Unmasking the Deception: The Interplay between Fake Reviews, Rating Discrepancy, and Consumer Demand

Yunhao Huang J. Miguel Villas-Boas Mingduo Zhao
UC Berkeley

Oct, 2023

Rating Discrepancy

Measured by standard deviation of ratings

Rating Discrepancy

Measured by standard deviation of ratings

Research Questions

1. Correlation: Rating discrepancy & fake reviews?
   - Why?
     - Companies reimburse buyers for providing 5-star ratings (Li et al., 2020; He et al., 2022)
     - Consumers rate 1-star ratings if experience worse than expected (Chakraborty et al., 2022)
   - Preview:
     - Rating discrepancy strongly correlates with fake review purchase, conditional on product characteristics

Research Questions

1. Correlation: Rating discrepancy & fake reviews?
   - Why?
     - Companies reimburse buyers for providing 5-star ratings (Li et al., 2020; He et al., 2022)
     - Consumers rate 1-star ratings if experience worse than expected (Chakraborty et al., 2022)
   - Preview:
     - Rating discrepancy strongly correlates with fake review purchase, conditional on product characteristics

Research Questions

1. Correlation: Rating discrepancy & fake reviews?
   - Why?
     - Companies reimburse buyers for providing 5-star ratings (Li et al., 2020; He et al., 2022)
     - Consumers rate 1-star ratings if experience worse than expected (Chakraborty et al., 2022)
   - Preview:
     - Rating discrepancy strongly correlates with fake review purchase, conditional on product characteristics

Research Questions

2. Rating discrepancy $\rightarrow$ demand?
   - Average rating $\rightarrow$ demand: Well-established
     - Chen & Xie (2008); Cabral & Hortacsu (2010); Moe & Trusov (2011); De Langhe et al. (2016); Luca (2016); Park et al. (2021); Pei & Mayzlin (2022); Reimers & Waldfogel (2021); Zhong (2022)
   - Discrepancy $\rightarrow$ demand: Mixed evidence
     - $+$: Sun (2012); Rozenkrants et al. (2017)
     - $-$: Luo et al. (2013); He and Bond (2015); We et al. (2015)
   - Preview:
     - Novel identification based on % rounding
     - 0.01 increase in std. $\rightarrow -6.5\%$ demand

Research Questions

2. Rating discrepancy $\rightarrow$ demand?
   - Average rating $\rightarrow$ demand: Well-established
     - Chen & Xie (2008); Cabral & Hortacsu (2010); Moe & Trusov (2011); De Langhe et al. (2016); Luca (2016); Park et al. (2021); Pei & Mayzlin (2022); Reimers & Waldfogel (2021); Zhong (2022)
   - Discrepancy $\rightarrow$ demand: Mixed evidence
     - $+$: Sun (2012); Rozenkrants et al. (2017)
     - $-$: Luo et al. (2013); He and Bond (2015); We et al. (2015)
   - Preview:
     - Novel identification based on % rounding
     - 0.01 increase in std. $\rightarrow$ -6.5% demand

Research Questions

2. Rating discrepancy $\rightarrow$ demand?
   - Average rating $\rightarrow$ demand: Well-established
     - Chen & Xie (2008); Cabral & Hortacsu (2010); Moe & Trusov (2011); De Langhe et al. (2016); Luca (2016); Park et al. (2021); Pei & Mayzlin (2022); Reimers & Waldfogel (2021); Zhong (2022)
   - Discrepancy $\rightarrow$ demand: Mixed evidence
     - $+$: Sun (2012); Rozenkrants et al. (2017)
     - $-$: Luo et al. (2013); He and Bond (2015); We et al. (2015)
   - Preview:
     - Novel identification based on % rounding
     - 0.01 increase in std. $\rightarrow$ -6.5% demand

Research Questions

3. Suspicion of fake reviews $\rightarrow$ demand?
   - Consumer response to misinformation
     - De Pechpeyrou & Odou (2012), Pennycook et al. (2020), Wu & Geylani (2020), Rao (2022), Simonov and Rao (2022), Fong et al. (2023)
   - Experiment: Information treatment (IV)
   - Preview:
     - 1% more suspicion $\rightarrow -6.5\%$ demand
     - Socially disadvantaged groups benefit more from information treatment

Research Questions

3. Suspicion of fake reviews $\rightarrow$ demand?
   - Consumer response to misinformation
     - De Pechpeyrou & Odou (2012), Pennycook et al. (2020), Wu & Geylani (2020), Rao (2022), Simonov and Rao (2022), Fong et al. (2023)
   - Experiment: Information treatment (IV)
   - Preview:
     - 1% more suspicion $\rightarrow -6.5\%$ demand
     - Socially disadvantaged groups benefit more from information treatment

Research Questions

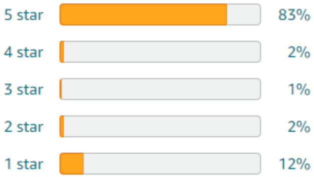
3. Suspicion of fake reviews $\rightarrow$ demand?
   - Consumer response to misinformation
     - De Pechpeyrou & Odou (2012), Pennycook et al. (2020), Wu & Geylani (2020), Rao (2022), Simonov and Rao (2022), Fong et al. (2023)
   - Experiment: Information treatment (IV)
   - Preview:
     - 1% more suspicion $\rightarrow$ -6.5% demand
     - Socially disadvantaged groups benefit more from information treatment

Research Questions

3. Suspicion of fake reviews $\rightarrow$ demand?
   - Consumer response to misinformation
     - De Pechpeyrou & Odou (2012), Pennycook et al. (2020), Wu & Geylani (2020), Rao (2022), Simonov and Rao (2022), Fong et al. (2023)
   - Experiment: Information treatment (IV)
   - Preview:
     - 1% more suspicion $\rightarrow$ -6.5% demand
     - Socially disadvantaged groups benefit more from information treatment

Outline

Unmasking the Deception

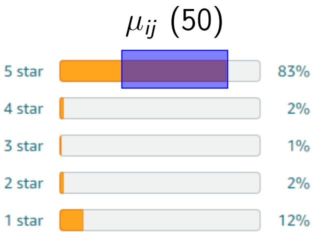


Observed rating ρ_j

- Consumer i only observe ρ_j
- Concerns: μ_{ij} are fake
- τ_{ij}: conjecture after removing μ_{ij} 5-star ratings

$$\rho_j = \mu_{ij}\delta_5 + (1 - \mu_{ij})\tau_{ij} \quad \Rightarrow \quad \tau_{ij} = \frac{\rho_j - \mu_{ij}\delta_5}{1 - \mu_{ij}}$$

Unmasking the Deception

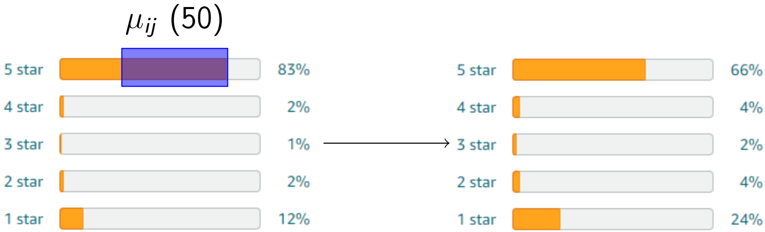


Observed rating ρ_j

- Consumer i only observe ρ_j
- Concerns: μ_{ij} are fake
- τ_{ij}: conjecture after removing μ_{ij} 5-star ratings

$$\rho_j = \mu_{ij}\delta_5 + (1 - \mu_{ij})\tau_{ij} \quad \Rightarrow \quad \tau_{ij} = \frac{\rho_j - \mu_{ij}\delta_5}{1 - \mu_{ij}}$$

Unmasking the Deception



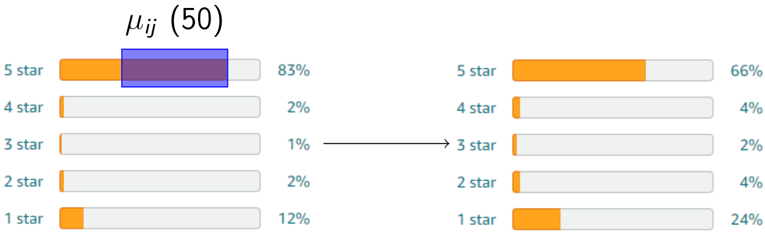
Observed rating ρ_j

Conjecture τ_{ij}

- Consumer i only observe ρ_j
- Concerns: μ_{ij} are fake
- τ_{ij}: conjecture after removing μ_{ij} 5-star ratings

$$\rho_j = \mu_{ij}\delta_5 + (1 - \mu_{ij})\tau_{ij} \quad \Rightarrow \quad \tau_{ij} = \frac{\rho_j - \mu_{ij}\delta_5}{1 - \mu_{ij}}$$

Unmasking the Deception



Observed rating ρ_j

Conjecture τ_{ij}

- Utility

$$U_{ij} = f(\tau_{ij}) = f\left(\frac{\rho_j - \mu_{ij}\delta_5}{1 - \mu_{ij}}\right)$$

decreases in μ_{ij}

Outline

Dataset I

Secondary dataset, from He et al. (2022)

- Cross-sectional sub-sample, October 2020
- $\sim$1,500 identified products with fake reviews, $\sim$2,000 competitors
- Variables:
  - **Whether purchased fake reviews**
  - Proportions of 1-star & 5-star (impute std.)
  - Avg. ratings, # ratings
  - Product characteristics

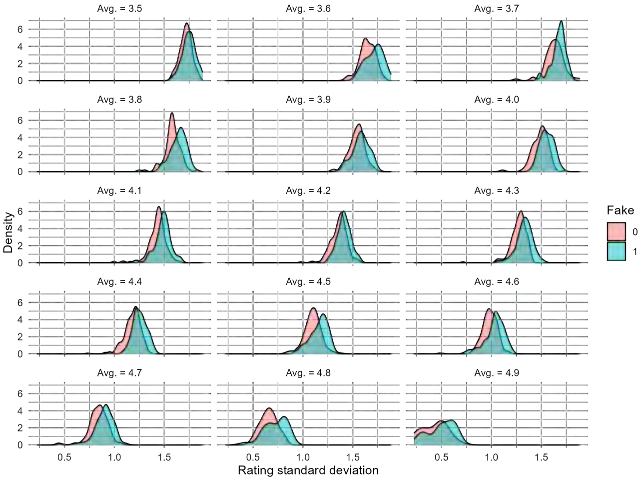
Dataset II

Scraped from Amazon

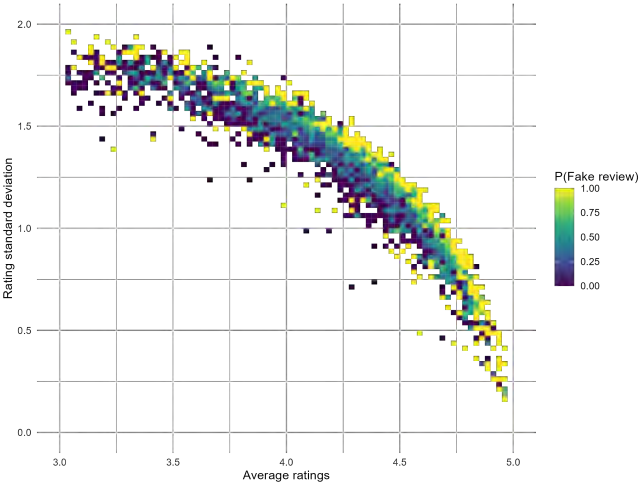
- January–September 2023
- "Home & Kitchen" category, best-seller top 100
- Variables
  - **Sales**
    - Inventory change
    - Imputed others based on Chevalier & Goolsbee (2003)
  - Rating distribution, # ratings
  - Product characteristics (prices, etc.)
  - Position in search list

Outline

Rating Discrepancy and Fake Review



Rating Discrepancy Strongly Predicts Fake Review



Outline

Identification using Percentage Rounding

$$\Delta Y_{kt} = \alpha_t + \beta \cdot \Delta\sigma_{kt} + \gamma^T X_{kt} + e_{kt}$$

- σ_{jt}: Rating std. based on rounded proportions
- Subset of dataset II
  - Adjacent days with same (rounded) average rating but different (rounded) rating distributions

Rating Discrepancy Negatively Affect Sales

	Outcome	
	(1) Log sales	(2) Log sales rankings
Std. ratings	-6.498^{***}	4.211^{***}
	(1.577)	(1.022)
Log price	-1.158^{**}	0.750^{**}
	(0.557)	(0.361)
Covariates	Yes	Yes
Observations	254	254
R^2	0.194	0.194
Adjusted R^2	0.097	0.097

$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Outline

Introduction

Theoretical Framework

Observational Data

Rating Discrepancy and Fake Review

Rating Discrepancy and Demand

Conjoint Experiment

Experiment Design

(a) Product 1

(b) Product 2

(c) Product 3

Asked to choose between each pair and

- Reason of choice
- Perceived rating discrepancy, prob. purchased fake reviews, proportion fake reviews, $\cdots$

Information Treatment

Here is a summary of an *Economists* ariticle. The article is based on a research paper (He, Hollenbeck, & Proserpio, 2022):

Title: A new study analyses the murky world of fake Amazon reviews

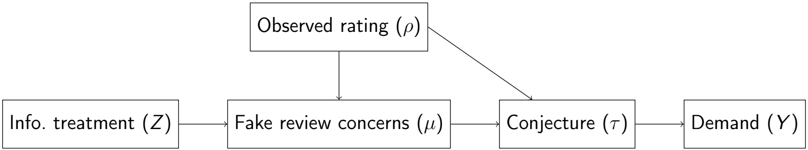
Summary: Companies can temporarily boost sales by paying for fake reviews on Amazon. Researchers identified 1,500 products with manipulated ratings through Facebook groups, where companies promised to reimburse buyers for purchasing their products and leaving **five-star** ratings. Purchasing fake reviews leads to a significant but short-term increase in average ratings and the number of reviews. After firms stop buying fake reviews, their average ratings fall, and the share of **one-star** reviews increases significantly.

- $N = 712$ participants, 350 treated and 362 control

Average Treatment Effects

Variable	(1) Both	(2) Information	(3) Placebo	(4) (2)−(3)	(5) *p*-value
Choose low variance product (1 v. 2)	0.733	0.797	0.671	0.126	0.000
Choose low variance product (1 v. 3)	0.753	0.811	0.696	0.115	0.000
Choose low variance product (2 v. 3)	0.712	0.763	0.663	0.100	0.003
List fake review reason (1 v. 2)	0.306	0.474	0.144	0.331	0.000
List fake review reason (1 v. 3)	0.323	0.483	0.169	0.314	0.000
List fake review reason (2 v. 3)	0.246	0.383	0.113	0.270	0.000
Probability purchased fake reviews (1)	39.146	45.957	32.561	13.396	0.000
Probability purchased fake reviews (2)	28.379	30.331	26.492	3.840	0.026
Probability purchased fake reviews (3)	24.753	25.760	23.779	1.981	0.242
Proportion fake reviews (1)	24.381	28.160	20.727	7.433	0.000
Proportion fake reviews (2)	18.093	19.509	16.724	2.785	0.030
Proportion fake reviews (3)	15.961	16.934	15.019	1.915	0.119
Perceived rating discrepancy (1)	37.294	36.957	37.619	−0.662	0.707
Perceived rating discrepancy (2)	32.955	32.091	33.790	−1.699	0.283
Perceived rating discrepancy (3)	32.235	32.206	32.262	−0.057	0.972
N	712	350	362		

2SLS Regression



- Y_i: Whether choose low-variance product
- C_i: Fake review concern measures
  1. List fake reviews as one reason for choice
  2. Difference in elicited probability of purchasing fake reviews
  3. Difference in elicited proportion of fake reviews
- Instrument C_i with treatment Z_i

2SLS Regression Results

	Outcome: Decide to purchase the low-variance product			
	(1)	(2)	(3)	(4)
	All	Product 1 v. 2	Product 1 v. 3	Product 2 v. 3
List fake review as a reason	0.305***	0.331***	0.314***	0.270***
	(0.019)	(0.032)	(0.033)	(0.031)
Constant	0.142***	0.144***	0.169***	0.113***
	(0.013)	(0.023)	(0.023)	(0.022)
Observations	2,136	712	712	712
R^2	0.112	0.129	0.113	0.098
Adjusted R^2	0.112	0.127	0.112	0.097

$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

2SLS Regression Results

	Outcome: Decide to purchase the low-variance product			
	(1)	(2)	(3)	(4)
	All	Product 1 v. 2	Product 1 v. 3	Product 2 v. 3
Difference in perceived	0.065^{***}	0.077^{***}	0.057^{**}	0.061^{**}
fake review probability (%)	(0.016)	(0.026)	(0.025)	(0.029)
Constant	0.815^{***}	0.815^{***}	0.840^{***}	0.790^{***}
	(0.011)	(0.019)	(0.018)	(0.020)
Observations	2,136	712	712	712
R^2	0.008	0.012	0.007	0.006
Adjusted R^2	0.008	0.010	0.006	0.005

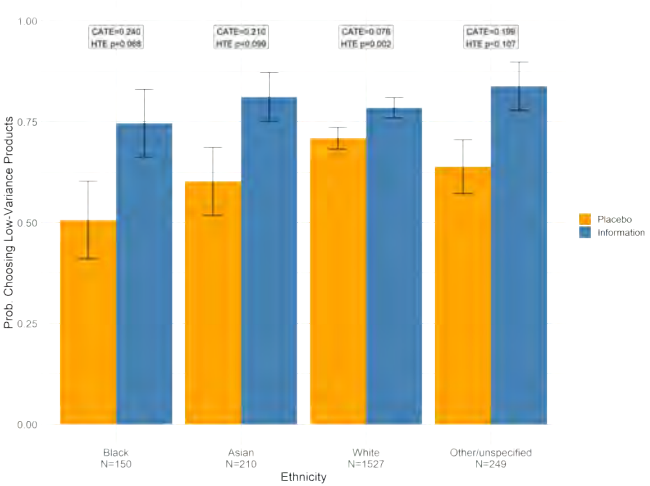
${}^{*}p<0.1$; ${}^{**}p<0.05$; ${}^{***}p<0.01$

2SLS Regression Results

	Outcome: Decide to purchase the low-variance product			
	(1) All	(2) Product 1 v. 2	(3) Product 1 v. 3	(4) Product 2 v. 3
Difference in perceived fake review proportion (%)	0.076***	0.107***	0.063**	0.058**
	(0.016)	(0.027)	(0.025)	(0.030)
Constant	0.804***	0.793***	0.840***	0.779***
	(0.011)	(0.019)	(0.018)	(0.021)
Observations	2,136	712	712	712
R^2	0.011	0.022	0.009	0.005
Adjusted R^2	0.010	0.021	0.007	0.004

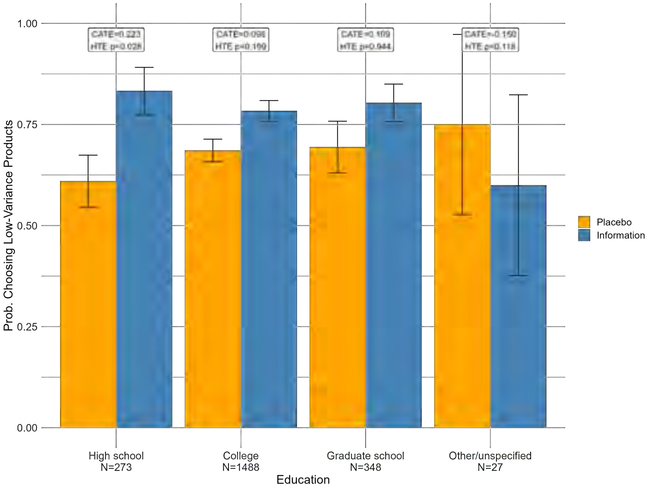
*$p<0.1$; **$p<0.05$; ***$p<0.01$

Heterogeneous Treatment Effects



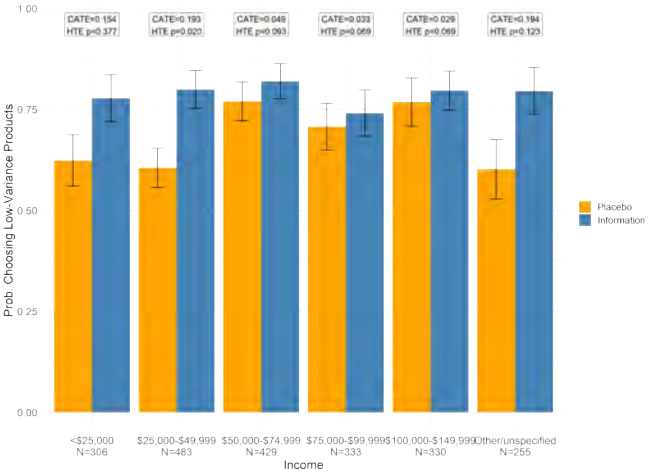
Socially disadvantaged groups have larger CATEs

Heterogeneous Treatment Effects



Socially disadvantaged groups have larger CATEs

Heterogeneous Treatment Effects



Socially disadvantaged groups have larger CATEs

Summary

- Rating discrepancy negatively affects sales
- New channel: Suspicion about fake reviews
- Rating discrepancy strongly predicts fake review purchase
- Fake review concerns significantly affect demand
- Socially disadvantaged groups adjust more to information