an AMPAS award. This suggests that, for a nomination to significantly alter users' expectations and induce disappointment, the signal must be both informative and widely recognized. By contrast, movies with only one nomination are less likely to experience this effect, as a single nomination may not substantially influence users' expectations.

As discussed above, this approach does not disentangle disappointment from the user selection effects associated with nominations; however, it provides valuable insights into the combined effects of selection and disappointment in this context. In the next Section, we introduce a recommendation-based matching approach to compare users with similar tastes and preferences for movies before and after the nominations. Complementary to this approach, we can control for observable user characteristics within the DiD framework used in this Section. This may not fully remove all user self-selection, but it allows us to assess whether the results in Table 3 are robust to the inclusion of user characteristics. This analysis is presented in Table B4 in Appendix, where we replicate the specification from Column (4) of Table 3 but add controls for users'

¹⁰For more information, see: https://www.metacritic.com/.

average rating prior to rating movie j, the total number of ratings posted by user i before rating movie j, user i's stringency and the proportion of current-year releases rated by user i prior to rating j. We also include user fixed effects in the final specifications. Notably, the significance and magnitude of the effect remain stable if not slightly larger in the most restrictive specification.

This result may seem counterintuitive at first; we might expect users who rate nominated movies after the nominations to be drawn in by the advertised high quality, potentially making them more critical of the nominated movies. However, as shown in Section 2.4, users who rate movies after the nominations tend to be less stringent and less attentive to new releases. This may partly explain why, when controlling for various user characteristics, we observe an even larger negative effect of nominations on ratings. We will further explore the role of selection in the next Section, comparing the coefficients of the matching with the ones previously estimated.

5 A Recommendation-based Matching Design

As shown in Equation 5, to identify the disappointment effect associated with the AMPAS nominations for a nominated movie *j*, we must account for taste-based selection, i.e. the fact that the nominations can attract moviegoers a different kind of moviegoer. To implement this approach and condition on α_{ij} , we use a recommendation-based matching design. Specifically, we train a recommendation system using only ratings posted before the AMPAS nominations to obtain a vector representation of users' "tastes" (their positions π_i in the feature space). We then integrate these results into a matching approach, comparing ratings of two users for the same nominated movie before and after the nomination date, matching on users' tastes and other characteristics that capture variation in baseline utility v_i .

In the remainder of this Section, we first describe the recommendation system and its characteristics. Next, we present the matching design and the main results of our analysis, which reveal a negative and significant disappointment effect following the AMPAS nominations. As a placebo test, we use the predicted ratings generated by the recommendation system as the outcome instead of the actual ratings; in this case, the matching yields null results, suggesting that the matched users indeed have similar preferences for the nominated movies. Then, we apply the same matching analysis to non-nominated movies and find no evidence of disappointment. Finally, we examine the evolution of the difference between actual and predicted ratings over time.

5.1 Training the Recommendation System

To compute a representation of user tastes, we rely on the literature on collaborative filtering and matrix factorization recommender systems (Koren et al., 2009; Billsus et al., 1998; Sarwar et al., 2000; Goldberg, Nichols, Oki and Terry, 1992; Paterek, 2007). Matrix factorization techniques aim to approximate the matrix of ratings as the product of two smaller matrices, one representing users and the other representing items, in our case, movies. The two smaller matrices, *P* and *Q*, are of dimensions $I \times M$ and $J \times M$, respectively, where *I* is the number of users, *J* is the number of movies, and *M* is the number of latent factors, which we set to 50.¹¹ We employ the singular value decomposition algorithm as described in Koren et al. (2009), where the embedding matrices are optimized via stochastic gradient descent.¹² The embedding matrices computed in this manner capture certain aspects of the items and users.

The latent factors represent a computerized alternative to intuitive categories that might influence movie recommendations, such as a user's preference for action, drama, or romance movies, or their affinity for specific actors. However, these embeddings are not directly interpretable (Koren et al., 2009). Importantly, two users with identical embeddings can be said to share the same tastes in the sense that they will have the same *predicted* ratings for all movies and receive the same recommendations from the algorithm.

This approach allows us to match users based on their embeddings, thus accounting for user tastes and isolating the disappointment effect from selection. To assess the quality of the matching, we run the matching procedure on a sample of ratings outside the treatment period and compare the root-mean-squared error (RMSE) and mean absolute error (MAE) with benchmarks from the recommender algorithm (results in Appendix C.2). For implementation, we use the singular value decomposition as implemented in the scikit-surprise python library. With this approach, ratings are modeled as follows:

$$\hat{r}_{ij} = \mu + b_i + b_j + q_i^T p_i,$$

where μ is the mean rating, and b_i and b_j are user- and movie-specific bias terms. Here, p_i represents a vector of latent factors (or "embeddings") for user *i*, while q_j represents a vector of latent factors for movie *j*. These latent factor vectors form the columns in the embedding matrices *P* and *Q*.

For each nomination year, we train the recommender system on all ratings up to 40 days before the nomination date, obtaining the user embeddings and predicted ratings for each year. We then merge the user

¹¹See Appendix C.2 for more information on the hyperparameter choice for training the recommender algorithm and the results in terms of efficiency.

¹²The same approach discovered by Simon Funk during the Netflix challenge. See: https://sifter.org/simon/journal/20061211.html.

embeddings and the predicted ratings with the dataset containing user ratings and other user characteristics.

5.2 Matching Users Before and After the AMPAS Nominations

We now present the results of our matching analysis, comparing ratings posted 40 days before and after the AMPAS nominations. To isolate the effect of potential disappointment from selection, we match users based on their 50-dimensional embedding vectors and other attributes reflecting their preferences. These embeddings are designed to capture a comprehensive representation of users' tastes and behaviors. By matching on embeddings instead of predicted ratings, we ensure that the users being compared share not only similar preferences for the specific movies of interest but also for a broad range of genres and types of movies. Thus, the matched users have comparable likelihoods of watching and rating the movie after its nomination.

Table 4 presents OLS and propensity score matching estimates using various sets of variables to match users. Column 1 (OLS) reports estimates with movie fixed effects, comparing ratings of users who watched the same nominated movies without adjusting for user characteristics. Columns 2 to 6 (M1–M5) provide propensity score matching estimates, incorporating user embeddings (50 dimensions) and exact matching on movie identifiers, along with quintiles of different user attributes: average user ratings, the number of movies watched, stringency levels (as defined in Section 2.3), and the share of movies watched by the user in the year of their release.

Across all specifications, the estimates reveal a statistically significant drop in ratings, with matching estimates being of a slightly larger magnitude compared to the OLS estimates. This difference may stem from the fact that more stringent users are disproportionately likely to watch nominated movies before the nominations. Since these users tend to give lower ratings, the DiD estimates may be influenced by this "positive" selection, which partially offsets the negative impact of disappointment.

In Tables C2, C3, C4, C5, and C6 in Appendix, we show that the results are robust when we vary the embedding dimensions from 10 to 200 and to using multivariate distance matching (with larger estimates than those from propensity score matching). The negative estimates are not driven by users who rate a movie shortly after the nominations. As shown in Appendix Table C7, when we exclude ratings posted within 5, 10, and 15 days following the nominations, the estimates remain negative, significant, and of similar magnitude.

We also conduct a series of placebo tests to reinforce these findings. Thus far, we have used embeddings to match users with similar tastes and preferences. However, the recommender system also enables us to compute predicted ratings for each user, based solely on ratings posted prior to the nominations. These predicted ratings, by design, are unaffected by disappointment or nomination-related effects.

	(OLS)	(M1)	(M2)	(M3)	(M4)	(M5)
Estimates	-0.039***	-0.042***	-0.043***	-0.046***	-0.040***	-0.047***
	(0.007)	(0.008)	(0.007)	(0.007)	(0.010)	(0.015)
User Embeddings		\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
\bar{r}_i^j			\checkmark	\checkmark	\checkmark	\checkmark
N_i^j				\checkmark	\checkmark	\checkmark
\bar{s}_i^j					\checkmark	\checkmark
Share ^j						\checkmark
N	63849	59937	59914	56098	40395	18218
Mean Dep. Var.	3.919	3.919	3.919	3.919	3.921	3.921

Table 4. Matching Users Before and After the AMPAS Nominations: Nominated Movies

Note: The sample includes MovieLens ratings from 1995 to 2019, posted within an 80-day window around the AMPAS nomination date, excluding ratings posted after the AMPAS award ceremony. The table displays estimated treatment effects for nominated movies. The first column is the OLS estimate from a regression of ratings on movie fixed effects and a dummy equal to one if a rating has been given after the nomination date. Columns M1 to M5 report ATT estimates from a propensity score matching with user embeddings (50 dimensions), matching exactly on the movie identifier as well as quintiles of different user characteristics. All characteristics are time-varying, based on the set of movies rated by the user prior to the rating of the nominated movie. Standard errors are obtained via bootstrap (100 replications). * p < 0.10, ** p < 0.05, *** p < 0.01.

	(OLS)	(M1)	(M2)	(M3)	(M4)	(M5)
Estimates	0.002	0.002	-0.003	-0.003	-0.003	-0.001
	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.005)
User Embeddings		\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
\bar{r}_i^j			\checkmark	\checkmark	\checkmark	\checkmark
N_i^j				\checkmark	\checkmark	\checkmark
\bar{s}_i^j					\checkmark	\checkmark
$Share_{i}^{j}$						\checkmark
N	63849	59937	59914	56098	40395	18218
Mean Dep. Var.	3.919	3.919	3.919	3.919	3.921	3.921

 Table 5. Matching Users Before and After the AMPAS Nominations: Predicted Ratings

Note: The sample includes predicted ratings for movies rated from 1995 to 2019, posted within an 80-day window around the AMPAS nomination date, excluding ratings posted after the AMPAS award ceremony. The table displays estimated treatment effects for nominated movies. The first column is the OLS estimate from a regression of predicted ratings on mvie fixed effects and a dummy equal to one if a predicted rating is related to a rating posted after the nomination date. Columns M1 to M5 report ATT estimates from a propensity score matching with user embeddings (50 dimensions), matching exactly on the movie identifier as well as quintiles of different user characteristics. All characteristics are time-varying, based on the set of movies rated by the user prior to the rating of the nominated movie. Standard errors are obtained via bootstrap (100 replications). * p < 0.10, *** p < 0.05, *** p < 0.01.

In Table 5, we replicate the analysis from Table 4 using predicted ratings as the outcome variable. As expected, the predicted ratings remain stable over time, indicating that users with similar preferences (as captured by embeddings and other observable characteristics) have consistent predicted ratings before and after the reference nomination date.

Additionally, in Appendix Table C8, we repeat the analysis for non-nominated movies and find no significant effects, except for one specification with a negative effect statistically significant at the 10 percent level. For these movies, ratings remain stable over time, and users with similar preferences provide consistent ratings before and after the reference nomination date.

Are some movies more affected by nomination-induced disappointment? And is the effect particularly strong for certain users? We address these questions in Appendix Tables C9, C10, C11, and C12, where we replicate the matching specification from Column (M5) in Table 4 across different subsets of movies and users. Appendix Table C9 shows that the disappointment effect is primarily driven by movies nominated for more than one AMPAS award. This finding aligns with the heterogeneity analysis discussed in the previous Section, suggesting that the disappointment effect is stronger when the signal is more likely to significantly influence users' expectations, that is, when movies receive multiple nominations. The negative effect is significant regardless of whether the movie ultimately wins an AMPAS award, but it appears to be larger for movies that are later awarded. This could be explained by the fact that movies that win awards tend to receive more nominations, and the importance of these nominations is particularly pronounced for such films. The heightened expectations created by multiple nominations, and the hype for the eventual awarding of some of these films, may amplify users' disappointment when their post-nomination viewing experience does not meet these elevated expectations.

The quality signal provided by AMPAS nominations and awards has evolved over time. In particular, various sources have noted a decline in the cultural relevance and public interest in these awards.¹³ In Appendix Table C10, we examine the heterogeneity of the disappointment effect across different years. Interestingly, we observe that the disappointment effect closely mirrors the relevance of the AMPAS awards, with a strong effect from 1995 to 2004, reflecting a high degree of disappointment (and thus a stronger influence of the Oscars on user expectations). For movies nominated between 2004 and 2011, the effect is smaller but still significant, whereas there is no significant effect for more recent movies from 2011 to 2019. This heterogeneity may be partly explained by the decreasing cultural relevance of the awards over time. However, it is also worth noting that the dates for nominations have shifted later in the year after the

¹³See: https://bunewsservice.com/the-death-of-awards-shows-do-people-no-longer-care/ and https:// movieweb.com/the-oscars-changed-lost-viewership/.

2010s, which could mean that users are less likely to be surprised by nominations and, consequently, less susceptible to disappointment. Finally, related to our next heterogeneity analysis, users who watch movies from different decades are not the same. Users on the MovieLens platform at its early stages may have been less experienced and more prone to disappointment, which could have further amplified the effect during the earlier years of the platform's development.

We investigate user heterogeneity in Appendix Tables C11 and C12, where we show that the effect is primarily driven by users who have posted fewer than 400 ratings before rating the nominated movie, and by those who are less stringent (as defined in Section 2.3). In contrast, the average ratings users have posted and the share of movies they rated in the same year are not significant drivers of heterogeneity. These findings suggest that users' disappointment is especially pronounced for less experienced users, and by those who, on average, tend to rate movies more leniently. Previous research has shown that these characteristics are typically associated with less expert users (Bondi et al., 2024), who are also more likely to rely on the AMPAS award, one of the most famous awards, as a guide to discover and watch movies.

We conclude this Section by offering some insights into the dynamics of disappointment, using the recommendations produced by the recommender system. Specifically, we examine the difference between the actual rating posted by user *i* for movie *j*, r_{ij} , and the predicted rating from the recommender system, \hat{r}_{ij} . We then explore how this difference evolves over time with a lead-lag model. In this model, the difference $r_{ij} - \hat{r}_{ij}$ is regressed on movie fixed effects, and a set of dummy variables for each 5-day window around the AMPAS nomination dates:

$$r_{ij} - \hat{r}_{ij} = \gamma_j + \sum_{\tau = -40 \text{ days}}^{40 \text{ days}} \delta_{\tau} \mathbf{1}(t = \tau) + \varepsilon_{ji}.$$
(8)

This allows us to observe how the gap between actual and predicted ratings changes relative to the nomination event. Figure 5 presents two event studies, tracking the differences before and after the nominations for both nominated and non-nominated movies. Nominated movies show a slight decline in the difference before the nominations, which could be due to minor changes in expectations as users begin to anticipate the potential for nominations. This is followed by a sharp drop starting 10 days after the nominations, which is consistent with the disappointment effect. The difference then increases again around 35 days after the nominations. This latter effect could be attributed to the fact that, in many years, the AMPAS awards occur after the 35-day mark. Since we remove ratings posted after the awards, the sample of movies used in the analysis differs for these later time points. In contrast, non-nominated movies exhibit a flatter trend, indicating that the nominations have no direct impact on their ratings.



(a) Nominated Movies

(b) Non-Nominated Movies



Notes: The sample includes MovieLens ratings from 1995 to 2019, posted within an 80-day window around the AMPAS nomination date, excluding ratings posted after the AMPAS award ceremony. MovieLens ratings are regressed on movie fixed effects and a set of five-day interval dummies centered on the AMPAS nomination date for nominated and non-nominated movies. The graphs display the estimated coefficients of these dummies, with the coefficient for the five-day interval of the nomination date normalized to zero. Standard errors, clustered by movie, are reported at the 10% level.

In Appendix Table C13, we extend this analysis by implementing a DiD design, so comparing again nominated and non-nominated movies, leveraging information from predicted ratings and embeddings. In all specifications, we include only users who provided at least one rating 40 days prior to the nominations (a necessary condition for calculating predicted ratings and embeddings). In Column (1), we report the same DiD specification as in Column (4) of Table 3, confirming that, even with a different sample, the results remain consistent and exhibit a very similar magnitude. In Column (2), we replicate the same DiD design but use predicted ratings, \hat{r}_{ij} , as the dependent variable. The results indicate that nominations have no significant effect on predicted ratings. This is expected, as the recommendation model does not incorporate ratings prior to the nominations. These findings suggest that, in the absence of a nomination, the ratings of nominated movies would not experience a post-nomination decline. Columns (3) and (4) present variations of the specification in Column (4) of Table 3, but with additional controls: predicted ratings in Column (3) and the 50 embeddings calculated by the recommendation system in Column (4). These serve as alternative methods to address potential selection concerns, and the results consistently show a decline in ratings with a similar magnitude. Finally, in Column (5), we apply the DiD specification to the difference between the actual rating and the predicted rating, $r_{ij} - \hat{r}_{ij}$, once again confirming the same negative effect.

This evidence further suggests that nominations lower the ratings of nominated movies. However, as in the previous DiD specification, this analysis compares nominated and non-nominated movies (treating the nominated movies as the treated units), and the precision of predictions may vary across movies: nominated movies may have more ratings and thus more precise predictions. For this reason, we still prefer the match-

ing approach presented in Table 4, which compares treated and non-treated units as users with similar tastes and preferences who watched the nominated movies, either before or after the nominations.

6 External Validation: IMDb Data

In the previous Sections, we leveraged the features of the dataset provided by MovieLens to investigate the role of disappointment in user ratings. The platform's specific characteristics and the comprehensive dataset allowed us to credibly isolate the disappointment effect. However, in many other contexts, disappointment effects likely exist but are far more difficult to measure reliably. To assess the external validity of our findings, we extend our analysis to IMDb, the largest online repository of user-generated movie content. As of November 2024, IMDb is the third most-visited website in the U.S. in the streaming and online TV category.¹⁴ To replicate the approach used on MovieLens, we scraped IMDb for all ratings posted for the nominated movies analyzed in previous sections. To gather additional user-related information, we also collected all ratings from users who rated these nominated movies. IMDb ratings range from 1 to 10, so we rescaled them to a 0.5–5-star scale to ensure comparability with MovieLens. Despite the absence of predicted ratings and user embeddings from a recommendation system, we employed a matching design based on observable user characteristics, mirroring our approach with MovieLens. Specifically, we included the average rating posted by each user, the number of movies they rated before rating the nominated movies, their stringency levels (calculated as in Section 2.3 but limited to the set of ratings and users scraped from IMDb¹⁵), and the share of user ratings for movies released in the same year.

Table 6 presents the OLS and matching estimates derived from IMDb data. Column 1 (OLS) includes movie fixed effects, comparing ratings of users who watched the same nominated movies without accounting for user characteristics. Columns 2–5 (M1–M4) provide matching estimates based on exact matching on movie identifiers and quintiles of the aforementioned user attributes, following a methodology similar to Table 4. Our findings confirm that IMDb users also rate nominated movies more negatively after the nominations. Notably, the magnitude of this disappointment effect is approximately three times larger than what we observed on MovieLens. The larger effect size on IMDb may arise from platform-specific factors. For instance, the public nature of IMDb reviews likely amplifies societal influences, such as groupthink, intensifying disappointment among users (a reverse "reputation effect" as in Li et al., 2024). Additionally, IMDb attracts a broader and more diverse user base, which includes a higher proportion of inexperienced

¹⁴For more details about IMDb, see: https://www.similarweb.com/website/imdb.com/.

¹⁵On MovieLens, we also have data on users that have never rated nominated movies, allowing us to measure how strict each user was compared to all other users in the sample. On IMDb, this measure is restricted to the subset of users who rated nominated movies.

	(OLS)	(M1)	(M2)	(M3)	(M4)
Estimates	-0.149***	-0.121***	-0.127***	-0.151***	-0.154***
	(0.012)	(0.015)	(0.017)	(0.023)	(0.025)
$ar{r}_i^j$		\checkmark	\checkmark	\checkmark	\checkmark
N_i^j			\checkmark	\checkmark	\checkmark
\bar{s}_i^j				\checkmark	\checkmark
$Share_i^j$					\checkmark
Ν	54838	36398	32232	21752	16129
Mean Dep. Var.	3.625	3.682	3.682	3.690	3.690

Table 6. Matching Users Before and After the AMPAS Nominations: IMDb Ratings

Note: The sample includes IMDb ratings posted within an 80-day window around the AMPAS nomination date, excluding ratings posted after the AMPAS award ceremony. The table displays estimated treatment effects for nominated movies. The first column is the OLS estimate from a regression of ratings on movie fixed effects and a dummy equal to one if a rating has been given after the nomination date. Columns M1 to M4 report ATT estimates matching exactly on the movie identifier as well as quintiles of different user characteristics. All characteristics are time-varying, based on the set of movies rated by the user prior to the rating of the nominated movie. Standard errors are obtained via bootstrap (100 replications). * p < 0.10, *** p < 0.05, *** p < 0.01.

users who may be more susceptible to heightened expectations set by nominations.

To further explore the robustness of our findings, we conducted a heterogeneity analysis for IMDb data, similar to the approach used for MovieLens. The results reveal that the patterns of heterogeneity on IMDb closely mimic those observed on MovieLens, albeit with a larger magnitude of the effect. Specifically, in Appendix Table D1, we show that the disappointment effect is more pronounced for movies that receive more than one nomination and for those that eventually win an AMPAS award. These findings suggest that the salience of multiple nominations or an eventual win may elevate expectations further, intensifying the subsequent disappointment among viewers. Interestingly, the results in Appendix Table D2 also highlight that the effect is largely driven by movies produced between 1995 and 2004, consistent with our findings on MovieLens. However, the disappointment effect remains statistically significant even for the most recent years on IMDb. This suggests that while the awards' influence may have faded, their capacity to shape consumer expectations, and trigger disappointment, has not disappeared entirely. Finally, we explored the role of user experience in driving the differences between IMDb and MovieLens results. Appendix Table D3 shows that IMDb users with fewer than five ratings prior to rating a nominated movie are the most affected by the disappointment effect. As the number of prior reviews increases, the magnitude of the effect decreases, pointing to the critical role of inexperience in amplifying disappointment. This pattern aligns with our findings on MovieLens, where more casual users (posting fewer ratings) also exhibit stronger disappointment. The higher prevalence of inexperienced users on IMDb, compared to the more expert reviewers on MovieLens, likely explains at least part of the larger effect size observed on IMDb.

7 Managerial Implications

Our findings carry managerial implications for the institutions awarding quality disclosures, for the firms that receive them and for the designers of rating and recommendation systems.

We highlight a critical dynamic in the interaction between quality certifications and consumer satisfaction, isolating a "disappointment effect" that can undermine the positive impact of certifications on consumer welfare. This has implications for industries that heavily rely on certifications and awards to signal quality, particularly in markets with experience goods such as movies, restaurants, and hospitality services. Examples include the Academy Awards in the film industry, Michelin stars in dining, and certification programs for sustainable products. Firms and companies operating in these sectors should be mindful of managing consumer expectations throughout the life cycle of their products (De Langhe et al., 2016; Carnehl, Stenzel and Schmidt, 2024). Properly aligning pre-purchase signals with the actual consumer experience can help mitigate the risk of elevated expectations leading to disappointment.

Entities that design certifications or awards should consider mechanisms to mitigate disappointment effects. For example, more transparent communication about what the certification represents or better consumer education could help calibrate expectations. Similarly, regulatory bodies overseeing certification schemes should weigh the trade-offs between bringing transparency to a market and sustaining consumer satisfaction.

Finally, online platforms and rating system organizers must recognize how disappointment can skew user reviews, reducing their reliability as unbiased quality signals. Platforms could implement algorithms or review mechanisms to better adjust for expectation-driven biases, ensuring ratings reflect actual quality more accurately. By acknowledging and addressing the disappointment effect, all stakeholders in the certification ecosystem can maximize the value of quality signals for both consumers and producers.

8 Conclusion

In markets where consumers are imperfectly informed about product quality, third-party quality disclosures, such as certifications and awards, play a crucial role in signaling quality and driving consumer demand (Chen and Xie, 2005). However, our analysis reveals that such disclosures may also have unintended consequences, particularly when they lead to inflated consumer expectations, resulting in disappointment.

Our findings show that Academy Award nominations generate a short-run disappointment effect, evidenced by a significant drop in ratings after the nomination announcement. This suggests that even when product quality remains constant, consumer satisfaction can decline due to heightened expectations. These results contribute to the broader literature on how quality disclosures impact consumer satisfaction and illustrate the complexity of managing consumer expectations (Li et al., 2024; Bondi et al., 2024).

Moreover, our study contributes to the growing body of work on reference-dependent preferences. By isolating the disappointment effect from user selection and strategic seller responses, we highlight how reference points interact with quality signals to influence consumer experiences. The implications extend beyond the movie industry, offering valuable insights for any context where certifications are used to signal quality.

Our matching approach, leveraging embeddings as multidimensional representations, can be applied in other settings with sparse and high-dimensional data. For example, in industrial organization and marketing, it could be applied using Nielsen household data; in labor economics, it could be used with high-dimensional employee mobility patterns; and with credit-card transaction data, it could be applied to fraud detection or a wide variety of policy evaluation questions.

Finally, our work also informs the literature on online reviewing behavior, where external factors such as quality disclosures can significantly influence user ratings. Previous studies have explored various biases in online reviews, including manipulation and social influences (Luca and Zervas, 2016; Moe and Schweidel, 2012; He et al., 2022). Our results suggest that quality disclosures not only affect consumer choices but also reshape the way users evaluate products post-purchase. This provides new insights into the dynamic relationship between expectations and consumer satisfaction.

By understanding and addressing the dynamics of disappointment, stakeholders in certification ecosystems, including firms, certification bodies, and online platforms, can refine their strategies to maximize the effectiveness of quality signals while preserving consumer welfare.

References

- Adomavicius, Gediminas, and Alexander Tuzhilin (2005) 'Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions.' *IEEE transactions on knowledge and data engineering* 17(6), 734–749
- Aridor, Guy, Duarte Gonçalves, Daniel Kluver, Ruoyan Kong, and Joseph Konstan (2022) 'The economics of recommender systems: Evidence from a field experiment on movielens.' *arXiv preprint arXiv:2211.14219*

- Athey, Susan, and Guido W Imbens (2019) 'Machine learning methods that economists should know about.' Annual Review of Economics 11(1), 685–725
- Backus, Matthew, Thomas Blake, Dimitriy Masterov, and Steven Tadelis (2022) 'Expectation, disappointment, and exit: evidence on reference point formation from an online marketplace.' *Journal of the European Economic Association* 20(1), 116–149
- Barberis, Nicholas C (2013) 'Thirty years of prospect theory in economics: A review and assessment.' Journal of economic perspectives 27(1), 173–196
- Billsus, Daniel, Michael J Pazzani et al. (1998) 'Learning collaborative information filters.' In 'Icml,' vol. 98 pp. 46–54
- Bondi, Tommaso (2023) 'Alone, together: A model of social (mis) learning from consumer reviews.' Available at SSRN 4453685
- Bondi, Tommaso, Michelangelo Rossi, and Ryan Stevens (2024) 'The good, the bad and the picky: Consumer heterogeneity and the reversal of product ratings.' *Management Science*
- Carnehl, Christoph, André Stenzel, and Peter Schmidt (2024) 'Pricing for the stars: Dynamic pricing in the presence of rating systems.' *Management Science* 70(3), 1755–1772
- 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), 1358–1398
- Chen, Yubo, and Jinhong Xie (2005) 'Third-party product review and firm marketing strategy.' *Marketing science* 24(2), 218–240
- De Langhe, Bart, Philip M Fernbach, and Donald R Lichtenstein (2016) 'Navigating by the stars: Investigating the actual and perceived validity of online user ratings.' *Journal of Consumer Research* 42(6), 817–833
- Dell, Melissa (2024) 'Deep learning for economists.' Technical Report, National Bureau of Economic Research
- Dellarocas, Chrysanthos, and Ritu Narayan (2006) 'A statistical measure of a population's propensity to engage in post-purchase online word-of-mouth'

- Dranove, David, and Ginger Zhe Jin (2010) 'Quality Disclosure and Certification: Theory and Practice.' Journal of Economic Literature 48(4), 935–63
- 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, and Georgios Zervas (2022) 'Consumer reviews and regulation: Evidence from nyc restaurants.' Technical Report, National Bureau of Economic Research
- Farronato, Chiara, Andrey Fradkin, Bradley J Larsen, and Erik Brynjolfsson (2024) 'Consumer protection in an online world: An analysis of occupational licensing.' *American Economic Journal: Applied Economics* 16(3), 549–579
- Fradkin, Andrey, Elena Grewal, and David Holtz (2021) 'Reciprocity and unveiling in two-sided reputation systems: Evidence from an experiment on airbnb.' *Marketing Science* 40(6), 1013–1029
- Friberg, Richard, and Erik Grönqvist (2012) 'Do expert reviews affect the demand for wine?' American Economic Journal: Applied Economics 4(1), 193–211
- Gergaud, Olivier, Karl Storchmann, and Vincenzo Verardi (2015) 'Expert opinion and product quality: Evidence from new york city restaurants.' *Economic Inquiry* 53(2), 812–835
- Goldberg, David, David Nichols, Brian M. Oki, and Douglas Terry (1992) 'Using collaborative filtering to weave an information tapestry.' *Commun. ACM* 35(12), 61–70
- 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
- He, Sherry, Brett Hollenbeck, and Davide Proserpio (2022) 'The market for fake reviews.' *Marketing Science* 41(5), 896–921
- Hilger, James, Greg Rafert, and Sofia Villas-Boas (2011) 'Expert opinion and the demand for experience goods: an experimental approach in the retail wine market.' *Review of Economics and Statistics* 93(4), 1289–1296
- Hu, Nan, Paul A Pavlou, and Jie Jennifer Zhang (2009) 'Why do online product reviews have a j-shaped distribution? overcoming biases in online word-of-mouth communication.' *Communications of the ACM* 52(10), 144–147

- Hu, Nan, Paul A Pavlou, and Jie Zhang (2017) 'On self-selection biases in online product reviews.' *MIS* quarterly 41(2), 449–475
- Hui, Xiang, Maryam Saeedi, Zeqian Shen, and Neel Sundaresan (2016) 'Reputation and regulations: Evidence from ebay.' *Management Science* 62(12), 3604–3616
- Koren, Yehuda, Robert Bell, and Chris Volinsky (2009) 'Matrix factorization techniques for recommender systems.' *Computer* 42(8), 30–37
- Kőszegi, Botond, and Matthew Rabin (2006) 'A model of reference-dependent preferences.' *The Quarterly Journal of Economics* 121(4), 1133–1165
- Kovács, Balázs, and Amanda J Sharkey (2014) 'The paradox of publicity: How awards can negatively affect the evaluation of quality.' *Administrative science quarterly* 59(1), 1–33
- Lagios, Nicolas, and Pierre-Guillaume Méon (2023) 'A matter of taste: The negative welfare effect of expert judgments.' *Working Papers CEB*
- 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
- Leibenstein, Harvey (1950) 'Bandwagon, snob, and veblen effects in the theory of consumers' demand.' *The quarterly journal of economics* 64(2), 183–207
- Li, Xingyi, Yiting Deng, Puneet Manchanda, and Bert De Reyck (2024) 'Can lower expert opinions lead to better consumer ratings?: The case of michelin stars.' *Available at SSRN 4275944*
- Li, Xinxin, and Lorin M Hitt (2008) 'Self-selection and information role of online product reviews.' *Information systems research* 19(4), 456–474
- Luca, Michael, and Georgios Zervas (2016) 'Fake it till you make it: Reputation, competition, and yelp review fraud.' *Management science* 62(12), 3412–3427
- McKenzie, Jordi (2023) 'The economics of movies (revisited): A survey of recent literature.' *Journal of Economic Surveys* 37(2), 480–525
- Moe, Wendy W, and David A Schweidel (2012) 'Online product opinions: Incidence, evaluation, and evolution.' *Marketing Science* 31(3), 372–386

- Nosko, Chris, and Steven Tadelis (2015) 'The limits of reputation in platform markets: An empirical analysis and field experiment.' Technical Report, National Bureau of Economic Research
- Paterek, Arkadiusz (2007) 'Improving regularized singular value decomposition for collaborative filtering.' *Proc. KDD Cup and Workshop* ACM Press, 39–42
- 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, 371–407
- Rey-Biel, Pedro, Ayelet Gneezy, Uri Gneezy, Dominique Lauga, and Joan Llull (2024) 'Expectations, satisfaction and utility from experience goods.' *Manuscript*
- Sarwar, Badrul, George Karypis, Joseph Konstan, and John Riedl (2000) 'Analysis of recommendation algorithms for e-commerce.' In 'Proceedings of the 2nd ACM Conference on Electronic Commerce' pp. 158–167
- Schoenmueller, Verena, Oded Netzer, and Florian Stahl (2020) 'The polarity of online reviews: Prevalence, drivers and implications.' *Journal of Marketing Research* 57(5), 853–877
- Sunder, Sarang, Kihyun Hannah Kim, and Eric A Yorkston (2019) 'What drives herding behavior in online ratings? the role of rater experience, product portfolio, and diverging opinions.' *Journal of Marketing* 83(6), 93–112
- Wang, Chong, Xiaoquan Zhang, and Il-Horn Hann (2018) 'Socially nudged: A quasi-experimental study of friends' social influence in online product ratings.' *Information Systems Research* 29(3), 641–655

APPENDIX - Quality Disclosures and Disappointment: Evidence from the Academy Awards

A APPENDIX - Empirical Setting and Dataset

	Mean	SD	Min	Max
Average Ratings				
All	3.212	.457	.8548	4.483
Before Nom.	3.182	.6037	.5	5
After Nom.	3.213	.4611	.8929	4.48
Number of Ratings				
All	1221	3441.41	30	72674
Before Nom.	127.2	504.3	1	1.5e+04
After Nom.	1085	3089.01	4	68582
Average Ratings (80-days window)				
All	3.205	.6629	.5	5
Before Nom.	3.197	.7444	.5	5
After Nom.	3.221	.7362	.5	5
Number of Ratings (80-days window)				
All	43.45	116.9	0	2827
Before Nom.	23.55	63.62	0	1454
After Nom.	20.31	55.61	0	1373

Table A1. Summary Statistics: Non-Nominated Movies

Note: The table presents descriptive statistics on all 9,397 movies, non-nominated for AMPAS awards, included in our sample. We report the average ratings and the number of ratings both before and after the nominations. The first two panels include all available ratings without any time restrictions, while the second two panels focus on ratings within a 80-day window surrounding the nominations.



Figure A1. Arrival of Ratings for Nominated and Non-Nominated Movies

Notes: 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.

	Mean	SD	Min	Max
Average User Ratings				
All	3.456	.1187	2.956	3.946
Before Nom.	3.411	.2138	1.038	5
After Nom.	3.459	.1255	2.65	3.955
Number of User Ratings				
All	758.5	395.881	62.09	2927.1
Before Nom.	813.5	636.7	1	1.8e+04
After Nom.	766.6	431.84	62.17	5602
Average User Stringency				
All	.0794	.0996	5153	.6804
Before Nom.	.1101	.2406	-2.243	2.2
After Nom.	.0768	.1102	4975	1.149
Average Share of User Ratings for Current-Year Releases				
All	.041	.0268	.0055	.427
Before Nom.	.1072	.1144	0	.8587
After Nom.	.03	.0132	.0047	.2183

Table A2. User Selection Before and After the Nominations: All Ratings

Note: The table presents descriptive statistics of users who rated movies nominated for AMPAS Awards included in our sample, with no time restrictions. We report the average user rating, the average number of ratings per user, the average rating stringency (a measure of how strictly users rate, based on how their ratings compare to the average rating of the movie at the time), and the average share of ratings for movies produced in the same year as the rating. The stringency of a user is defined as a moving average of the stringency of their past five ratings. These statistics are presented separately for users who rated the movies before and after the nominations.

B APPENDIX - Comparing Ratings for Nominated and Non-Nominated Movies

Table B1. Difference-in-Differences: Removing Days Right After the Nominations

	(All)	(5 Days After)	(10 Days After)	(15 Days After)
Nominated $_{i} \times post_{t}^{Nom}$	-0.018**	-0.019**	-0.024***	-0.033***
U I	(0.007)	(0.008)	(0.008)	(0.009)
Movie FEs	\checkmark	\checkmark	\checkmark	\checkmark
Date FEs	\checkmark	\checkmark	\checkmark	\checkmark
R^2	0.21	0.21	0.21	0.21
Ν	497,266	469,327	437,886	405,986
Mean Dep. Var.	3.503	3.502	3.503	3.502

	(All)	$(n_i^{critics} > 0)$	$(\bar{r}_i^{critics} > 50)$	$(\bar{r}_i^{IMDb} > 5)$
Nominated $_j \times post_t^{Nom}$	-0.018**	-0.015**	-0.013*	-0.016**
	(0.007)	(0.007)	(0.007)	(0.007)
Movie FEs	\checkmark	\checkmark	\checkmark	\checkmark
Date FEs	\checkmark	\checkmark	\checkmark	\checkmark
R^2	0.21	0.20	0.16	0.18
Ν	497,266	467,818	397,498	484,683
Mean Dep. Var.	3.503	3.511	3.628	3.532

Table B2. Difference-in-Differences: Varying Non-Nominated Movies in the Control Group

	(1)	(2)	(3)	(4)	(5)	(6)
$Nominated_j^{PA.=0} \times post_t^{Nom}$	0.020					
$Nominated_j^{PA.>0} \times post_t^{Nom}$	(0.023)	-0.021***				
$Nominated_j^{=1} \times post_t^{Nom}$		(0.007)	-0.012			
$Nominated_j^{>1} \times post_t^{Nom}$			(0.011)	-0.021**		
$Nominated_j^{award=0} \times post_t^{Nom}$				(0.000)	-0.018** (0.009)	
$Nominated_j^{award>0} \times post_t^{Nom}$					· · · ·	-0.018* (0.011)
Movie FEs	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
Date FEs	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
R^2	0.18	0.21	0.19	0.21	0.19	0.20
Ν	398,139	490,303	429,955	458,487	455,277	433,165
Mean Dep. Var.	3.389	3.497	3.421	3.475	3.451	3.446

Table B3. Difference-in-Differences: Nominated Movie Heterogeneity

Note: The sample includes MovieLens ratings from 1995 to 2019, posted within an 80-day window around the AMPAS nomination date, excluding ratings posted after the AMPAS award ceremony. Standard errors, clustered by movie, are shown in parentheses. * p < 0.10, *** p < 0.05, *** p < 0.01.

	(1)	(2)	(3)	(4)	(5)
Nominated $_j \times post_t^{Nom}$	-0.018**	-0.018**	-0.016**	-0.017**	-0.027***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
\bar{r}_i^j	0.689***	0.693***	0.244***	0.244***	-0.156***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.015)
N_i^j		0.000***	-0.000***	-0.000***	-0.000***
		(0.000)	(0.000)	(0.000)	(0.000)
\bar{s}_i^j			-0.520***	-0.521***	-0.201***
			(0.005)	(0.005)	(0.006)
Share ^j				-0.189***	-0.258***
				(0.022)	(0.049)
Movie FEs	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
Date FEs	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
User FEs					\checkmark
R^2	0.30	0.30	0.33	0.33	0.43
Ν	496,044	496,044	493,817	493,817	487,847
Mean Dep. Var.	3.502	3.502	3.503	3.503	3.498

Table B4. Difference-in-Differences: Controlling for User Selection

C APPENDIX - A Recommendation-based Matching Design

C.1 The Hyperparameter Choice for the Recommender Algorithm and its Efficiency

The calculation of embedding values for the matching steps involves several degrees of freedom, with four key hyperparameters to configure: the number of embedding dimensions, the number of epochs, the learning rate, and the regularization rate. The number of embedding dimensions determines the size of the vector that captures users' preferences. The number of epochs refers to the number of iterations the optimization routine performs to refine the embedding values. The learning rate controls the magnitude of updates the model applies during each epoch, influencing the speed of convergence. Finally, the regularization rate imposes a penalty on large embedding values to mitigate overfitting.

Appendix Table C1 indicate that a learning rate of 0.005 and a regularization rate of 0.02 yield the best performance. Comparing the two panels, increasing the number of epochs from 10 to 20 offers some improvement in performance, but the root mean squared error is less sensitive to this change than to adjustments in the learning rate and regularization rate. Regarding the number of embedding dimensions, the performance gains from increasing from 50 to 100 or 200 dimensions are minimal. Based on these findings, we use the hyperparameters LR = 0.005, Reg = 0.02, No. epochs = 20, and No. embeddings used in all matching estimations. Notably, the values LR = 0.005, Reg = 0.02, and No. epochs = 20 are also the default settings of the scikit-surprise library, which we used to compute the embeddings. However, the default setting for the number of embedding dimensions in scikit-surprise is 100.

No. of Epochs = 10				
No. of Factors	(LR=0.002, RR=0.02)	(LR=0.002, RR=0.01)	(LR=0.005, RR=0.02)	(LR=0.005, RR=0.01)
10	.8586	.8689	.8185	.864
20	.8561	.8688	.8128	.8623
50	.8513	.8688	.8074	.8598
100	.8487	.8686	.8035	.8582
200	.8457	.8684	.801	.8566
No. of Epochs = 20				
No. of Factors	(LR=0.002, RR=0.02)	(LR=0.002, RR=0.01)	(LR=0.005, RR=0.02)	(LR=0.005, RR=0.01)
10	.8268	.8644	.7903	.847
20	.8226	.8635	.7836	.8461
50	.8162	.8617	.7789	.8449
100	.8118	.86	.7779	.8439
200	.8088	.8578	.7781	.8427

Table C1. Recommender Algorithm Performance with Different Specifications

To confirm the robustness of our matching procedure to variations in the number of embedding dimensions, Tables C2,C3,C4 and C5 present the matching results and placebo results using embeddings with 10, 20, 100, and 200 dimensions. The findings remain robust overall, with the estimated treatment effect ranging between -0.034 and -0.04, while the treatment effect in the placebo sample remains insignificant across all specifications.

C.2 Recommendation-based Matching - Robustness

	(OLS)	(M1)	(M2)	(M3)	(M4)	(M5)
Estimates	-0.039***	-0.036***	-0.052***	-0.044***	-0.040***	-0.046***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.010)	(0.016)
User Embeddings		\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
\bar{r}_i^j			\checkmark	\checkmark	\checkmark	\checkmark
N_i^j				\checkmark	\checkmark	\checkmark
\bar{S}_{i}^{j}					\checkmark	\checkmark
Share ^j						\checkmark
Ň	63849	59937	59975	55965	40354	18237
Mean Dep. Var.	3.919	3.919	3.919	3.919	3.921	3.921

Table C2. Matching Users Before and After the AMPAS Nominations: embeddings with 10-dimensions

Note: The sample includes MovieLens ratings from 1995 to 2019, posted within an 80-day window around the AMPAS nomination date, excluding ratings posted after the AMPAS award ceremony. The table displays estimated treatment effects for nominated movies. The first column is the OLS estimate from a regression of ratings on mvie fixed effects and a dummy equal to one if a rating has been given after the nomination date. Columns M1 to M5 report ATT estimates from a multivariate-distance matching with user embeddings (10 dimensions), matching exactly on the movie identifier as well as quintiles of different user characteristics. All characteristics are time-varying, based on the set of movies rated by the user prior to the rating of the nominated movie. Standard errors are obtained via bootstrap (100 replications). * p < 0.10, ** p < 0.05, *** p < 0.01.

	(OLS)	(M1)	(M2)	(M3)	(M4)	(M5)
Estimates	-0.039***	-0.036***	-0.042***	-0.037***	-0.032***	-0.042***
	(0.007)	(0.009)	(0.007)	(0.008)	(0.010)	(0.016)
User Embeddings		\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
\bar{r}_i^j			\checkmark	\checkmark	\checkmark	\checkmark
N_i^j				\checkmark	\checkmark	\checkmark
\bar{s}_{i}^{j}					\checkmark	\checkmark
Share ^j						\checkmark
Ň	63849	60043	59974	56068	40266	18165
Mean Dep. Var.	3.919	3.919	3.919	3.919	3.921	3.921

Table C3. Matching Users Before and After the AMPAS Nominations: embeddings with 20-dimensions

Note: The sample includes MovieLens ratings from 1995 to 2019, posted within an 80-day window around the AMPAS nomination date, excluding ratings posted after the AMPAS award ceremony. The table displays estimated treatment effects for nominated movies. The first column is the OLS estimate from a regression of ratings on mvie fixed effects and a dummy equal to one if a rating has been given after the nomination date. Columns M1 to M5 report ATT estimates from a multivariate-distance matching with user embeddings (20 dimensions), matching exactly on the movie identifier as well as quintiles of different user characteristics. All characteristics are time-varying, based on the set of movies rated by the user prior to the rating of the nominated movie. Standard errors are obtained via bootstrap (100 replications). * p < 0.10, ** p < 0.05, *** p < 0.01.

	(OLS)	(M1)	(M2)	(M3)	(M4)	(M5)
Estimates	-0.039***	-0.040***	-0.042***	-0.046***	-0.040***	-0.048***
	(0.007)	(0.008)	(0.007)	(0.007)	(0.010)	(0.017)
User Embeddings		\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
\bar{r}_i^j			\checkmark	\checkmark	\checkmark	\checkmark
N_i^j				\checkmark	\checkmark	\checkmark
\bar{s}_{i}^{j}					\checkmark	\checkmark
$Share_i^j$						\checkmark
Ň	63849	60015	59979	56210	40384	18266
Mean Dep. Var.	3.919	3.919	3.919	3.919	3.921	3.921

Table C4. Matching Users Before and After the AMPAS Nominations: embeddings with 100-dimensions

Note: The sample includes MovieLens ratings from 1995 to 2019, posted within an 80-day window around the AMPAS nomination date, excluding ratings posted after the AMPAS award ceremony. The table displays estimated treatment effects for nominated movies. The first column is the OLS estimate from a regression of ratings on mvie fixed effects and a dummy equal to one if a rating has been given after the nomination date. Columns M1 to M5 report ATT estimates from a multivariate-distance matching with user embeddings (100 dimensions), matching exactly on the movie identifier as well as quintiles of different user characteristics. All characteristics are time-varying, based on the set of movies rated by the user prior to the rating of the nominated movie. Standard errors are obtained via bootstrap (100 replications). * p < 0.10, ** p < 0.05, *** p < 0.01.

	(OLS)	(M1)	(M2)	(M3)	(M4)	(M5)
Estimates	-0.039***	-0.038***	-0.039***	-0.043***	-0.039***	-0.048***
	(0.007)	(0.008)	(0.008)	(0.008)	(0.011)	(0.015)
User Embeddings		\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
\bar{r}_i^j			\checkmark	\checkmark	\checkmark	\checkmark
N_i^j				\checkmark	\checkmark	\checkmark
\bar{s}_{i}^{j}					\checkmark	\checkmark
$Share_i^j$						\checkmark
N	63849	59980	59859	56322	40338	18261
Mean Dep. Var.	3.919	3.919	3.919	3.919	3.921	3.921

Table C5. Matching Users Before and After the AMPAS Nominations: embeddings with 200-dimensions

Note: The sample includes MovieLens ratings from 1995 to 2019, posted within an 80-day window around the AMPAS nomination date, excluding ratings posted after the AMPAS award ceremony. The table displays estimated treatment effects for nominated movies. The first column is the OLS estimate from a regression of ratings on mvie fixed effects and a dummy equal to one if a rating has been given after the nomination date. Columns M1 to M5 report ATT estimates from a multivariate-distance matching with user embeddings (200 dimensions), matching exactly on the movie identifier as well as quintiles of different user characteristics. All characteristics are time-varying, based on the set of movies rated by the user prior to the rating of the nominated movie. Standard errors are obtained via bootstrap (100 replications). * p < 0.10, ** p < 0.05, *** p < 0.01.

	(OLS)	(M1)	(M2)	(M3)	(M4)	(M5)
Estimates	-0.039***	-0.061***	-0.047***	-0.049***	-0.041***	-0.047***
	(0.007)	(0.006)	(0.006)	(0.007)	(0.010)	(0.015)
User Embeddings		\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
\bar{r}_i^j			\checkmark	\checkmark	\checkmark	\checkmark
N_i^j				\checkmark	\checkmark	\checkmark
\bar{s}_i^{j}					\checkmark	\checkmark
$Share_{i}^{j}$						\checkmark
N	63849	63339	62996	58308	41361	18610
Mean Dep. Var.	3.919	3.919	3.919	3.919	3.921	3.921

Table C6. Matching Users Before and After the AMPAS Nominations: Ratings (multivariate-distance matching)

Note: The sample includes MovieLens ratings from 1995 to 2019, posted within an 80-day window around the AMPAS nomination date, excluding ratings posted after the AMPAS award ceremony. The table displays estimated treatment effects for nominated movies. The first column is the OLS estimate from a regression of ratings on mvie fixed effects and a dummy equal to one if a rating has been given after the nomination date. Columns M1 to M5 report ATT estimates from a multivariate-distance matching with user embeddings (50 dimensions), matching exactly on the movie identifier as well as quintiles of different user characteristics. All characteristics are time-varying, based on the set of movies rated by the user prior to the rating of the nominated movie. Standard errors are obtained via bootstrap (100 replications). * p < 0.10, ** p < 0.05, *** p < 0.01.

 Table C7. Matching Users Before and After the AMPAS Nominations: Removing Days Right After the Nominations

	(All)	(5 Days After)	(10 Days After)	(15 Days After)
Estimates	-0.047***	-0.040**	-0.042**	-0.042**
	(0.015)	(0.016)	(0.020)	(0.020)
User Embeddings	\checkmark	\checkmark	\checkmark	\checkmark
$ar{r}_i^j$	\checkmark	\checkmark	\checkmark	\checkmark
N_i^j	\checkmark	\checkmark	\checkmark	\checkmark
\bar{s}_i^j	\checkmark	\checkmark	\checkmark	\checkmark
$Share_i^j$	\checkmark	\checkmark	\checkmark	\checkmark
N	18218	16038	13733	10964
Mean Dep. Var.	3.921	3.922	3.921	3.921

Note: The sample includes MovieLens ratings from 1995 to 2019, posted within an 80-day window around the AMPAS nomination date, excluding ratings posted after the AMPAS award ceremony. The table displays estimated treatment effects for nominated movies. In all columns we report ATT estimates from a propensity score matching with user embeddings (50 dimensions), matching exactly on the movie identifier as well as quintiles of different user characteristics (as in Column (M5) in Table 4). Standard errors are obtained via bootstrap (100 replications). * p < 0.10, ** p < 0.05, *** p < 0.01.

	(OLS)	(M1)	(M2)	(M3)	(M4)	(M5)
Estimates	0.001	-0.006	0.001	-0.015	-0.031*	-0.023
	(0.007)	(0.009)	(0.008)	(0.011)	(0.017)	(0.030)
User Embeddings		\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
\bar{r}_i^j			\checkmark	\checkmark	\checkmark	\checkmark
N_i^j				\checkmark	\checkmark	\checkmark
\bar{s}_i^j					\checkmark	\checkmark
$Share_i^j$						\checkmark
N	99608	83266	68369	42141	19578	6331
Mean Dep. Var.	3.397	3.397	3.397	3.397	3.399	3.399

Table C8. Matching Users Before and After the AMPAS Nominations: Non-Nominated Movies

Note: The sample includes MovieLens ratings from 1995 to 2019, posted within an 80-day window around the AMPAS nomination date, excluding ratings posted after the AMPAS award ceremony. The table displays estimated treatment effects for non-nominated movies, excluding all users who have rated nominated movies. The first column is the OLS estimate from a regression of ratings on mvie fixed effects and a dummy equal to one if a rating has been given after the nomination date. Columns M1 to M5 report ATT estimates from a propensity score matching with user embeddings (50 dimensions), matching exactly on the movie identifier as well as quintiles of different user characteristics. All characteristics are time-varying, based on the set of movies rated by the user prior to the rating of the non-nominated movie. Standard errors are obtained via bootstrap (100 replications). * p < 0.10, ** p < 0.05, *** p < 0.01.

	(Nom = 1)	(<i>Nom</i> > 1)	(Award = 0)	(Award > 0)
Estimates	-0.049	-0.043**	-0.037*	-0.061***
	(0.038)	(0.020)	(0.021)	(0.021)
User Embeddings	\checkmark	\checkmark	\checkmark	\checkmark
\bar{r}_i^j	\checkmark	\checkmark	\checkmark	\checkmark
N_i^j	\checkmark	\checkmark	\checkmark	\checkmark
\bar{s}_i^j	\checkmark	\checkmark	\checkmark	\checkmark
$Share_i^j$	\checkmark	\checkmark	\checkmark	\checkmark
N	4964	13277	10049	8200
Mean Dep. Var.	3.795	3.993	3.847	4.037

Table C9. Matching Users Before and After the AMPAS Nominations: Number of Nominations and Awards

Note: The sample includes MovieLens ratings from 1995 to 2019, posted within an 80-day window around the AMPAS nomination date, excluding ratings posted after the AMPAS award ceremony. The table displays estimated treatment effects for nominated movies. In all columns we report ATT estimates from a propensity score matching with user embeddings (50 dimensions), matching exactly on the movie identifier as well as quintiles of different user characteristics (as in Column (M5) in Table 4). Standard errors are obtained via bootstrap (100 replications). * p < 0.10, ** p < 0.05, *** p < 0.01.

Table C10. Matching Users Before and After the AMPAS Nominations: Movie Production Year

(1995-2004)	(2004-2011)	(2011-2019)
-0.080***	-0.059**	0.004
(0.028)	(0.023)	(0.025)
\checkmark	\checkmark	\checkmark
6863	5574	5478
4.024	3.913	3.822
	(1995 - 2004) -0.080*** (0.028) \checkmark \checkmark \checkmark \checkmark \checkmark 6863 4.024	$\begin{array}{cccc} (1995-2004) & (2004-2011) \\ \hline -0.080^{***} & -0.059^{**} \\ (0.028) & (0.023) \\ \checkmark & \checkmark \\ 6863 & 5574 \\ 4.024 & 3.913 \end{array}$

Note: The sample includes MovieLens ratings from 1995 to 2019, posted within an 80-day window around the AMPAS nomination date, excluding ratings posted after the AMPAS award ceremony. The table displays estimated treatment effects for nominated movies. In all columns we report ATT estimates from a propensity score matching with user embeddings (50 dimensions), matching exactly on the movie identifier as well as quintiles of different user characteristics (as in Column (M5) in Table 4). Standard errors are obtained via bootstrap (100 replications). * p < 0.10, ** p < 0.05, *** p < 0.01.

	$(N_i^j < 400)$	$(N_i^j > 400)$	$(\bar{s}_i^j < 0)$	$(\bar{s}_i^j > 0)$
Estimates	-0.047**	-0.037	-0.063***	-0.035
	(0.022)	(0.025)	(0.020)	(0.026)
User Embeddings	\checkmark	\checkmark	\checkmark	\checkmark
\bar{r}_i^j	\checkmark	\checkmark	\checkmark	\checkmark
N_i^j	\checkmark	\checkmark	\checkmark	\checkmark
\bar{s}_i^j	\checkmark	\checkmark	\checkmark	\checkmark
$Share_i^j$	\checkmark	\checkmark	\checkmark	\checkmark
Ν	9225	8402	8827	8185
Mean Dep. Var.	3.996	3.859	4.129	3.729

Table C11. Matching Users Before and After the AMPAS Nominations: User Characteristics $(N_i^j \text{ and } \bar{s}_i^j)$

Note: The sample includes MovieLens ratings from 1995 to 2019, posted within an 80-day window around the AMPAS nomination date, excluding ratings posted after the AMPAS award ceremony. The table displays estimated treatment effects for nominated movies. In all columns we report ATT estimates from a propensity score matching with user embeddings (50 dimensions), matching exactly on the movie identifier as well as quintiles of different user characteristics (as in Column (M5) in Table 4). Standard errors are obtained via bootstrap (100 replications). * p < 0.10, ** p < 0.05, *** p < 0.01.

Table C12. Matching Users Before and After the AMPAS Nominations: User Characteristics $(\bar{r}_i^j \text{ and } Share_i^j)$

	$(\bar{r}_i^j < 3.5)$	$(\bar{r}_i^j > 3.5)$	$(Share_{i}^{j} < 0.05)$	$(Share_{i}^{j} > 0.05)$
Estimates	-0.051*	-0.049**	-0.046**	-0.047**
	(0.028)	(0.021)	(0.023)	(0.022)
User Embeddings	\checkmark	\checkmark	\checkmark	\checkmark
\bar{r}_i^j	\checkmark	\checkmark	\checkmark	\checkmark
N_i^j	\checkmark	\checkmark	\checkmark	\checkmark
\bar{s}_i^j	\checkmark	\checkmark	\checkmark	\checkmark
$Share_i^j$	\checkmark	\checkmark	\checkmark	\checkmark
Ν	6679	10353	8099	8807
Mean Dep. Var.	3.698	4.101	3.918	3.923

Note: The sample includes MovieLens ratings from 1995 to 2019, posted within an 80-day window around the AMPAS nomination date, excluding ratings posted after the AMPAS award ceremony. The table displays estimated treatment effects for nominated movies. In all columns we report ATT estimates from a propensity score matching with user embeddings (50 dimensions), matching exactly on the movie identifier as well as quintiles of different user characteristics (as in Column (M5) in Table 4). Standard errors are obtained via bootstrap (100 replications). * p < 0.10, ** p < 0.05, *** p < 0.01.

	(<i>r_{ij}</i>)	(\hat{r}_{ij})	(<i>r</i> _{<i>ij</i>})	(<i>r</i> _{<i>ij</i>})	$(r_{ij}-\hat{r}_{ij})$
Nominated $_j \times post_t^{Nom}$	-0.017*	-0.003	-0.014*	-0.017**	-0.014*
	(0.009)	(0.004)	(0.008)	(0.009)	(0.008)
Movie FEs	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
Date FEs	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
Predicted Ratings			\checkmark		
Embeddings				\checkmark	
R^2	0.25	0.50	0.39	0.25	0.12
Ν	254,482	254,482	254,482	254,482	254,482
Mean Dep. Var.	3.503	3.539	3.503	3.503	-0.036

Table C13. Difference-in-Differences: Exploiting Recommended Systems Predictions and Embeddings

Note: The sample includes MovieLens ratings from 1995 to 2019, posted within an 80-day window around the AMPAS nomination date, excluding ratings posted after the AMPAS award ceremony. In all specifications, we only include users with at least one rating 40 days before the nominations (a necessary condition to calculate predicted ratings and embeddings). Standard errors, clustered by movie, are shown in parentheses. * p < 0.10, *** p < 0.05, *** p < 0.01.

D APPENDIX - External Validity: IMDb Data

	(<i>Nom</i> = 1)	(<i>Nom</i> > 1)	(Award = 0)	(<i>Award</i> > 0)
Estimates	-0.122**	-0.162***	-0.084*	-0.218***
	(0.051)	(0.032)	(0.049)	(0.043)
$ar{r}_i^j$	\checkmark	\checkmark	\checkmark	\checkmark
N_i^j	\checkmark	\checkmark	\checkmark	\checkmark
\bar{s}_i^j	\checkmark	\checkmark	\checkmark	\checkmark
$Share_i^j$	\checkmark	\checkmark	\checkmark	\checkmark
Ν	3538	12591	7902	8227
Mean Dep. Var.	7.390	7.375	7.207	7.603

 Table D1. Matching Users Before and After the AMPAS Nominations (IMDb Ratings): Number of Nominations and Awards

Note: The sample includes IMDb ratings posted within an 80-day window around the AMPAS nomination date, excluding ratings posted after the AMPAS award ceremony. The table displays estimated treatment effects for nominated movies. In all columns we report ATT estimates matching exactly on the movie identifier as well as quintiles of different user characteristics (as in Column (M5) in Table 4). Standard errors are obtained via bootstrap (100 replications). * p < 0.10, ** p < 0.05, *** p < 0.01.

Table D2. Matching Users Before and After the AMPAS Nominations (IMDb Ratings): Movie Production

 Year

	(1995-2004)	(2004-2011)	(2011-2019)
Estimates	-0.249***	-0.149***	-0.128***
	(0.081)	(0.052)	(0.035)
\bar{r}_i^j	\checkmark	\checkmark	\checkmark
N_i^j	\checkmark	\checkmark	\checkmark
\bar{s}_i^j	\checkmark	\checkmark	\checkmark
$Share_i^j$	\checkmark	\checkmark	\checkmark
N	2537	4317	9048
Mean Dep. Var.	7.646	7.499	7.226

Note: The sample includes IMDb ratings posted within an 80-day window around the AM-PAS nomination date, excluding ratings posted after the AMPAS award ceremony. The table displays estimated treatment effects for nominated movies. In all columns we report ATT estimates matching exactly on the movie identifier as well as quintiles of different user characteristics (as in Column (M5) in Table 4). Standard errors are obtained via bootstrap (100 replications). * p < 0.10, ** p < 0.05, *** p < 0.01.

	$(N_i^j < 5)$	$(5 > N_i^j > 20)$	$(N_i^j > 20)$
Estimates	-0.187***	-0.154***	-0.097**
	(0.042)	(0.053)	(0.039)
$ar{r}_i^j$	\checkmark	\checkmark	\checkmark
N_i^j	\checkmark	\checkmark	\checkmark
\bar{s}_i^j	\checkmark	\checkmark	\checkmark
Share ^j	\checkmark	\checkmark	\checkmark
Ν	8031	3577	3797
Mean Dep. Var.	7.166	7.283	7.599

 Table D3.
 Matching Users Before and After the AMPAS Nominations (IMDb Ratings): User Number of Ratings

Note: The sample includes IMDb ratings posted within an 80-day window around the AMPAS nomination date, excluding ratings posted after the AMPAS award ceremony. The table displays estimated treatment effects for nominated movies. In all columns we report ATT estimates matching exactly on the movie identifier as well as quintiles of different user characteristics (as in Column (M5) in Table 4). Standard errors are obtained via bootstrap (100 replications). * p < 0.10, ** p < 0.05, *** p < 0.01.