Controlling Fake Reviews

Yuta Yasui

October 13, 2020

Introduction

- **Customer ratings** play key roles in platform markets:
  - Hollenbeck et al (2019): ratings vs advertisement in hotel industry
  - Reimers and Waldfogel (2020): ratings vs professional reviews for books

Introduction

- **Customer ratings** play key roles in platform markets:
  - Hollenbeck et al (2019): ratings vs advertisement in hotel industry
  - Reimers and Waldfogel (2020): ratings vs professional reviews for books
- Platform markets are **growing**,
  - so does the **incentive** to make **fake reviews**.

Introduction

- Platforms and regulators are concerned about fake reviews:

Introduction

- Platforms and regulators are concerned about fake reviews:
  - Amazon strictly prohibits incentivized reviews since 2016.
  - In 2019, FTC filed the first case challenging fake paid reviews

Introduction

- Platforms and regulators are concerned about fake reviews:
  - Amazon strictly prohibits incentivized reviews since 2016.
  - In 2019, FTC filed the first case challenging fake paid reviews
- We can still find fake reviews

Question: How should a platform deal with fake reviews?

- Should it reduce fake reviews? (Are fake reviews harmful?)

Question: How should a platform deal with fake reviews?

- Should it reduce fake reviews? (Are fake reviews harmful?)
  - Rational buyers might not be fooled by the fake reviews.

Question: How should a platform deal with fake reviews?

- Should it reduce fake reviews? (Are fake reviews harmful?)
  - Rational buyers might not be fooled by the fake reviews.
  - A boosted rating might work as a signal of good quality.
    - It might pay off only for high quality sellers through future sales. (similarly to Nelson; 70,74)

Question: How should a platform deal with fake reviews?

- Should it reduce fake reviews? (Are fake reviews harmful?)
  - Rational buyers might not be fooled by the fake reviews.
  - A boosted rating might work as a signal of good quality.
    - It might pay off only for high quality sellers through future sales. (similarly to Nelson; 70,74)

- Instruments of the platform:
  1. filtering policy on suspicious reviews
  2. weights on old/new reviews

Overview

Overview

- A strict filtering policy **reduces**
  - #(fake reviews) in expectation,
  - impacts of fake reviews on the rating.

Overview

- A strict filtering policy **reduces**
  - #(fake reviews) in expectation,
  - impacts of fake reviews on the rating.
- For rational consumers:
- For naive consumers:

Overview

- A strict filtering policy **reduces**
  - #(fake reviews) in expectation,
  - impacts of fake reviews on the rating.
- For rational consumers:
  - a rating with **fake reviews** can be more **informative** than one without fake reviews
- For naive consumers:

Overview

- A strict filtering policy **reduces**
  - #(fake reviews) in expectation,
  - impacts of fake reviews on the rating.
- For rational consumers:
  - a rating with **fake reviews** can be more **informative** than one without fake reviews
- For naive consumers:
  - a strict **filtering** policy **reduces bias** for the naive consumers (as long as positive number of fake reviews are observed).

Overview

- A strict filtering policy **reduces**
  - #(fake reviews) in expectation,
  - impacts of fake reviews on the rating.
- For rational consumers:
  - a rating with **fake reviews** can be more **informative** than one without fake reviews
  - **old reviews** should be **weighted more** than the optimal level without fake reviews.
- For naive consumers:
  - a strict **filtering** policy **reduces bias** for the naive consumers (as long as positive number of fake reviews are observed).

Overview

- A strict filtering policy **reduces**
  - #(fake reviews) in expectation,
  - impacts of fake reviews on the rating.
- For rational consumers:
  - a rating with **fake reviews** can be more **informative** than one without fake reviews
  - **old reviews** should be **weighted more** than the optimal level without fake reviews.
- For naive consumers:
  - a strict **filtering** policy **reduces bias** for the naive consumers (as long as positive number of fake reviews are observed).
- #(fake reviews) is **increasing in the quality** and **decreasing in the rating**.
  - Implications for empirical analysis

Literature

Design of Rating Systems

- **[certification]** Lizzeri (1999), Harbaugh and Rasmusen (2018), DeMarzo, Kremer, Skrzypacz (2019), Hopenhayn and Saeedi (2019), Hui et al (2018), Zapechelnyuk (2020)
- **[scoring][one-shot]** Ball (2019), **[dynamic]** Vellodi (2019): entry/exit, directed search;

  Horner and Lambert (2018), Bonatti and Cisternas (2020): **signal jamming**
  This paper:
  - Fake reviews **with refunds**
  - Impact of a filtering policy on the rating's precision
  - Naive consumers

Promotion and Signaling (Q: The higher quality, the more promotion?)

Literature

Design of Rating Systems

- **[certification]** Lizzeri (1999), Harbaugh and Rasmusen (2018), DeMarzo, Kremer, Skrzypacz (2019), Hopenhayn and Saeedi (2019), Hui et al (2018), Zapechelnyuk (2020)
- **[scoring][one-shot]** Ball (2019), **[dynamic]** Vellodi (2019): entry/exit, directed search;

  Horner and Lambert (2018), Bonatti and Cisternas (2020): **signal jamming**
  This paper:

  - Fake reviews **with refunds**
  - Impact of a filtering policy on the rating's precision
  - Naive consumers

Promotion and Signaling (Q: The higher quality, the more promotion?)

Literature

Design of Rating Systems

- **[certification]** Lizzeri (1999), Harbaugh and Rasmusen (2018), DeMarzo, Kremer, Skrzypacz (2019), Hopenhayn and Saeedi (2019), Hui et al (2018), Zapechelnyuk (2020)
- **[scoring][one-shot]** Ball (2019), **[dynamic]** Vellodi (2019): entry/exit, directed search;

  Horner and Lambert (2018), Bonatti and Cisternas (2020): **signal jamming**
  This paper:

  - Fake reviews **with refunds**
  - Impact of a filtering policy on the rating's precision
  - Naive consumers

Promotion and Signaling (Q: The higher quality, the more promotion?)

- **[One shot promotion]** Nelson (1970, 1974), Kihlstrom and Riordan (1984), Milgrom and Roberts (1986), Horstmann and Moorthy (2003), Mayzlin (2006), Dellarocas (2006):

Literature

Design of Rating Systems

- **[certification]** Lizzeri (1999), Harbaugh and Rasmusen (2018), DeMarzo, Kremer, Skrzypacz (2019), Hopenhayn and Saeedi (2019), Hui et al (2018), Zapechelnyuk (2020)
- **[scoring][one-shot]** Ball (2019), **[dynamic]** Vellodi (2019): entry/exit, directed search;

  Horner and Lambert (2018), Bonatti and Cisternas (2020): **signal jamming**
  This paper:
  - Fake reviews **with refunds**
  - Impact of a filtering policy on the rating's precision
  - Naive consumers

Promotion and Signaling (Q: The higher quality, the more promotion?)

- [**One shot promotion**] Nelson (1970, 1974), Kihlstrom and Riordan (1984), Milgrom and Roberts (1986), Horstmann and Moorthy (2003), Mayzlin (2006), Dellarocas (2006):
- [**Repeated promotion**] Horstmann and MacDonald (1994):
  This paper:
  - New source of signaling promotion caused by dynamics
  - Promotion's dependence on the rating/reputation
  - Implication on quality and reputation in empirical research

Motivating example

Fake reviews with "verified purchase" on Amazon

Note:

Motivating example

Fake reviews with “verified purchase” on Amazon

1. The seller posts info of the product and offers full refund (+ extra)

Note:

Motivating example

Fake reviews with "verified purchase" on Amazon

1. The seller posts info of the product and offers full refund (+ extra)
2. Fake reviewers buy the product and write a good review on Amazon.

Note:

Motivating example

Fake reviews with "verified purchase" on Amazon

1. The seller posts info of the product and offers full refund (+ extra)
2. Fake reviewers buy the product and write a good review on Amazon.
3. After verifying the review, the seller refunds the product via PayPal.

Note:

Motivating example

Fake reviews with "verified purchase" on Amazon

1. The seller posts info of the product and offers full refund (+ extra)
2. Fake reviewers buy the product and write a good review on Amazon.
3. After verifying the review, the seller refunds the product via PayPal.
4. Amazon detects and deletes a part of the fake reviews

Note:

Motivating example

Fake reviews with "verified purchase" on Amazon

1. The seller posts info of the product and offers full refund (+ extra)
2. Fake reviewers buy the product and write a good review on Amazon.
3. After verifying the review, the seller refunds the product via PayPal.
4. Amazon detects and deletes a part of the fake reviews

Note:

- The platform takes a **transaction fee** from each fake reviews

Motivating example

Fake reviews with "verified purchase" on Amazon

1. The seller posts info of the product and offers full refund (+ extra)
2. Fake reviewers buy the product and write a good review on Amazon.
3. After verifying the review, the seller refunds the product via PayPal.
4. Amazon detects and deletes a part of the fake reviews

Note:

- The platform takes a **transaction fee** from each fake reviews
  - (Revenue from the fake sales) $<$ (Refund of the fake sales)

Model (1/3)

- Time: $t \in [0, \infty)$

Model (1/3)

- Time: $t \in [0, \infty)$
- Players: a long lived seller, many short lived buyers
  - (a platform can control parameters before the game starts)

Model (1/3)

- Time: $t \in [0, \infty)$
- Players: a long lived seller, many short lived buyers
  - (a platform can control parameters before the game starts)
- Action at time t
  - Seller:
    - choose the amount of the fake reviews: $F_t \in \mathbb{R}$
    - (sell q units of the product: **fixed/normalized to 1**)
  - Buyers:
    - buy the product, or not
      $\rightarrow$ form the equilibrium price: p_t

Model (1/3)

- Time: $t \in [0, \infty)$
- Players: a long lived seller, many short lived buyers
  - (a platform can control parameters before the game starts)
- Action at time t
  - Seller:
    - choose the amount of the fake reviews: $F_t \in \mathbb{R}$
    - (sell q units of the product: **fixed/normalized to 1**)
  - Buyers:
    - buy the product, or not
      $\rightarrow$ form the equilibrium price: p_t
- State:
  - θ_t: seller's type (quality of the product) at t
  - Y_t: seller's rating at t

Model (1/3)

- Time: $t \in [0, \infty)$
- Players: a long lived seller, many short lived buyers
  - (a platform can control parameters before the game starts)
- Action at time t
  - Seller:
    - choose the amount of the fake reviews: $F_t \in \mathbb{R}$
    - (sell q units of the product: **fixed/normalized to 1**)
  - Buyers:
    - buy the product, or not
      $\rightarrow$ form the equilibrium price: p_t
- State:
  - θ_t: seller's type (quality of the product) at t
  - Y_t: seller's rating at t
- Information:
  - Seller at time t: the whole history so far $= (\theta_s, Y_s, F_s, p_s)_{s \in [0, t]}$
  - Buyers at time t: current rating $= Y_t$

Model (1/3)

- Time: $t \in [0, \infty)$
- Players: a long lived seller, many short lived buyers
  - (a platform can control parameters before the game starts)
- Action at time t
  - Seller:
    - choose the amount of the fake reviews: $F_t \in \mathbb{R}$
    - (sell q units of the product: **fixed/normalized to 1**)
  - Buyers:
    - buy the product, or not
      $\rightarrow$ form the equilibrium price: p_t
- State:
  - θ_t: seller's type (quality of the product) at t
  - Y_t: seller's rating at t
- Information:
  - Seller at time t: the whole history so far $= (\theta_s, Y_s, F_s, p_s)_{s \in [0, t]}$
  - Buyers at time t: current rating $= Y_t$

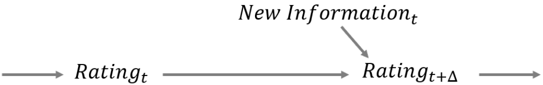
Model (1/3)

- Time: $t \in [0, \infty)$
- Players: a long lived seller, many short lived buyers
  - (a platform can control parameters before the game starts)
- Action at time t
  - Seller:
    - choose the amount of the fake reviews: $F_t \in \mathbb{R}$
    - (sell q units of the product: **fixed/normalized to 1**)
  - Buyers:
    - buy the product, or not
      $\rightarrow$ form the equilibrium price: p_t
- State:
  - θ_t: seller's type (quality of the product) at t
  - Y_t: seller's rating at t
- Information:
  - Seller at time t: the whole history so far $= (\theta_s, Y_s, F_s, p_s)_{s \in [0, t]}$
  - Buyers at time t: current rating $= Y_t$

Model (1/3)

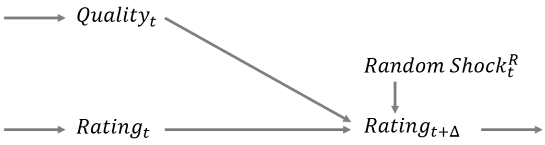
- Time: $t \in [0, \infty)$
- Players: a long lived seller, many short lived buyers
  - (a platform can control parameters before the game starts)
- Action at time t
  - Seller:
    - choose the amount of the fake reviews: $F_t \in \mathbb{R}$
    - (sell q units of the product: **fixed/normalized to 1**)
  - Buyers:
    - buy the product, or not
      $\rightarrow$ form the equilibrium price: $p_t = E[\theta_t | Y_t] \equiv M_t$ Details
- State:
  - θ_t: seller's type (quality of the product) at t
  - Y_t: seller's rating at t
- Information:
  - Seller at time t: the whole history so far $= (\theta_s, Y_s, F_s, p_s)_{s \in [0, t]}$
  - Buyers at time t: current rating $= Y_t$

Model (2/3)



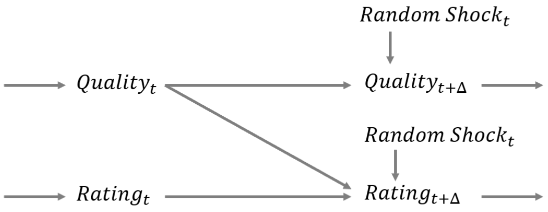
- State transition:
  - Rating Y_t follows $Y_{t+dt} = Y_t(1 - \phi dt) + (\text{new reviews})$
  - Quality θ_t follows $\theta_{t+dt} = \theta_t(1 - \kappa dt) + \mu\kappa dt + \sigma_\theta dZ_t^\theta$
    - exogenous for players (seller/buyers) and for the platform

Model (2/3)



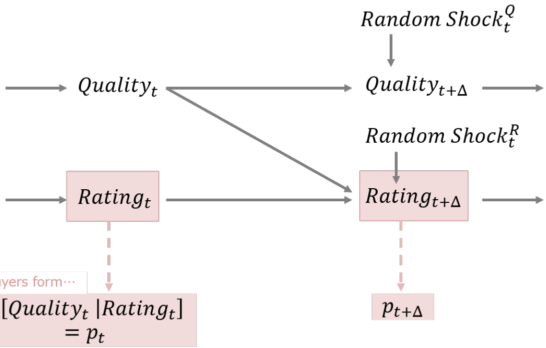
- State transition:
  - Rating Y_t follows $Y_{t+dt} = Y_t\,(1 - \phi dt) + aF_t dt + \theta_t dt + \sigma_\xi dZ_t^\xi$
  - Quality θ_t follows $\theta_{t+dt} = \theta_t\,(1 - \kappa dt) + \mu\kappa dt + \sigma_\theta dZ_t^\theta$
    - exogenous for players (seller/buyers) and for the platform

Model (2/3)



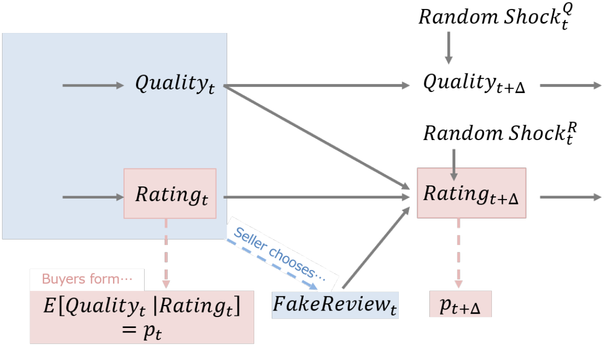
- State transition:
  - Rating Y_t follows $Y_{t+dt} = Y_t\,(1-\phi dt) + aF_t dt + \theta_t dt + \sigma_\xi dZ_t^\xi$
  - Quality θ_t follows $\theta_{t+dt} = \theta_t\,(1-\kappa dt) + \mu\kappa dt + \sigma_\theta dZ_t^\theta$
    - exogenous for players (seller/buyers) and for the platform

Model (2/3)



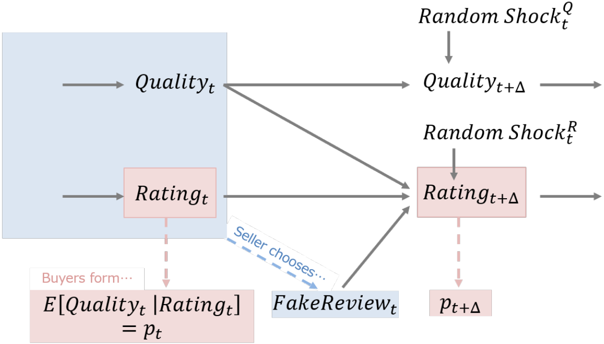
- State transition:
  - Rating Y_t follows $Y_{t+dt} = Y_t\,(1-\phi dt) + aF_t dt + \theta_t dt + \sigma_\xi dZ_t^\xi$
  - Quality θ_t follows $\theta_{t+dt} = \theta_t\,(1-\kappa dt) + \mu\kappa dt + \sigma_\theta dZ_t^\theta$
    - exogenous for players (seller/buyers) and for the platform

Model (2/3)



- State transition:
  - Rating Y_t follows $Y_{t+dt} = Y_t(1 - \phi dt) + aF_t dt + \theta_t dt + \sigma_\xi dZ_t^\xi$
    - $a > 0$: effectiveness of fake reviews. (low a = stringent censorship)
  - Quality θ_t follows $\theta_{t+dt} = \theta_t(1 - \kappa dt) + \mu\kappa dt + \sigma_\theta dZ_t^\theta$
    - exogenous for players (seller/buyers) and for the platform

Model (2/3)



- State transition:
  - Rating Y_t follows $dY_t = -\phi Y_t dt + aF_t dt + \theta_t dt + \sigma_\xi dZ^\xi_t$
    - $a > 0$: effectiveness of fake reviews. (low a = stringent censorship)
  - Quality θ_t follows $d\theta_t = -\kappa(\theta_t - \mu)\, dt + \sigma_\theta dZ^\theta_t$
    - exogenous for players (seller/buyers) and for the platform

Model (3/3)

- Seller's instantaneous payoff:

$$\pi_t = \underbrace{(1-\tau)\,p_t\,(1+F_t)}_{\text{revenue}} - \underbrace{p_t \cdot F_t}_{\text{refund}} - \underbrace{\frac{c}{2}F_t^2}_{\text{other costs}}$$

$$= (1-\tau)\,p_t - \tau p_t \cdot F_t - \frac{c}{2}F_t^2$$

  - τ: transaction fee imposed by the platform.

Model (3/3)

- Seller's instantaneous payoff:

$$\pi_t = \underbrace{(1-\tau)\, p_t \,(1+F_t)}_{\text{revenue}} - \underbrace{p_t \cdot F_t}_{\text{refund}} - \underbrace{\frac{c}{2}F_t^2}_{\text{other costs}}$$

$$= (1-\tau)\, p_t - \tau p_t \cdot F_t - \frac{c}{2}F_t^2$$

  - τ: transaction fee imposed by the platform.

Model (3/3)

- Seller's instantaneous payoff:

$$\pi_t = \underbrace{(1-\tau)\, p_t\, (1+F_t)}_{\text{revenue}} - \underbrace{p_t \cdot F_t}_{\text{refund}} - \underbrace{\frac{c}{2} F_t^2}_{\text{other costs}}$$

$$= (1-\tau)\, p_t - \tau p_t \cdot F_t - \frac{c}{2} F_t^2$$

  - τ: transaction fee imposed by the platform.

Model (3/3)

- Seller's instantaneous payoff:

$$\pi_t = \underbrace{(1-\tau)\, p_t\, (1+F_t)}_{\text{revenue}} - \underbrace{p_t \cdot F_t}_{\text{refund}} - \underbrace{\frac{c}{2} F_t^2}_{\text{other costs}}$$

$$= (1-\tau)\, p_t - \tau p_t \cdot F_t - \frac{c}{2} F_t^2$$

  - τ: transaction fee imposed by the platform.

Model (3/3)

- Seller's instantaneous payoff:

$$\pi_t = \underbrace{(1-\tau)\, p_t\, (1+F_t)}_{\text{revenue}} - \underbrace{p_t \cdot F_t}_{\text{refund}} - \underbrace{\frac{c}{2} F_t^2}_{\text{other costs}}$$
$$= (1-\tau)\, p_t - \tau p_t \cdot F_t - \frac{c}{2} F_t^2$$

  - τ: transaction fee imposed by the platform.

Model (3/3)

- Seller's instantaneous payoff:

$$\pi_t = \underbrace{(1-\tau)\, p_t\, (1+F_t)}_{\text{revenue}} - \underbrace{p_t \cdot F_t}_{\text{refund}} - \underbrace{\frac{c}{2} F_t^2}_{\text{other costs}}$$
$$= (1-\tau)\, p_t - \tau p_t \cdot F_t - \frac{c}{2} F_t^2$$

  - τ: transaction fee imposed by the platform.
- The market determines $p_t = E\left[\theta_t | Y_t\right] \equiv M_t$

$$\pi_t = (1-\tau)\, M_t - \tau M_t \cdot F_t - \frac{c}{2} F_t^2$$

Model (3/3)

- Seller's instantaneous payoff:

$$\pi_t = \underbrace{(1-\tau)\,p_t\,(1+F_t)}_{\text{revenue}} - \underbrace{p_t \cdot F_t}_{\text{refund}} - \underbrace{\frac{c}{2}F_t^2}_{\text{other costs}}$$
$$= (1-\tau)\,p_t - \tau p_t \cdot F_t - \frac{c}{2}F_t^2$$

  - τ: transaction fee imposed by the platform.
- The market determines $p_t = E\left[\theta_t | Y_t\right] \equiv M_t$

$$\pi_t = (1-\tau)\,M_t - \tau M_t \cdot F_t - \frac{c}{2}F_t^2$$

  - $\tau = 0$: a. la. Holmstrom (1999); $F_t = \bar{F} > 0$ for all t, at eqm.

Model (3/3)

- Seller's instantaneous payoff:

$$\pi_t = \underbrace{(1-\tau)\, p_t\, (1+F_t)}_{\text{revenue}} - \underbrace{p_t \cdot F_t}_{\text{refund}} - \underbrace{\frac{c}{2} F_t^2}_{\text{other costs}}$$

$$= (1-\tau)\, p_t - \tau p_t \cdot F_t - \frac{c}{2} F_t^2$$

  - τ: transaction fee imposed by the platform.

- The market determines $p_t = E\left[\theta_t | Y_t\right] \equiv M_t$

$$\pi_t = (1-\tau)\, M_t - \tau M_t \cdot F_t - \frac{c}{2} F_t^2$$

  - $\tau = 0$: a. la. Holmstrom (1999); $F_t = \bar{F} > 0$ for all t, at eqm.
  - $\tau > 0$: F_t depends on θ_t and M_t
    - Key: $M_t \cdot F_t$ in the cost term.
      [An alternative micro-foundation is in the paper]

Definition of Equilibrium

Definition (Stationary Linear Markov equilibrium)

A linear Markov strategy $F = (F_t)_{t \geq 0}$ s.t. $F_t = \alpha\theta_t + \beta Y_t + \gamma$ is a stationary linear Markov equilibrium if

1. Buyers take the seller's strategy into account
2. Seller maximizes its own expected discounted value
3. $(\theta_t,\ Y_t)_{t \geq 0}$ induced by F is stationary Gaussian

Definition of Equilibrium

Definition (Stationary Linear Markov equilibrium)

A linear Markov strategy $F = (F_t)_{t \geq 0}$ s.t. $F_t = \alpha\theta_t + \beta Y_t + \gamma$ is a stationary linear Markov equilibrium if

1. $M_t = E^F[\theta_t | Y_t]$
2. Seller maximizes its own expected discounted value
3. $(\theta_t,\ Y_t)_{t \geq 0}$ induced by F is stationary Gaussian

Definition of Equilibrium

Definition (Stationary Linear Markov equilibrium)

A linear Markov strategy $F = (F_t)_{t\geq 0}$ s.t. $F_t = \alpha\theta_t + \beta Y_t + \gamma$ is a stationary linear Markov equilibrium if

1. $M_t = E^F\left[\theta_t | Y_t\right]$
2. $F = \arg\max_{\left(\tilde{F}_t\right)_{t\geq 0}} E_0\left[\int_0^\infty e^{-tr}\left(\left(1-\tau\right)M_t - \tau M_t \cdot \tilde{F}_t - \frac{c}{2}\tilde{F}_t^2\right)dt\right]$
3. $\left(\theta_t,\ Y_t\right)_{t\geq 0}$ induced by F is stationary Gaussian

Definition of Equilibrium

Definition (Stationary Linear Markov equilibrium)

A linear Markov strategy $F = (F_t)_{t \geq 0}$ s.t. $F_t = \alpha\theta_t + \beta Y_t + \gamma$ is a stationary linear Markov equilibrium if

1. $M_t = E^F[\theta_t | Y_t]$
2. $F = \arg\max_{(\tilde{F}_t)_{t\geq 0}} E_0 \left[\int_0^\infty e^{-tr} \left((1-\tau) M_t - \tau M_t \cdot \tilde{F}_t - \frac{c}{2} \tilde{F}_t^2 \right) dt \right]$
3. $(\theta_t,\ Y_t)_{t \geq 0}$ induced by F is stationary Gaussian

- Note: The last condition is not exogenously given.

Stationarity of Equilibrium

- Transition of $(\theta_t,\ Y_t)$ (in discrete analogue):

$$\theta_{t+dt} = \theta_t\left(1 - \kappa dt\right) + \mu\kappa dt + \sigma_\theta dZ_t^\theta$$

$$Y_{t+dt} = Y_t\left(1 - \phi dt\right) + aF_t dt + \theta_t dt + \sigma_\xi dZ_t^\xi$$

Stationarity of Equilibrium

- Transition of $(\theta_t,\ Y_t)$ (in discrete analogue):

$$\theta_{t+dt} = \theta_t\,(1-\kappa dt) + \mu\kappa dt + \sigma_\theta dZ_t^\theta$$

$$Y_{t+dt} = Y_t\,(1-\phi dt) + aF_t dt + \theta_t dt + \sigma_\xi dZ_t^\xi$$

Stationarity of Equilibrium

- Transition of $(\theta_t,\ Y_t)$ (in discrete analogue):

$$\theta_{t+dt} = \theta_t\,(1 - \kappa dt) + \mu\kappa dt + \sigma_\theta dZ_t^\theta$$

$$Y_{t+dt} = Y_t\,(1 - \phi dt) + a(\alpha\theta_t + \beta Y_t + \gamma)dt + \theta_t dt + \sigma_\xi dZ_t^\xi$$

Stationarity of Equilibrium

- Transition of $(\theta_t,\ Y_t)$ (in discrete analogue):

$$\theta_{t+dt} = \theta_t\left(1 - \kappa dt\right) + \mu\kappa dt + \sigma_\theta dZ_t^\theta$$

$$Y_{t+dt} = Y_t\left(1 - \phi dt\right) + a\left(\alpha\theta_t + \beta Y_t + \gamma\right) dt + \theta_t dt + \sigma_\xi dZ_t^\xi$$

Stationarity of Equilibrium

- Transition of $(\theta_t,\ Y_t)$ (in discrete analogue):

$$\theta_{t+dt} = \theta_t\,(1 - \kappa dt) + \mu\kappa dt + \sigma_\theta dZ_t^\theta$$

$$Y_{t+dt} = Y_t\,(1 - (\phi - a\beta)\,dt) + \theta_t\,(1 + a\alpha)\,dt + a\gamma dt + \sigma_\xi dZ_t^\xi$$

Stationarity of Equilibrium

- Transition of $(\theta_t,\ Y_t)$ (in discrete analogue):

$$\theta_{t+dt} = \theta_t \left(1 - \kappa dt\right) + \mu\kappa dt + \sigma_\theta dZ_t^\theta$$

$$Y_{t+dt} = Y_t \left(1 - (\phi - a\beta) dt\right) + \theta_t \left(1 + a\alpha\right) dt + a\gamma dt + \sigma_\xi dZ_t^\xi$$

Stationarity of Equilibrium

- Transition of $(\theta_t,\ Y_t)$ (in discrete analogue):

$$\theta_{t+dt} = \theta_t\left(1 - \kappa dt\right) + \mu\kappa dt + \sigma_\theta dZ_t^\theta$$

$$Y_{t+dt} = Y_t\left(1 - (\phi - a\beta)dt\right) + \theta_t\left(1 + a\alpha\right)dt + a\gamma dt + \sigma_\xi dZ_t^\xi$$

  - $(\theta_t,\ Y_t)$ is stationary Gaussian if $\phi - a\beta > 0$

Stationarity of Equilibrium

- Transition of (θ_t, Y_t) (in discrete analogue):

$$\theta_{t+dt} = \theta_t (1 - \kappa dt) + \mu\kappa dt + \sigma_\theta dZ_t^\theta$$

$$Y_{t+dt} = Y_t (1 - (\phi - a\beta)dt) + \theta_t (1 + a\alpha) dt + a\gamma dt + \sigma_\xi dZ_t^\xi$$

  - (θ_t, Y_t) is stationary Gaussian if $\phi - a\beta > 0$
- When (θ_t, Y_t) is stationary Gaussian, then

$$M_t \equiv E[\theta_t | Y_t] = \underbrace{E[\theta_t]}_{\equiv\mu} + \underbrace{\frac{Cov(\theta_t, Y_t)}{Var(Y_t)}}_{\equiv\lambda}[Y_t - \underbrace{E[Y_t]}_{\equiv\bar{Y}}]$$

Stationarity of Equilibrium

- Transition of $(\theta_t,\ Y_t)$ (in discrete analogue):

$$\theta_{t+dt} = \theta_t\left(1 - \kappa dt\right) + \mu\kappa dt + \sigma_\theta dZ_t^\theta$$

$$Y_{t+dt} = Y_t\left(1 - (\phi - a\beta)dt\right) + \theta_t\left(1 + a\alpha\right)dt + a\gamma dt + \sigma_\xi dZ_t^\xi$$

  - $(\theta_t,\ Y_t)$ is stationary Gaussian if $\phi - a\beta > 0$
- When $(\theta_t,\ Y_t)$ is stationary Gaussian, then

$$M_t \equiv E\left[\theta_t \middle| Y_t\right] = \underbrace{E\left[\theta_t\right]}_{\equiv\mu} + \underbrace{\frac{Cov\left(\theta_t,\ Y_t\right)}{Var\left(Y_t\right)}}_{\equiv\lambda}\left[Y_t - \underbrace{E\left[Y_t\right]}_{\equiv\bar{Y}}\right]$$

Characterize Equilibrium

- HJB equation:

$$
\begin{aligned}
rV(\theta,\ Y) = \sup_{F \in \mathbb{R}} &(1-\tau) M \cdot q - \tau M \cdot F - \frac{c}{2} F^2 \\
&- \kappa (\theta - \mu) V_\theta + \{-\phi Y_t + aF + \theta\} V_Y \\
&+ \frac{\sigma_\theta^2}{2} V_{\theta\theta} + \frac{\sigma_\xi^2}{2} V_{YY} \\
\text{s.t. }\ & M = \mu + \lambda [Y - \bar{Y}]
\end{aligned}
$$

Characterize Equilibrium

- HJB equation:

$$
\begin{aligned}
rV(\theta,\ Y) = \sup_{F\in\mathbb{R}} & (1-\tau)\, M \cdot q - \tau M \cdot F - \frac{c}{2} F^2 \\
& - \kappa(\theta-\mu)\, V_\theta + \{-\phi Y_t + aF + \theta\}\, V_Y \\
& + \frac{\sigma_\theta^2}{2} V_{\theta\theta} + \frac{\sigma_\xi^2}{2} V_{YY} \\
\text{s.t. } & M = \mu + \lambda[Y - \bar{Y}]
\end{aligned}
$$

- The equilibrium is characterized by guess-and-verify of
  - $F = \alpha\theta + \beta Y + \gamma$ (linear strategy)
  - $V = v_0 + v_1\theta + v_2 Y + v_3\theta^2 + v_4 Y^2 + v_5 Y\theta$ (quadratic value function)
  - $\phi - a\beta > 0$ (stationarity)

Theorem (Existence and uniqueness)

There exists a stationary linear Markov equilibrium.

- Reminder: $F_t = \alpha\theta_t + \beta Y_t + \gamma$

Theorem (Existence and uniqueness)

There exists a stationary linear Markov equilibrium. In this equilibrium, $\alpha > 0$, $\beta < 0$, $\lambda > 0$.

- Reminder: $F_t = \alpha\theta_t + \beta Y_t + \gamma$

Theorem (Existence and uniqueness)

There exists a stationary linear Markov equilibrium. In this equilibrium, $\alpha > 0$, $\beta < 0$, $\lambda > 0$. The equilibrium is unique and continuously differentiable in parameters if a loose condition in parameters holds.

- Reminder: $F_t = \alpha\theta_t + \beta Y_t + \gamma$

Theorem (Existence and uniqueness)

There exists a stationary linear Markov equilibrium. In this equilibrium, $\alpha > 0$, $\beta < 0$, $\lambda > 0$. The equilibrium is unique and continuously differentiable in parameters if a loose condition in parameters holds.

- Reminder: $F_t = \alpha\theta_t + \beta Y_t + \gamma$
- Uniqueness holds if $\phi > \kappa$ (rating evolves faster than underlying quality).

Theorem (Existence and uniqueness)

There exists a stationary linear Markov equilibrium. In this equilibrium, $\alpha > 0$, $\beta < 0$, $\lambda > 0$. The equilibrium is unique and continuously differentiable in parameters if a loose condition in parameters holds.

- Reminder: $F_t = \alpha\theta_t + \beta Y_t + \gamma$
- Uniqueness holds if $\phi > \kappa$ (rating evolves faster than underlying quality).
- $\beta < 0$:
  - Driving force: $Y_t \uparrow \Rightarrow p_t \uparrow \Rightarrow$ marginal cost of fake reviews $\uparrow$
  - Countervailing effect: $Y_t \uparrow \Rightarrow Y_{t+\Delta} \uparrow \Rightarrow \frac{\partial V}{\partial Y_{t+\Delta}} \uparrow$ by $\frac{\partial^2 V}{\partial Y^2} > 0$ ▸ Details
- $\alpha > 0$:
  - $\theta_t \uparrow \Rightarrow Y_{t+\Delta} \uparrow \Rightarrow \frac{\partial V}{\partial Y_{t+\Delta}} \uparrow$ by $\frac{\partial^2 V}{\partial Y^2} > 0$

Consistency to data:

Implication to empirical literature:

Consistency to data:

- $\beta < 0$ is consistent with Luca and Zervas (2016)
  - More manipulation after a drop of a rating

Implication to empirical literature:

Consistency to data:

- $\beta < 0$ is consistent with Luca and Zervas (2016)
  - More manipulation after a drop of a rating

Implication to empirical literature:

1. The rating should **not** be used as a proxy for the quality

Consistency to data:

- $\beta < 0$ is consistent with Luca and Zervas (2016)
    - More manipulation after a drop of a rating

Implication to empirical literature:

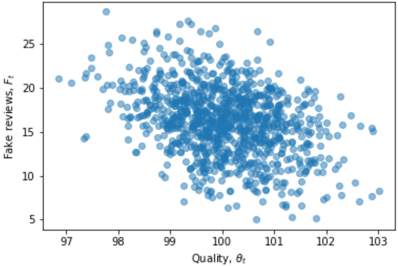
1. The rating should **not** be used as a proxy for the quality
2. Even with true quality index, researchers need to control rating or reputation.

Consistency to data:

- $\beta < 0$ is consistent with Luca and Zervas (2016)
    - More manipulation after a drop of a rating

Implication to empirical literature:

1. The rating should **not** be used as a proxy for the quality
2. Even with true quality index, researchers need to control rating or reputation.

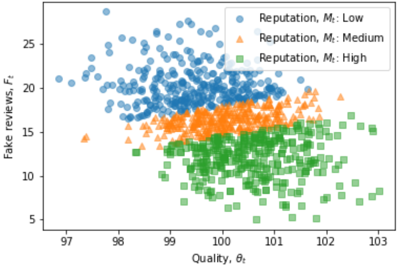


Consistency to data:

- $\beta < 0$ is consistent with Luca and Zervas (2016)
  - More manipulation after a drop of a rating

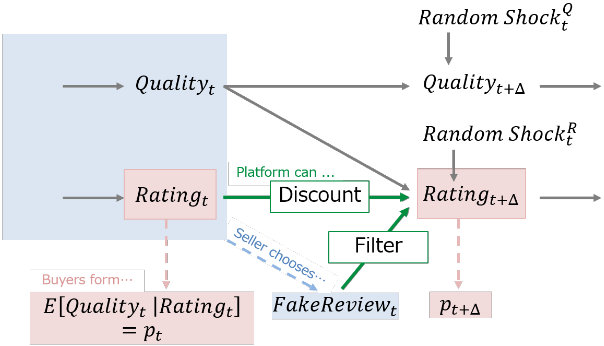
Implication to empirical literature:

1. The rating should **not** be used as a proxy for the quality
2. Even with true quality index, researchers need to control rating or reputation.



Comparative Statics

- Assume that the platform can change
  - a (filtering policy) and
  - ϕ (weights on old/new review)



$$Y_{t+dt} = Y_t(1 - \phi dt) + aF_t dt + \theta_t dt + \sigma_\xi dZ_t^\xi$$

- [Comparative statics about τ and σ_ξ is found in the paper]

Proposition

(i) $E[F_t]$ is increasing in a.
(ii) $a \cdot \alpha$, $a \cdot \beta$, and $a \cdot \gamma$ go to zero as $a \to 0$.

- Reminder: $aF_t = a\alpha\theta_t + a\beta Y_t + a\gamma =$ the effect of fake reviews
- Stringent censorship can **reduce** the **expected amount** and the **effects** of fake reviews.
  - Note: $(\alpha, \beta, \gamma) \nrightarrow 0$ even when $E[F_t] \to 0$ or $(a\alpha, a\beta, a\gamma) \to 0$

Q1: **Should** the platform reduce fake reviews?

Criteria: $\rho^2 = \frac{Cov(\theta_t, Y_t)^2}{Var(\theta_t)Var(Y_t)}$

- Motivation:

Q1: **Should** the platform reduce fake reviews?

Criteria: $\rho^2 = \frac{Cov(\theta_t, Y_t)^2}{Var(\theta_t) Var(Y_t)}$

- Motivation:
- Regulators often want to make rating systems informative.

Q1: **Should** the platform reduce fake reviews?

Criteria: $\rho^2 = \frac{Cov(\theta_t, Y_t)^2}{Var(\theta_t)Var(Y_t)}$

- Motivation:
- Regulators often want to make rating systems informative.
- For the platform, if the rating system is not informative, the sellers and buyers might move out to other platforms.

Q1: **Should** the platform reduce fake reviews?

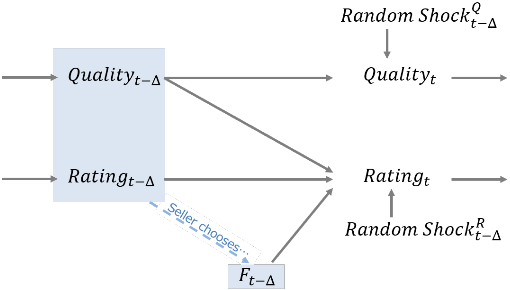
Criteria: $\rho^2 = \frac{Cov(\theta_t, Y_t)^2}{Var(\theta_t)Var(Y_t)}$

- Motivation:
- Regulators often want to make rating systems informative.
- For the platform, if the rating system is not informative, the sellers and buyers might move out to other platforms.
  - Maximization of ρ^2 is equivalent to minimizing $E\left[(p_t - \theta_t)^2\right]$

$$E\left[(p_t - \theta_t)^2\right] = \underbrace{Var(\theta)}_{\text{exogenous}}\left(1 - \rho^2\right)$$

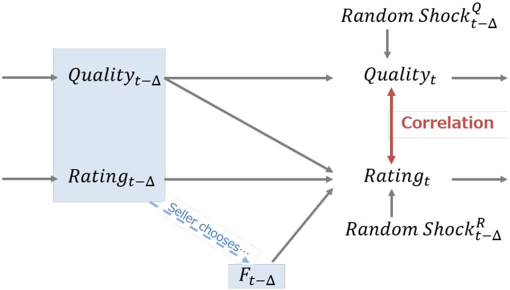
Q1: **Should** the platform reduce fake reviews?

Criteria: $\rho^2 = \frac{(\phi - a\beta)}{(\kappa + \phi - a\beta)} \frac{(a\alpha + 1)^2}{((a\alpha+1)^2 + \kappa(\sigma_\xi/\sigma_\theta)^2(\kappa + \phi - a\beta))}$ (given any α, β, δ)



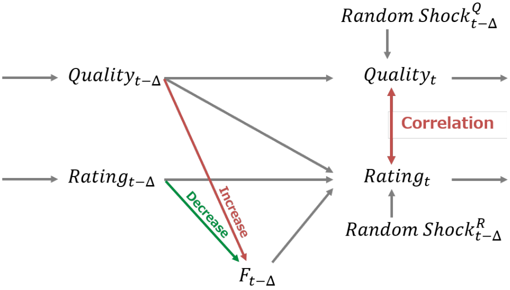
Q1: **Should** the platform reduce fake reviews?

Criteria: $\rho^2 = \frac{(\phi - a\beta)}{(\kappa + \phi - a\beta)} \frac{(a\alpha + 1)^2}{((a\alpha+1)^2 + \kappa(\sigma_\xi/\sigma_\theta)^2(\kappa + \phi - a\beta))}$ (given any α, β, δ)



Q1: **Should** the platform reduce fake reviews?

Criteria: $\rho^2 = \frac{(\phi - a\beta)}{(\kappa + \phi - a\beta)} \frac{(a\alpha + 1)^2}{((a\alpha+1)^2 + \kappa(\sigma_\xi/\sigma_\theta)^2(\kappa + \phi - a\beta))}$ (given any α, β, δ)



- Impacts of the fake reviews:
  1. $a \cdot \alpha > 0$ enhances the positive relationship between the true quality θ_t and the rating Y_t.
  2. $a \cdot \beta < 0$ cancels out the variation in the old rating, $Y_{t-\Delta}$.
     - More discount on old reviews. (ie, Faster transition of the rating)

Proposition (Informativeness of fake reviews)

*The rating **with fake reviews** is **more informative** than one without when*

1. *a is sufficiently large, or*
2. *(i) a is sufficiently small and*
   (ii) ϕ is too small

- The first effect (from $a \cdot \alpha > 0$) dominates for large a, and
- The second effect (from $a \cdot \beta < 0$) dominates for small a.
  - The second effect is good if ϕ is too small

Proposition (Informativeness of fake reviews)

*The rating **with fake reviews** is **more informative** than one without when*

1. *a is sufficiently large, or*
2. *(i) a is sufficiently small and*
   (ii) ϕ is too small

- The first effect (from $a \cdot \alpha > 0$) dominates for large a, and
- The second effect (from $a \cdot \beta < 0$) dominates for small a.
  - The second effect is good if ϕ is too small

Proposition (Informativeness of fake reviews)

*The rating **with fake reviews** is **more informative** than one without when*

1. *a is sufficiently large, or*
2. *(i) a is sufficiently small and*
   (ii) ϕ is too small

- The first effect (from $a \cdot \alpha > 0$) dominates for large a, and
- The second effect (from $a \cdot \beta < 0$) dominates for small a.
  - The second effect is good if ϕ is too small

Proposition (Informativeness of fake reviews)

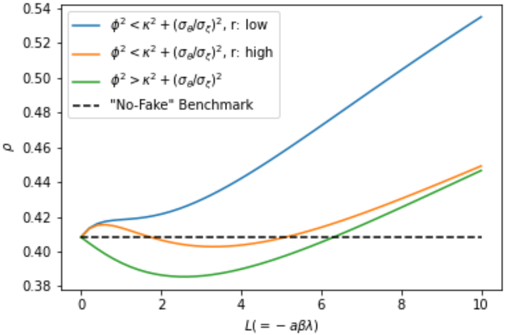
*The rating **with fake reviews** is **more informative** than one without when*

1. *a is sufficiently large, or*
2. *(i) a is sufficiently small and*
   (ii) $\phi^2 < \frac{\sigma_\theta^2}{\sigma_\xi^2} + \kappa^2$

- The first effect (from $a \cdot \alpha > 0$) dominates for large a, and
- The second effect (from $a \cdot \beta < 0$) dominates for small a.
  - The second effect is good if ϕ is too small

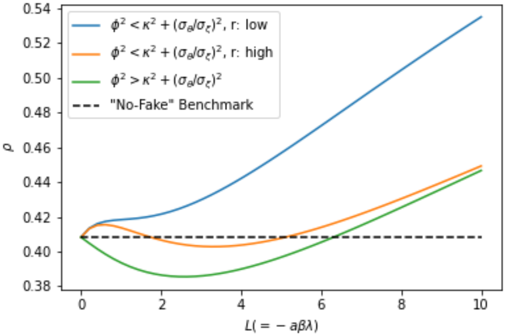
Sketch of the proof:

- $\rho^2 = \frac{(\phi - a\beta)}{(\kappa + \phi - a\beta)} \frac{(a\alpha+1)^2}{((a\alpha+1)^2 + \kappa(\sigma_\xi/\sigma_\theta)^2(\kappa+\phi-a\beta))}$ (given any α, β, δ)



Sketch of the proof:

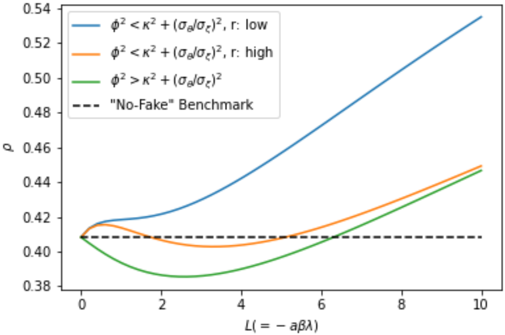
- $\rho^2 = \frac{(\phi+L)}{(\kappa+\phi+L)} \frac{(A(L)+1)^2}{((A(L)+1)^2+\kappa(\sigma_\xi/\sigma_\theta)^2(\kappa+\phi+L))} = \rho^2\left(L,\, \phi\right)$ (given **eqm** α, β, δ)
  - L (eqm effect on the transition speed) is positive and increasing in a.



Sketch of the proof:

- $\rho^2 = \frac{(\phi+L)}{(\kappa+\phi+L)} \frac{(A(L)+1)^2}{((A(L)+1)^2+\kappa(\sigma_\xi/\sigma_\theta)^2(\kappa+\phi+L))} = \rho^2(L, \phi)$ (given **eqm** α, β, δ)
  - L (eqm effect on the transition speed) is positive and increasing in a.

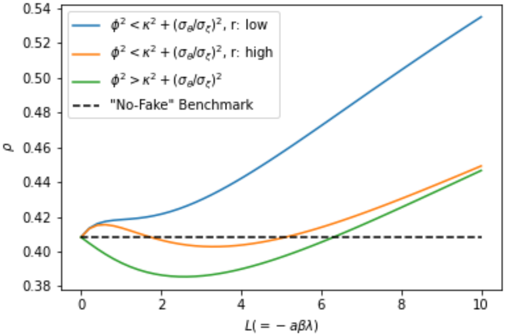
1. $\lim_{L\to\infty} \rho^2 = 1$



Sketch of the proof:

- $\rho^2 = \frac{(\phi+L)}{(\kappa+\phi+L)} \frac{(A(L)+1)^2}{((A(L)+1)^2+\kappa(\sigma_\xi/\sigma_\theta)^2(\kappa+\phi+L))} = \rho^2\left(L,\, \phi\right)$ (given **eqm** α, β, δ)
  - L (eqm effect on the transition speed) is positive and increasing in a.

1. $\lim_{L\to\infty} \rho^2 = 1$
2. $\frac{\partial \rho^2}{\partial L}\big|_{L=0} > 0$ iff $\phi^2 < \frac{\sigma_\theta^2}{\sigma_\xi^2} + \kappa^2$



Q2: How should the platform adjust ϕ?

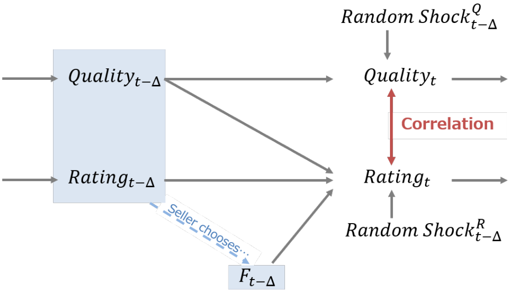
- ϕ: transition speed of the rating, relative weight on new reviews

Q2: How should the platform adjust ϕ?

- ϕ: transition speed of the rating, relative weight on new reviews
- Comparison with the optimal ϕ without fake reviews, ϕ^0

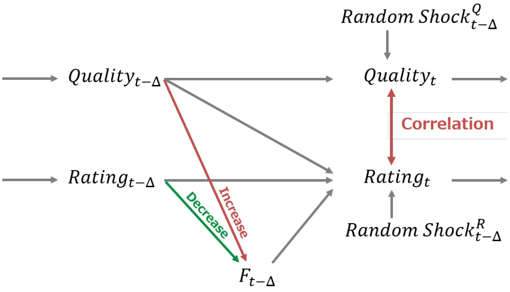
Q2: How should the platform adjust ϕ?

- ϕ: transition speed of the rating, relative weight on new reviews
- Comparison with the optimal ϕ without fake reviews, ϕ^0
  - Higher ϕ,
    - faster update on the underlying quality
    - less robust to the random shocks
  - Platform choose ϕ^0 to balance those effects.



Q2: How should the platform adjust ϕ?

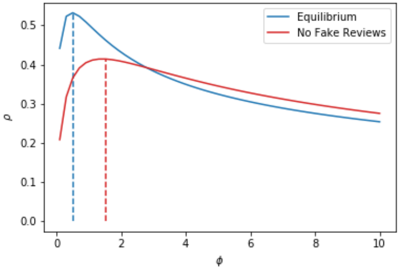
- ϕ: transition speed of the rating, relative weight on new reviews
- Comparison with the optimal ϕ without fake reviews, ϕ^0
  - Higher ϕ,
    - faster update on the underlying quality
    - less robust to the random shocks
  - Platform choose ϕ^0 to balance those effects.



Proposition

(i) At eqm (with fake reviews), ρ^2 is decreasing in ϕ at $\phi = \phi^0$.

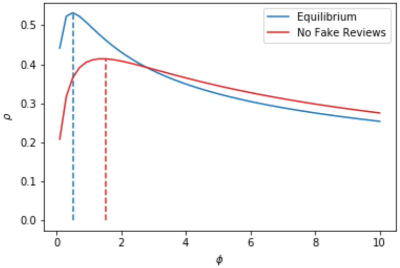
- (i) w/ fake reviews: effective transition speed is $\phi - a\beta$
  - $\rightarrow$ the platform should adjust ϕ downward.



Proposition

(i) At eqm (with fake reviews), ρ^2 is decreasing in ϕ at $\phi = \phi^0$.
(ii) Furthermore, for sufficiently small r, the maximum of ρ^2 with fake reviews is higher than without.

- (i) w/ fake reviews: effective transition speed is $\phi - a\beta$
  - $\rightarrow$ the platform should adjust ϕ downward.
- (ii) small $r \Rightarrow$ high weight on the future$\Rightarrow$ high $\alpha \Rightarrow$ rating is informative with fake review, given ϕ^0
  - The platform can further adjust ϕ from ϕ^0 .



Naive Consumers

- Naive consumers believe that
- Reputation:
- Seller's payoff:

Naive Consumers

- Naive consumers believe that
  - they face a stationary Gaussian distribution of (θ_t, Y_t)
- Reputation:
- Seller's payoff:

Naive Consumers

- Naive consumers believe that
  - they face a stationary Gaussian distribution of $(\theta_t,\ Y_t)$
  - there is no fake reviews by the seller. (ie, assume $\alpha = \beta = \gamma = 0$)
- Reputation:
- Seller's payoff:

Naive Consumers

- Naive consumers believe that
  - they face a stationary Gaussian distribution of $(\theta_t,\ Y_t)$
  - there is no fake reviews by the seller. (ie, assume $\alpha = \beta = \gamma = 0$)
- Reputation:
  - rational consumers: $M_t = E^F\left[\theta_t | Y_t\right] = \mu + \lambda\left(\alpha,\ \beta\right)\left[Y_t - \bar{Y}\left(\alpha,\ \beta,\ \gamma\right)\right]$
- Seller's payoff:

Naive Consumers

- Naive consumers believe that
  - they face a stationary Gaussian distribution of $(\theta_t,\, Y_t)$
  - there is no fake reviews by the seller. (ie, assume $\alpha = \beta = \gamma = 0$)
- Reputation:
  - rational consumers: $M_t = E^F\left[\theta_t | Y_t\right] = \mu + \lambda(\alpha,\, \beta)\left[Y_t - \bar{Y}(\alpha,\, \beta,\, \gamma)\right]$
  - naive consumers: $\widetilde{M}_t = \widetilde{E}\left[\theta_t | Y_t\right] = \mu + \lambda(0,\, 0)\left[Y_t - \bar{Y}(0,\, 0,\, 0)\right]$
    - belief based an wrong joint distribution of $(\theta_t,\, Y_t)$
- Seller's payoff:

Naive Consumers

- Naive consumers believe that
  - they face a stationary Gaussian distribution of $(\theta_t,\, Y_t)$
  - there is no fake reviews by the seller. (ie, assume $\alpha = \beta = \gamma = 0$)
- Reputation:
  - rational consumers: $M_t = E^F\left[\theta_t | Y_t\right] = \mu + \lambda\left(\alpha,\, \beta\right)\left[Y_t - \bar{Y}\left(\alpha,\, \beta,\, \gamma\right)\right]$
  - naive consumers: $\widetilde{M}_t = \widetilde{E}\left[\theta_t | Y_t\right] = \mu + \lambda\left(0,\, 0\right)\left[Y_t - \bar{Y}\left(0,\, 0,\, 0\right)\right]$
    - belief based an wrong joint distribution of $(\theta_t,\, Y_t)$
- Seller's payoff:
  - $\pi_t = (1 - \tau)\, p_t - \tau p_t \cdot F_t - \frac{c}{2} F_t^2$

Naive Consumers

- Naive consumers believe that
  - they face a stationary Gaussian distribution of $(\theta_t,\, Y_t)$
  - there is no fake reviews by the seller. (ie, assume $\alpha = \beta = \gamma = 0$)
- Reputation:
  - rational consumers: $M_t = E^F\left[\theta_t | Y_t\right] = \mu + \lambda(\alpha,\, \beta)\left[Y_t - \bar{Y}(\alpha,\, \beta,\, \gamma)\right]$
  - naive consumers: $\widetilde{M}_t = \widetilde{E}\left[\theta_t | Y_t\right] = \mu + \lambda(0,\, 0)\left[Y_t - \bar{Y}(0,\, 0,\, 0)\right]$
    - belief based an wrong joint distribution of $(\theta_t,\, Y_t)$
- Seller's payoff:
  - $\pi_t = (1-\tau)\, p_t - \tau p_t \cdot F_t - \frac{c}{2} F_t^2$
  - $p_t = \eta M_t + (1-\eta)\, \widetilde{M}_t$ where $\eta \in [0,\, 1]$

Naive Consumers

- Naive consumers believe that
  - they face a stationary Gaussian distribution of $(\theta_t,\, Y_t)$
  - there is no fake reviews by the seller. (ie, assume $\alpha = \beta = \gamma = 0$)
- Reputation:
  - rational consumers: $M_t = E^F\left[\theta_t | Y_t\right] = \mu + \lambda\left(\alpha,\, \beta\right)\left[Y_t - \bar{Y}\left(\alpha,\, \beta,\, \gamma\right)\right]$
  - naive consumers: $\widetilde{M}_t = \widetilde{E}\left[\theta_t | Y_t\right] = \mu + \lambda\left(0,\, 0\right)\left[Y_t - \bar{Y}\left(0,\, 0,\, 0\right)\right]$
    - belief based an wrong joint distribution of $(\theta_t,\, Y_t)$
- Seller's payoff:
  - $\pi_t = \left(1 - \tau\right) p_t - \tau p_t \cdot F_t - \frac{c}{2} F_t^2$
  - $p_t = \eta M_t + \left(1 - \eta\right) \widetilde{M}_t$ where $\eta \in [0,\, 1]$
  - Interpretation:

Naive Consumers

- Naive consumers believe that
  - they face a stationary Gaussian distribution of (θ_t, Y_t)
  - there is no fake reviews by the seller. (ie, assume $\alpha = \beta = \gamma = 0$)
- Reputation:
  - rational consumers: $M_t = E^F[\theta_t | Y_t] = \mu + \lambda(\alpha, \beta)\left[Y_t - \bar{Y}(\alpha, \beta, \gamma)\right]$
  - naive consumers: $\widetilde{M}_t = \widetilde{E}[\theta_t | Y_t] = \mu + \lambda(0, 0)\left[Y_t - \bar{Y}(0, 0, 0)\right]$
    - belief based an wrong joint distribution of (θ_t, Y_t)
- Seller's payoff:
  - $\pi_t = (1-\tau)p_t - \tau p_t \cdot F_t - \frac{c}{2}F_t^2$
  - $p_t = \eta M_t + (1-\eta)\widetilde{M}_t$ where $\eta \in [0, 1]$
  - Interpretation:
    - η captures the rationality of each consumer

Naive Consumers

- Naive consumers believe that
  - they face a stationary Gaussian distribution of (θ_t, Y_t)
  - there is no fake reviews by the seller. (ie, assume $\alpha = \beta = \gamma = 0$)
- Reputation:
  - rational consumers: $M_t = E^F[\theta_t | Y_t] = \mu + \lambda(\alpha, \beta)\left[Y_t - \bar{Y}(\alpha, \beta, \gamma)\right]$
  - naive consumers: $\widetilde{M}_t = \widetilde{E}[\theta_t | Y_t] = \mu + \lambda(0, 0)\left[Y_t - \bar{Y}(0, 0, 0)\right]$
    - belief based an wrong joint distribution of (θ_t, Y_t)
- Seller's payoff:
  - $\pi_t = (1 - \tau) p_t - \tau p_t \cdot F_t - \frac{c}{2} F_t^2$
  - $p_t = \eta M_t + (1 - \eta)\widetilde{M}_t$ where $\eta \in [0, 1]$
  - Interpretation:
    - η captures the rationality of each consumer
    - η is the ratio of rational consumers in the market. ▸ Details

Theorem

Existence and uniqueness given the same condition as the baseline model

Theorem

Existence and uniqueness given the same condition as the baseline model

Proposition

Existence of the naive consumers decreases $E[F_t]$.

Theorem

Existence and uniqueness given the same condition as the baseline model

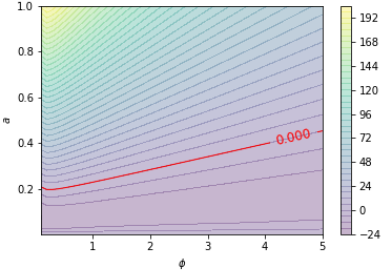
Proposition

Existence of the naive consumers decreases $E[F_t]$.

- Intuition:
  - Naive consumers set higher price, but
  - **Rational** consumers are **more sensitive to the rating** than naive consumers.
    - Rational consumers takes $a\alpha > 0$ into account.
  - Less marginal benefit with naive consumers.
  - Less fake reviews with naive consumers.

Criteria for the naive consumers:

$$Bias = E\left[\widetilde{E}\left[\theta_t | Y_t\right] - \theta_t\right]$$

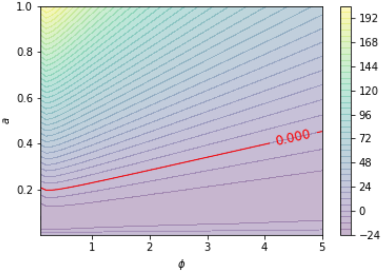


Criteria for the naive consumers:

$$Bias = E\left[\widetilde{E}\left[\theta_t|Y_t\right] - \theta_t\right]$$

Lemma

$Bias \geq 0$ iff $E\left[F_t\right] \geq 0$.



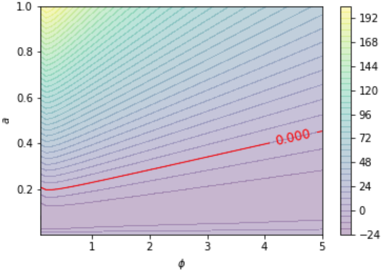
Criteria for the naive consumers:

$$Bias = E\left[\widetilde{E}\left[\theta_t | Y_t\right] - \theta_t\right]$$

Lemma

$Bias \geq 0$ iff $E\left[F_t\right] \geq 0$.

Suppose there are only naive consumers in the market.



Criteria for the naive consumers:

$$Bias = E\left[\widetilde{E}\left[\theta_t | Y_t\right] - \theta_t\right]$$

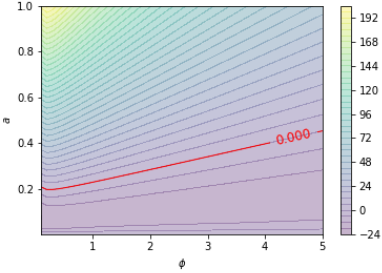
Lemma

$Bias \geq 0$ iff $E\left[F_t\right] \geq 0$.

Suppose there are only naive consumers in the market.

Proposition

A strict filtering policy reduces $Bias$ as long as $Bias \geq 0$.



Criteria for the naive consumers:

$$Bias = E\left[\widetilde{E}\left[\theta_t | Y_t\right] - \theta_t\right]$$

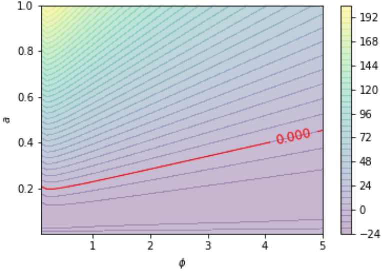
Lemma

$Bias \geq 0$ iff $E\left[F_t\right] \geq 0$.

Suppose there are only naive consumers in the market.

Proposition

A strict filtering policy reduces Bias as long as $E\left[F_t\right] \geq 0$.



Summary

Positive Analysis:

Normative Analysis:

Summary

Positive Analysis:

- The number of fake reviews is increasing in quality, decreasing in reputation.
- The stringent censorship
  - reduces fake reviews in expectation, but
  - reduces the effects of fake reviews.

Normative Analysis:

Summary

Positive Analysis:

- The number of fake reviews is increasing in quality, decreasing in reputation.
- The stringent censorship
  - reduces fake reviews in expectation, but
  - reduces the effects of fake reviews.

Normative Analysis:

- For rational consumers:
- For naive consumers:

Summary

Positive Analysis:

- The number of fake reviews is increasing in quality, decreasing in reputation.
- The stringent censorship
  - reduces fake reviews in expectation, but
  - reduces the effects of fake reviews.

Normative Analysis:

- For rational consumers:
  - a rating with fake reviews can be more informative than without fake reviews
  - Transition speed of the rating should be slower than the optimal level without fake reviews.
- For naive consumers:

Summary

Positive Analysis:

- The number of fake reviews is increasing in quality, decreasing in reputation.
- The stringent censorship
  - reduces fake reviews in expectation, but
  - reduces the effects of fake reviews.

Normative Analysis:

- For rational consumers:
  - a rating with fake reviews can be more informative than without fake reviews
  - Transition speed of the rating should be slower than the optimal level without fake reviews.
- For naive consumers:
  - As long as $E[F_t] \geq 0$, the more stringent censorship, the less bias for the naive consumers.

Intuition of the Equilibrium Strategy

Back to Theorem

- Reminder: $V = v_0 + v_1\theta + v_2 Y + v_3\theta^2 + v_4 Y^2 + v_5\theta Y$
- FOC: $F_t = -\frac{\tau}{c}M_t + \frac{a}{c}\{v_2 + 2Y_t v_4 + \theta v_5\}$
- $\beta < 0$
  - $\beta = -\frac{\tau}{c}\lambda + 2\frac{a}{c}v_4 = -\frac{\tau}{c}\lambda + \frac{a}{c}\frac{-\beta\lambda\tau}{(-a\beta+r+2\phi)}$
  - $\beta < 0$ since today's cost saving incentive dominates.
- $v_4 = \frac{-\beta\lambda\tau}{2(-a\beta\lambda+r+2\phi)} > 0$
  - Higher reputation, less promotion, less costly: $\tau M_t F_t = \tau\alpha\theta_t Y_t + \tau\beta Y_t^2 + \text{constant}$
- $\alpha > 0$
  - $\alpha = \frac{a}{c}v_5$
  - $v_5 = \frac{1}{\kappa+r+\phi}\{ \underbrace{2(a\alpha + bq)}_{\text{High High future } Y_t} v_4 - \alpha\lambda\tau\}$
    - Driving Force: Higher θ today, higher Y in the future, value is quadratically increasing in Y.
    - Counteracting effect: Higher quality, higher $F_{t+\Delta}$(if $\alpha > 0$). Less complementarity from $-\tau M_t F_t$.

Microfoundation of the price: $p_t = M_t$

Back to Model

- (Reminder: $M_t \equiv E\left[\theta_t | Y_t\right]$)
- Suppose there is a mass (2) of buyers.
- Consumer $i \in [0, 2]$ feels

$$u_i = \begin{cases} \theta + \epsilon_i - p & \text{if the consumer purchase the product} \\ 0 & \text{otherwise} \end{cases}$$

  - $\epsilon_i \sim_{i.i.d.} F(\cdot)$ where $F(\cdot)$ is symmetric around zero
- Given Y, rational consumer purchases iff $M + \epsilon_i - p \geq 0$
- Market clearing

$$\begin{aligned} 1 &= 2q \cdot (1 - F(p - M)) \\ \Leftrightarrow p &= M \end{aligned}$$

Mixture of the rational/naive consumers

Back to Model

- $M = E[\theta|Y]$: rational consumer's belief (on the seller's quality)
- $\widetilde{M} = \widetilde{E}[\theta|Y]$: naive consumer's belief (on the seller's quality)

Rationale:

- 2η rational consumers and $2(1-\eta)$ naive consumers in mkt
- Consumer $i \in [0, 2]$ feels

$$u_i = \begin{cases} \theta + \epsilon_i - p & \text{if the consumer purchase the product} \\ 0 & \text{otherwise} \end{cases}$$

  - $\epsilon_i \sim U(-C, C)$: iid over the consumer types.
- Rational consumer purchases iff $M_t + \epsilon_i - p \geq 0$
- naive consumer purchases iff $\widetilde{M}_t + \epsilon_i - p \geq 0$
- Market clearing

$$1 = 2\eta \cdot (1 - F(p - M)) + 2(1-\eta) \cdot \left(1 - F\left(p - \widetilde{M}\right)\right)$$

$$\Leftrightarrow p = \eta M + (1-\eta)\widetilde{M}$$