

The Effect of Internet Reviews on the Demand of Experience Goods: Case of Korean Movie Industry

Yun, Yun, and Lee

SKKU Global Econ

November 2023

Introduction

- Standard consumer theory assume perfect information of consumer.
- However, for some goods, one cannot know their exact characteristic, and hence don't know their potential utility before consumption.
- These goods are **experience goods**. In this case, **consumers refer to other's consumption experience** such as reviews to avoid uncertainty.

Introduction

- Our question is how these reviews are related with experience goods demand.
- This give sense to demand side and expect to be theoretic background of experience goods market. This might give intuition for welfare aspect in future studies.
- Korean movie will be main target since it is matured market among experience goods market, and reviews are easily accessible to consumers.
- This research tries to catch how each components of review affects to movie demand using panel regression.

Literature

- The influence of expert reviews on consumer demand for experience goods: A case study of movie critics(David and Chirstopher, 2005)
 - Studied the effects of expert reviews on the movie demand using DID
 - Effect of expert reviews was small but still detectable
- The dynamics of online word-of-mouth and product sales-An empirical investigation of the movie industry(Wenjing et al, 2008)
 - Studied the effects of words of mouth(WOM) effect on the consumer's purchase decision

Speciality

- Use **mean** and **standard deviation** of reviews to catch two different information from reviews.
- Data collectable in **panel form** from “Korea Movie Council”, enabling to use **fixed effect** to control potential *unobserved effect*.
- Check which one is more important between *accumulated movie review* and *last day's movie review*.

Stylized Facts



영화정보 ⓘ

W 왓차

N 넷플릭스

T 티빙

W 웨이브

주요정보

출연/제작

영상/포토

평점

네티즌 평점 9.0점 (2686명)

MY

★★★★★ 0

로그인 해주세요.



추천댓글 ? 찬반순 최신순 과거순

★★★★★ 10

이 영화가 2007년에 나왔다니 지금 개봉해도 손색이 없겠다.

2019. 11. 7. 21:59

👍 5 🗨 0

★★★★★ 8

이게 시리즈중 제일 괜찮음

2021. 7. 8. 20:50

👍 4 🗨 0

★★★★★ 7

일본의 완구로봇을 모델로, 그 로봇만화시리즈를 바탕으로 영화로 만들어선지 기억나는 명대사가 있다
'북한것으로 보입니다'
(국방성 네트워크가 해킹당했을때)

실관람객 평점 ⓘ

8.60 / 10



10명 참여

남자

8.67

여자

8.5

점수별 비율 ⓘ

★ 9-10 60%

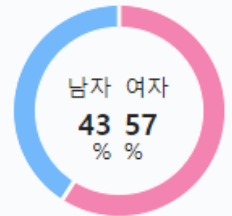
★ 7-8 20%

★ 5-6 20%

★ 3-4 0%

★ 1-2 0%

성별 비율 ⓘ



감상평

별점을 선택해주세요.



감상평을 작성해주세요.

등록

• 공감순 • 최신순

스포일러 포함 ☐

★★★★★ 9

관람객 트랜스포머4를 보고 1은 명작이었음을 다시 깨닫는다
yume**** · 2014.06.26. 03:02 · 신고

👍 446

🗨 21

★★★★★ 10

관람객 재밌게본 영화예요~^^

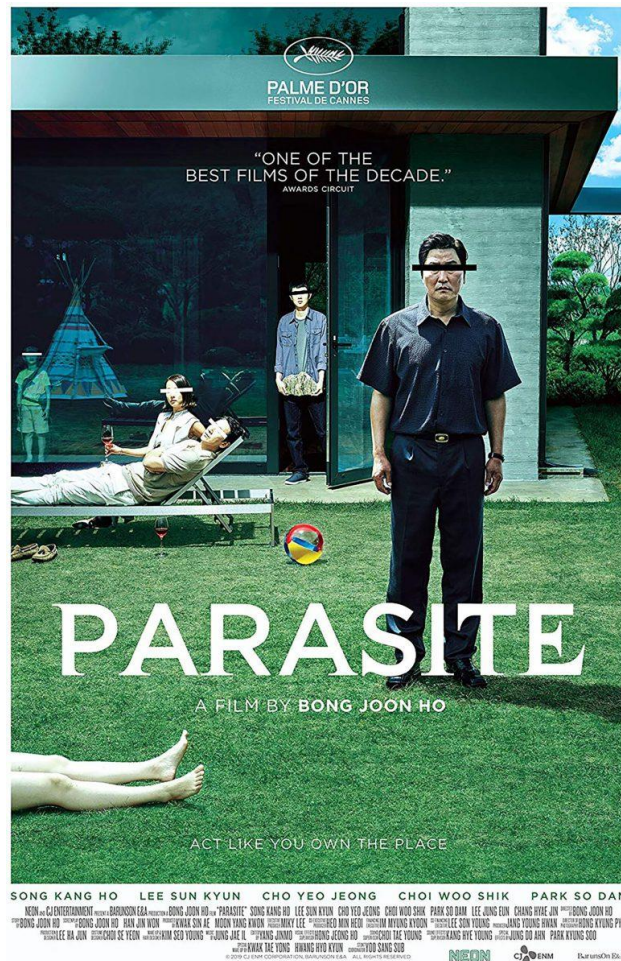
Hypothesis

- Mean of the review will have positive relation with demand, and standard deviation would have negative relation.
- Different importance between *recent* reviews and *overall* reviews.

Hypothesis

Movie Review in 2024/01/13

Overall mean : 8.8



★★★★★ 9

슬픈 대한민국

하니베어 2023. 12. 5. 10:35

★☆☆☆☆ 1

별로 재미 없다,,

이윤호 2023. 11. 17. 11:25

★☆☆☆☆ 1

쓰레기영화를 이렇게들 좋아하다니

젠름맨리턴즈 2023. 8. 15. 19:49

★★★★★ 10

가해자는 없으나 피해자는 있다. 그것이 피해자가 가해자로 변환되는 지점이다.

ADIA 2023. 11. 18. 20:08

★☆☆☆☆ 2

지하층에 살수 밖에 없는 가난한 사람들의 심정을 더욱더 처참하게 만든 영화~~~~

파랑새 2023. 4. 3. 15:24

Hypothesis

Movie Review in 2024/01/13

Overall mean : 6.8



★★★★☆ 5

2부는 재미 있을려나 ㅋㅋ

[Tough Cookie](#) 2023. 11. 26. 01:48

★★★★☆ 7

난 재밌게 봤는데 이상하네~

[로즈님](#) 2023. 12. 3. 18:06

★★★★★ 9

막 봤는데 영화관에서 안본게 아쉽네요~ 엄청 잘만들 영화인데 저평가가 되어 안타까바~
2부는 꼭 극장에서 봐야 겠어요~한국영화 멋지넹^^

[개똥천사](#) 2023. 12. 12. 11:09

★★★★☆ 6

잘 풀 거라고 믿으면 2부를 기다리며 어렵다고 오버는 말자.

[날손](#) 2023. 11. 25. 17:01

Hypothesis

Movie Review in 2024/01/13

Overall mean : 8.8



★★★★★ 10

관람객 미지의 목소리에서 나를 찾아가는
nulj**** · 2024.01.09. 12:07 · 신고

★★★★★ 10

관람객 1편에 비해 노래도 스토리도 아쉽지만 재밌었어요.
kivo**** · 2023.12.31. 13:51 · 신고

★★★★★ 9

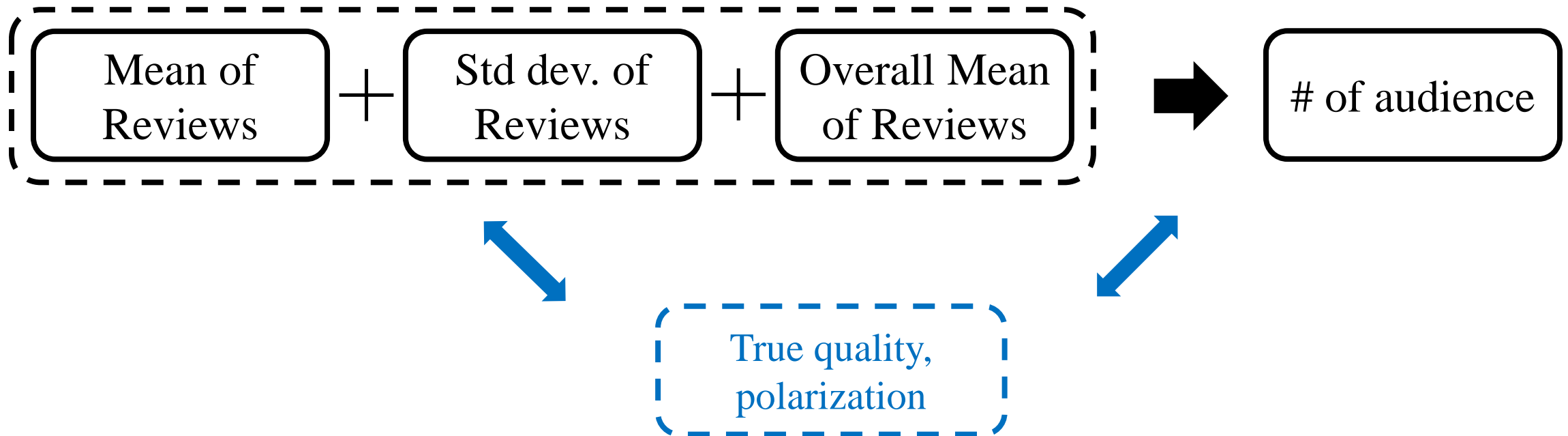
관람객 Frozen 2, 2019
wal9**** · 2024.01.07. 16:15 · 신고

★★★★★ 10

관람객 좋은 영화 재밌게 관람했어요.
hdj6**** · 2023.12.28. 21:05 · 신고

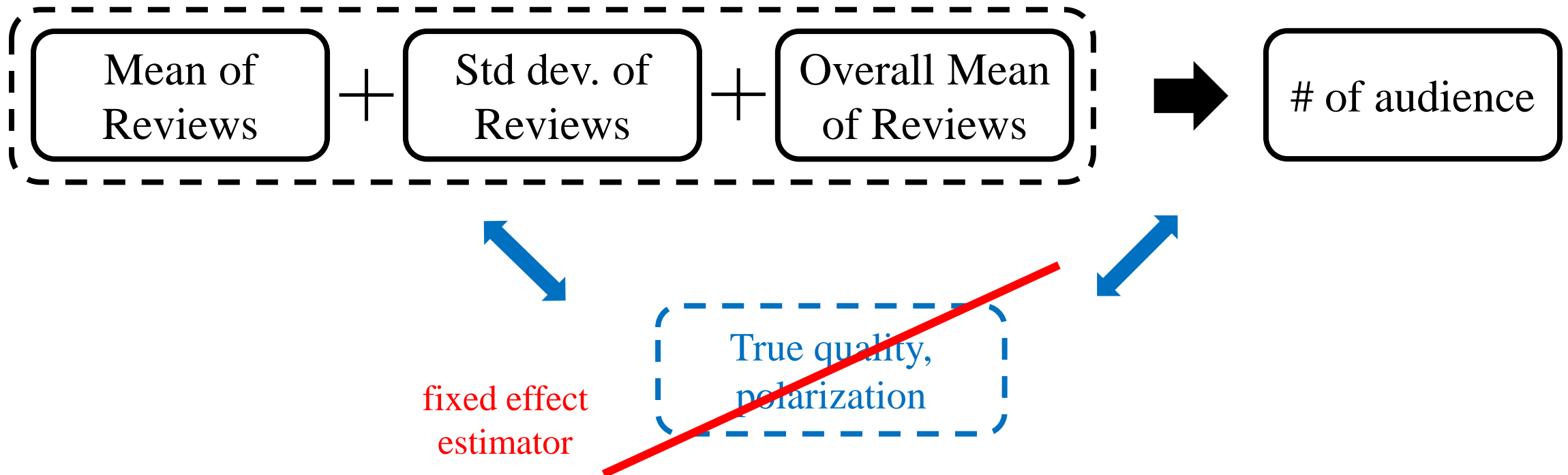
Empirical Strategy

- One challenge to observe the effect of the movie reviews is that those have some correlation with real movie qualities and polarization. (**Omitted Variable Problem**)



Empirical Strategy

- Assuming effect of true quality and polarization of movie on the each day's movie demand is constant, those *unobserved effect* can be controlled using **fixed effect estimator**.



Empirical Strategy

- Our model is

$$\log(Audience_{i,t}) = \beta_0 + \beta_1 \log(lastmean_{i,t-1}) + \beta_2 \log(laststd_{i,t-1}) \\ + \beta_3 \log(overallmean_{i,t-1}) + \alpha_i + \sum_{j=1}^6 \gamma_j DoW_{i,t,j} + \sum_{k=3}^8 \omega_{k-1} dayk_t + \varepsilon_{i,t}$$

where i denotes the index of movie, and t denotes the time ($t \geq 2$).

- logged value of each variables were used to compare which one is more important.
- $Audience_{i,t}$ is movie i 's number of audiences in day t .
- $lastmean/laststd_{i,t-1}$ is the mean and standard deviation of reviews in day $t - 1$.
- $overallmean$ is the mean of movie reviews which was written before day $t - 1$.

Empirical Strategy

- Our model is

$$\begin{aligned} \log(Acudience_{i,t}) = & \beta_0 + \beta_1 \log(lastmean_{i,t-1}) + \beta_2 \log(laststd_{i,t-1}) \\ & + \beta_3 \log(overallmean_{i,t-1}) + \alpha_i + \sum_{j=1}^6 \gamma_j DoW_{i,t,j} + \omega_1 day3_t + \cdots + \omega_6 day8_t + \varepsilon_{i,t} \end{aligned}$$

where i denotes the index of movie, and t denotes the time ($t \geq 2$).

- α_i is time invariant characteristic of movie i such as **True quality** and **like and dislike**, causing **omitted variable problem**. This is excluded by **within fixed effect estimator**, which is going to be used in our research.

Empirical Strategy

- Our model is

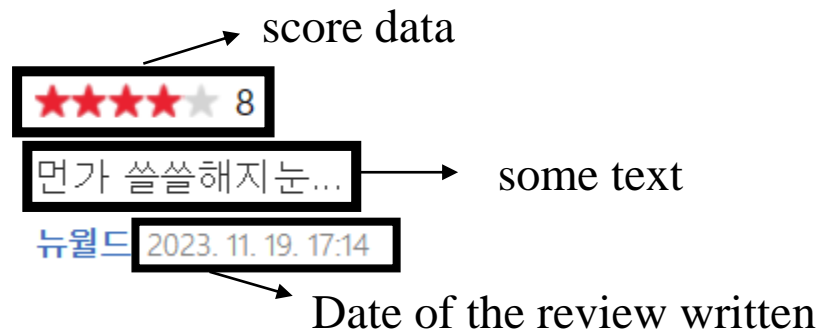
$$\log(Acudience_{i,t}) = \beta_0 + \beta_1 \log(lastmean_{i,t-1}) + \beta_2 \log(laststd_{i,t-1}) \\ + \beta_3 \log(overallmean_{i,t-1}) + \alpha_i + \sum_{j=1}^6 \gamma_j Dow_{i,t,j} + \omega_1 day3_t + \dots + \omega_6 day8_t + \varepsilon_{i,t}$$

where i denotes the index of movie, and t denotes the time ($t \geq 2$).

- $Dow_{i,t}$ is dummy variable for day of week with Monday base. 1 if t is certain day of week, 0 otherwise. These controls different movie demand in each day of week ; *day effect*.
- $day3_t, \dots, day8_t$ is dummy variable which captures day after movie release. 1 if t is k^{th} day after movie release, 0 otherwise. These controls the fact that movie demand decreases as the time goes by .

Data

- **Box office data** of 702 Korean movies from Korea Movie Council which was on the screen from 2006 to 2023 was used.
- Each movie data has the box office data from its opening date to 7 days later from its opening date. (Day 1 ~ Day 7)
- **Review data** was derived from Daum Movie Review, which is one of the major Korean movie critique site.



Data

- Summary statistics

	mean	std	min	median	max
# of audience	10349.36	14578.52	131	53863	1466225
Last day mean	7.314677	1.690849	0.197531	7.571429	9.972973
Last day std	2.849925	0.997636	0.164399	3.003924	6.363961
Overall mean	7.874559	1.201359	1.434783	8.059233	9.933052

Data

- Correlation coefficient matrix

	mean	std	Overall mean
mean	1.000000	-0.700555	0.754345
std	-0.700555	1.000000	-0.541912
overallmean	0.754345	-0.541912	1.000000

Results and implication

	Our Model	Our Model without DoW and day	PooledOLS
Last Mean	-0.0157 (0.0454)	0.0149 (0.0701)	1.0021*** (0.127)
Last Std	0.0258 (0.0193)	0.04 (0.0337)	0.5816*** (0.0568)
Overall Mean	1.3431*** (0.1662)	4.1146*** (0.3722)	-0.1084 (0.2105)
DoW_Tuesday	0.0353 (0.0225)	-	-
...
day7	-0.7558*** (0.0363)	-	-
Const	8.1647*** (0.3428)	2.2255*** (0.1084)	8.4311*** (0.371)
R-Squared	0.6812	0.1084	0.0419
F-Statistic	598.77***	170.9***	71.635***
F-statistic's P-value	0.000	0.000	0.000

- Our hypothesis is *partially true*:
which means the hypothesis about
mean is true, but not the hypothesis
about standard deviation.

- *: 10% significance, **: 5% significance, ***: 1% significance. Standard errors of each estimates are reported in the parentheses, and heteroskedasticity robust.

Results and implication

	Our Model	Our Model without DoW and day	PooledOLS
Last Mean	-0.0157 (0.0454)	0.0149 (0.0701)	1.0021*** (0.127)
Last Std	0.0258 (0.0193)	0.04 (0.0337)	0.5816*** (0.0568)
Overall Mean	1.3431*** (0.1662)	4.1146*** (0.3722)	-0.1084 (0.2105)
DoW_Tuesday	0.0353 (0.0225)	-	-
...
day7	-0.7558*** (0.0363)	-	-
Const	8.1647*** (0.3428)	2.2255*** (0.1084)	8.4311*** (0.371)
R-Squared	0.6812	0.1084	0.0419
F-Statistic	598.77***	170.9***	71.635***
F-statistic's P-value	0.000	0.000	0.000

- People tend to watch *overall mean*. This means people does not like to scroll down the webpages to see the recent reviews. In other words, the *cost* of scrolling down for people is quite high.

- *: 10% significance, **: 5% significance, ***: 1% significance. Standard errors of each estimates are reported in the parentheses, and heteroskedasticity robust.

Results and implication

	Our Model	Our Model without DoW and day	PooledOLS
Last Mean	-0.0157 (0.0454)	0.0149 (0.0701)	1.0021*** (0.127)
Last Std	0.0258 (0.0193)	0.04 (0.0337)	0.5816*** (0.0568)
Overall Mean	1.3431*** (0.1662)	4.1146*** (0.3722)	-0.1084 (0.2105)
DoW_Tuesday	0.0353 (0.0225)	-	-
...
day7	-0.7558*** (0.0363)	-	-
Const	8.1647*** (0.3428)	2.2255*** (0.1084)	8.4311*** (0.371)
R-Squared	0.6812	0.1084	0.0419
F-Statistic	598.77***	170.9***	71.635***
F-statistic's P-value	0.000	0.000	0.000

- The result is significantly changed as we use the fixed effect estimator, not standard OLS estimator: this means that the bias comes from omitted variables such as true movie quality and polarization was high.

- *: 10% significance, **: 5% significance, ***: 1% significance. Standard errors of each estimates are reported in the parentheses, and heteroskedasticity robust.

Results and implication

	Our Model	Our Model without DoW and day	PooledOLS
Last Mean	-0.0157 (0.0454)	0.0149 (0.0701)	1.0021*** (0.127)
Last Std	0.0258 (0.0193)	0.04 (0.0337)	0.5816*** (0.0568)
Overall Mean	1.3431*** (0.1662)	4.1146*** (0.3722)	-0.1084 (0.2105)
DoW_Tuesday	0.0353 (0.0225)	-	-
...
day7	-0.7558*** (0.0363)	-	-
Const	8.1647*** (0.3428)	2.2255*** (0.1084)	8.4311*** (0.371)
R-Squared	0.6812	0.1084	0.0419
F-Statistic	598.77***	170.9***	71.635***
F-statistic's P-value	0.000	0.000	0.000

- *: 10% significance, **: 5% significance, ***: 1% significance. Standard errors of each estimates are reported in the parentheses, and heteroskedasticity robust.

- The estimate also changes when the DoW effect and the day effect has been controlled. This means the effect of overall mean might be *overestimated* if we did not considered the effect of them.

Conclusion

- The result shows that our hypothesis is *partially true*: the hypothesis about mean was true, but not the hypothesis about standard deviation.
- Also, the result supports that people tend to only watch the overall mean, and not to scroll down the pages. This implies that people might feel the cost of scrolling down is *high*.

Limits

- Even if we tried to control the **omitted variable problem**, there might be another potential problem..
- **First, Selection Bias**: As we used panel data to deal with omitted variable problem, **some movies without enough movie review data** was excluded: this means that our result can be only applied to the movie which has enough popularity level.
- **Second, Marketing**: Time-invariant omitted variables, such as real movie quality we've mentioned before, would be subtracted. However, the effect of the **marketing** would be time-variant, and this can cause **another omitted variable problem**.

Limits

- Even if we tried to control the **omitted variable problem**, there might be another potential problem..
- **Third, model misspecification:** When we set our panel regression model, we implicitly assumed that people would decide whether to see or not and watch the movie in the *next day*. However, this might not be true: the term between the decision day and watching day can be different.