

# レビューの不正操作に関するサーベイ

坂口洋英

慶應義塾大学

➤ オンライン上の口コミやレビューは消費者に情報源として大きな便益を与えている

➤ 消費者の75%はレビューを参照(総務省,2012)

➤ レビュー・口コミの影響

➤ Chevalier and Mayzlin (2006)から続く1トピック

➤ 売上にレビューが果たす役割は様々な研究で実証されている

➤ 一方で、日本では2010年代からこれらの不正操作が社会問題化

➤ ペニーオークション詐欺事件

➤ 食べログやらせ事件

➤ 近年ではAmazonにおける“やらせ”・“サクラ”のレビューが問題視されている

➤ フェイクレビューと呼称される

➤ NHKで特集が組まれることも

➤ 「Amazon レビュー」の検索で、1ページ目にサクラ関連記事が提示



<https://securitynews.so-net.ne.jp> トピックス ▾

Amazonなどの「サクラレビュー（やらせ）」を見破る方法と ...

2021/07/02 — やらせレビューを見破るサクラチェッカー。商品を手に取ってみることができないネット通販は、レビューだけが頼り。特にAmazonのレビュー欄は、...

<https://diamond.jp> > ... > それ、ネット詐欺です! ▾

ニセの口コミに騙された...アマゾンの「サクラレビュー」を ...

2021/09/24 — AmazonなどのECサイトで商品を選ぶときに、多くの人が頼りにするのが、購入した人たちの口コミやユーザーレビュー。ところが時々、明らかなウソが...

<https://news.mynavi.jp> > ... > スマホアプリ/サービス ▾

Amazonレビューの“やらせ度”を表示、無料のスマホアプリは ...

2021/06/15 — Amazonで販売されている商品レビューの信用度を分析して表示する無料のスマホアプリ「サードモールド」が登場しました。表示されるスコアおよび評価は...

Amazonのアプリから簡単に呼び... 実際に「やらせ」の商品をチェック...

- ▶ レビューの不正操作がもたらす被害
  - ▶ 消費者が欺かれ、望まない消費を行う恐れがある
  - ▶ 消費者のレビューという仕組み自体への信頼を失わせる
    - ▶ 財の選択に係るサーチコストを増大させ、消費者と企業両方に被害をもたらすことが懸念される
- ▶ 米国等では明確に違法とされるものの、国内では景品表示法のガイドライン改訂にとどまっている。
  - ▶ FTC (連邦取引委員会) は違反一回あたり最大\$43,792を課している
  - ▶ 法曹界からは規制強化の意見書が提出(日本弁護士連合会,2017)
- ▶ 規制にあたっては、不正操作による被害の定量化が必要。
- ▶ フェイクレビューがもたらす被害はフェイクを行う主体と消費者に依存
  - ▶ 悪徳業者が粗悪品を売るためにフェイクレビューを使う: 被害が大きい
  - ▶ 質がいいが知名度の低い企業がフェイクレビューを宣伝に使う: 誰も損しない
    - ▶ 直観的ではないものの、支持する理論研究も存在(Dellarocas,2006;Yasui,2020)
  - ▶ 消費者が騙されやすい場合: 被害が大きい
  - ▶ 消費者がフェイクレビューの存在を割り引いて考える場合: 被害が小さい
- ▶ 先行研究は海外においてある程度蓄積されているものの、特に国内に関しては、エビデンスの蓄積は不十分
- ▶ 規制を議論するうえでは、これらに関するエビデンスが不可欠

- どうやってレビューがフェイクだと判別するのか
  - 個別レベルで分かっているものは存在しない
  - 代理変数(Mayzlin et al.,2014;Luca and Zervas,2016)
  - 商品単位で把握 (He et al., 2021)
- どんな企業がフェイクレビューを行っているのか
  - 競合相手のいる企業(Mayzlin et al.,2014;Luca and Zervas,2016)
  - フェイクレビューのコストが小さい企業(Mayzlin et al.,2014;Luca and Zervas,2016)
    - 間接的に質の低い企業の可能性
  - 評価が落ちたタイミングで高評価のフェイクレビューで自社を高く評価されているように装う企業(Luca and Zervas, 2016;He et al., 2021)
    - より直接的に質の低い企業の可能性
- どんな影響があるのか
  - フェイクレビューにより商品のセールスランクが上昇(He et al.,2021)
  - フェイクレビューにより低評価をつけるユーザーが増加(He et al.,2021)
    - 被害が生じている可能性

- 米国ホテル産業におけるレビューを分析
  - Tripadvisor(誰でもレビューが書ける)
  - Expedia(泊まった消費者しかレビューを書けない)
- Tripadvisorへのフェイクレビューを投稿コストが相対的に低い
- Tripadvisor(TA)とExpedia(EXP)の得点分布の差に注目

$$\frac{Nstar\ Reviews_{ij}^{TA}}{Total\ Reviews_{ij}^{TA}} - \frac{Nstar\ Reviews_{ij}^{Exp}}{Total\ Reviews_{ij}^{TA}}$$

- TAに相対的に多くフェイクレビューが投稿されるなら左辺が大きくなる。
  - 自社への高評価:Nが5や4の時に左辺が大きくなる
  - 他社から自社への攻撃:Nが1や2のときに左辺が大きくなる
- 結果、
  - ブランドイメージを気にしないでいい非チェーン系ホテルほど自社を持ち上げる正のフェイクレビューを行いやすい
  - 他のホテルのイメージを気にするマルチユニットオーナーほど自社への正のフェイクレビューを行いにくい
  - 同じ地域に競合相手がいると低評価のフェイクレビューによる攻撃を受けやすい
    - その相手がブランドイメージを気にしないでいい独立系だとさらに攻撃を受けやすい
    - 他のホテルのイメージを気にするマルチユニットオーナーだと攻撃がやわらぐ

- ▶ 米国レストラン産業におけるレビューを分析
  - ▶ Yelpはフィルタリングしたレビューを公開
- ▶ フィルタリングされたレビューはフェイクレビューである可能性が高い
  - ▶ フィルタリングとは別におとり捜査に捕捉されたレストランのレビューの大半がフィルタリングされていた
  - ▶ 代理変数として妥当性が担保される
  - ▶ 固定効果とフィルタリングされやすいレビューの属性をコントロールすることで対処

$$\underbrace{f_{jt}}_{\text{フィルタリングされたレビューの数}} = \alpha_0 n_{jt} + x'_{jt} \tilde{\beta} + \underbrace{\tilde{b}_j}_{\text{固定効果}} + \tilde{\mu}_t + \tilde{\epsilon}_{jt} + \sum_i \underbrace{z'_{ijt}}_{\text{レビューの属性}} \gamma + \hat{u}_{ijt}$$

- ▶ 結果、
  - ▶ ブランドイメージを気にしないでいい非チェーン系レストランほど自社を持ち上げる正のフェイクレビューを行いやすい
  - ▶ 評価が落ちたタイミングで正のフェイクレビューが行われやすい
    - ▶ 経済的インセンティブに基づく行動
  - ▶ 一つあたりのレビュー効果が薄くなるレビューの数が多いレストランはフェイクレビューを投稿しにくく、攻撃されにくい
  - ▶ 同じ地域に同じ種類の料理のレストランがいると低評価のフェイクレビューによる攻撃を受けやすい
  - ▶ 同じ地域にチェーン店や違う種類のレストランがいると低評価をのフェイクレビューの攻撃を受けにくい

- Amazon.com(米Amazon)におけるレビューを分析
- Facebookのレビュー依頼コミュニティをRAにより調査
- フェイクレビューを募集している商品を把握
  - 比較的発売から時間が経っていない商品が多い
    - 新製品ではない
  - 競合製品と比べ安い
  - 中国セラーが8割

- フェイクレビューの募集の前後の変化を把握
  - 評価は一時的に上がるものの長期的に下がる
  - 低評価のレビューをつけるユーザーが長期的に増加

- 自然実験を利用してDIDによりフェイクレビューが売り上げを伸ばしていることを実証
  - 極端に削除の多い時期がある
  - フェイクレビューがすぐ削除されてしまう
  - コントロールグループとして利用

Figure 4: 7-day average ratings, 7-day average number of reviews, and cumulative average ratings before and after fake reviews recruiting begins. The red dashed line indicates the last week of data before we observe Facebook fake review recruiting.

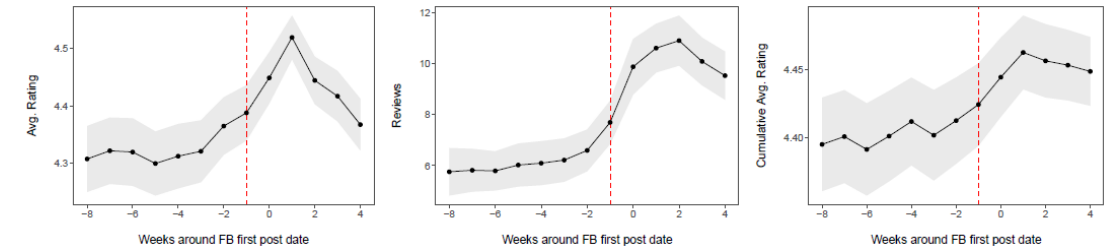


Figure 8: 7-day average number of average ratings, reviews, and average share of one-star reviews before and after fake reviews recruiting stops. The red dashed line indicates the last week of data in which we observe Facebook fake review recruiting.

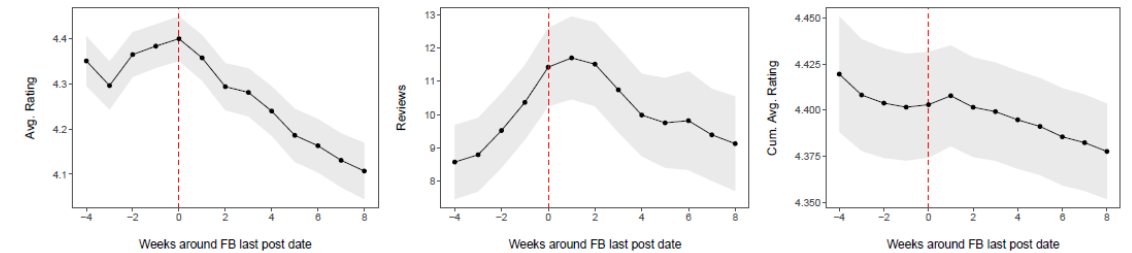
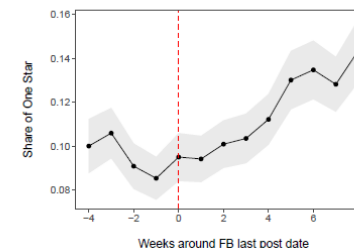


Figure 14: 7-day average share of one-star reviews before and after fake reviews recruiting stops. The red dashed line indicates the last time we observe Facebook fake review recruiting.



- Amazon.co.jp(日本Amazon)におけるレビューを分析
- 高頻度でデータを取得 & 削除されたレビューを代理変数
  - 日次レベルで取得
    - 2万以上の商品、約120万のレビュー
    - サバイバルバイアス排除のためリアルタイムで観測できたもののみ使用
  - Luca and Zervas(2016)を踏襲
  - ある程度妥当性は担保される(He et al.,2021)
- 削除されやすいレビューを把握
  - 高評価なもの
  - 発売から早期に投稿されるもの
  - 文章やタイトルが長く、写真がついているもの
  - ひらがな率が高いもの
  - デフォルトアイコン
    - もっともらしく見える特徴
    - フェイクレビューが巧妙化している可能性

	(1)	(2)
	deleted	deleted
*5	0.04*** (14.93)	0.04*** (15.73)
*4	0.01*** (3.31)	0.01*** (3.48)
*2	-0.00 (-0.72)	-0.00 (-1.62)
*1	0.00 (0.29)	0.00 (0.08)
得票数	-0.00*** (-5.29)	-0.00*** (-4.56)
購入者	0.03*** (6.60)	0.03*** (7.40)
写真あり	0.04*** (9.26)	0.04*** (10.14)
レビューア投稿数	-0.00* (-2.08)	-0.00* (-1.96)
レビューア得票数	-0.00*** (-6.55)	-0.00*** (-6.83)
レビューテキストひらがな率	0.03*** (8.29)	0.03*** (8.42)
アイコンがデフォルトのもの	-0.03*** (-13.64)	-0.03*** (-14.50)
レビューテキストの長さ	0.00*** (8.93)	0.00*** (9.06)
レビュータイトルの長さ	0.00*** (5.38)	0.00*** (4.78)
投稿時の発売からの経過週数 (レビュー基準)	-0.00*** (-3.79)	
投稿時の発売からの経過週数 (取り扱い開始日基準)		-0.00*** (-3.82)
切片	0.15*** (13.59)	0.15*** (13.80)
<i>N</i>	308146	343244

*t* statistics in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$



- Amazon.co.jp(日本Amazon)におけるレビューを分析
- 高頻度でデータを取得 & 削除されたレビューを代理変数
  - 日次レベルで取得
    - 2万以上の商品、約120万のレビュー
    - サバイバルバイアス排除のためリアルタイムで観測できたもののみ使用
  - Luca and Zervas(2016)の代理変数・モデルを踏襲
  - ある程度妥当性は担保される(He et al.,2021)
- 削除されやすいレビューを把握
  - 高評価なもの
  - 発売から早期に投稿されるもの
  - 文章やタイトルが長く、写真がついているもの
  - ひらがな率が高いもの
  - デフォルトアイコン
    - もっともらしく見える特徴
    - フェイクレビューが巧妙化している可能性

	(1)	(2)
	deleted	deleted
*5	0.04*** (14.93)	0.04*** (15.73)
*4	0.01*** (3.31)	0.01*** (3.48)
*2	-0.00 (-0.72)	-0.00 (-1.62)
*1	0.00 (0.29)	0.00 (0.08)
得票数	-0.00*** (-5.29)	-0.00*** (-4.56)
購入者	0.03*** (6.60)	0.03*** (7.40)
写真あり	0.04*** (9.26)	0.04*** (10.14)
レビューア投稿数	-0.00* (-2.08)	-0.00* (-1.96)
レビューア得票数	-0.00*** (-6.55)	-0.00*** (-6.83)
レビューテキストひらがな率	0.03*** (8.29)	0.03*** (8.42)
アイコンがデフォルトのもの	-0.03*** (-13.64)	-0.03*** (-14.50)
レビューテキストの長さ	0.00*** (8.93)	0.00*** (9.06)
レビュータイトルの長さ	0.00*** (5.38)	0.00*** (4.78)
投稿時の発売からの経過週数 (レビュー基準)	-0.00*** (-3.79)	
投稿時の発売からの経過週数 (取り扱い開始日基準)		-0.00*** (-3.82)
切片	0.15*** (13.59)	0.15*** (13.80)
N	308146	343244

t statistics in parentheses

\* p < 0.05, \*\* p < 0.01, \*\*\* p < 0.001

- フェイクレビューを行いやすい商品
  - Mayzlin et al.(2014)を踏襲
  - 質の低いもの
  - 中国系セラーによるもの

	(1)	(2)
	mayzline	mayzline
* 1-3(非削除)	0.00*	
	(2.29)	
* 4-5(非削除)		-0.00***
		(-3.48)
メーカーストアあり	-0.08***	-0.07***
	(-5.35)	(-4.84)
中国系セラー	0.07***	0.07***
	(4.63)	(4.64)
セラー評価(点)	-0.00	-0.00
	(-0.92)	(-0.65)
切片	0.18*	0.18*
	(2.35)	(2.29)
<i>N</i>	2539	2539

*t* statistics in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

- パネルデータ分析により
  - 評判を落とすタイミングで高評価の削除された怪しいレビューが増加
  - 初期に怪しいレビューが増加

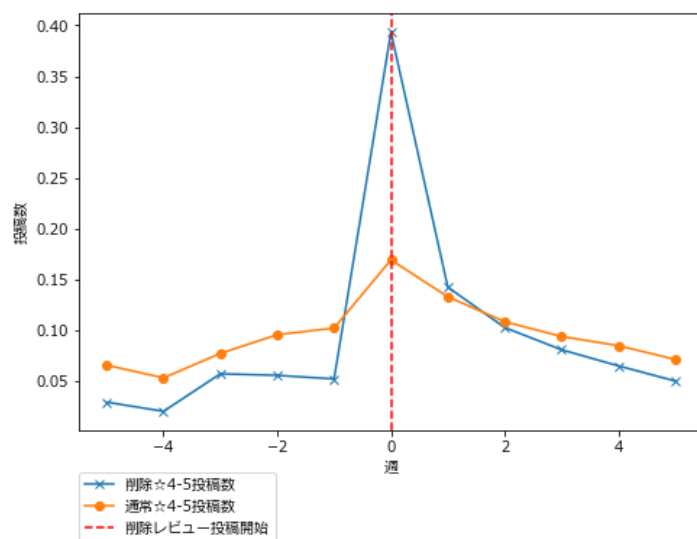
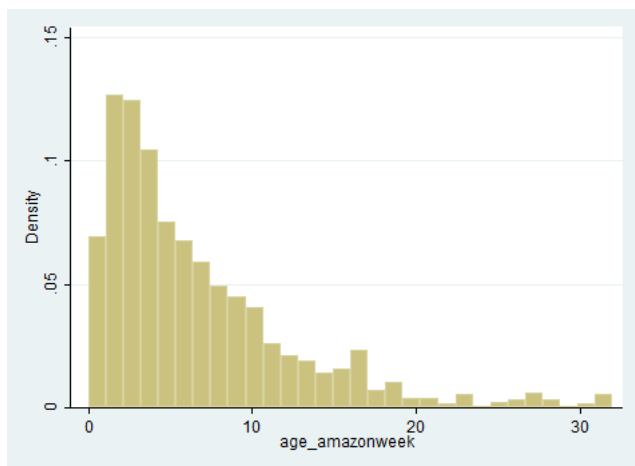
	(1)	(2)	(3)	(4)
	* 4-5(削除)	* 4-5(非削除)	* 4-5(削除)	4-5(非削除)
* 4-5(非削除・1週ラグ)	-0.26***	0.28***	-0.29***	0.30***
	(-4.09)	(4.66)	(-4.31)	(4.99)
* 1-3(非削除・1週ラグ)	0.00	-0.16***	0.03	-0.18***
	(0.08)	(-4.73)	(0.79)	(-4.25)
対数レビュー本数(非削除・ラグ)	0.05	-0.16	0.04	-0.16
	(0.43)	(-1.50)	(0.46)	(-1.70)
レビューベース商品年齢	-0.01	0.01***		
	(-1.95)	(3.83)		
取り扱い開始日ベース商品年齢			-0.01**	0.01***
			(-2.71)	(5.13)
切片	-0.11	0.16	-0.14	0.17
	(-0.42)	(0.64)	(-0.45)	(0.57)
<i>N</i>	68298	68298	50098	50098

*t* statistics in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

## ➤ 怪しいレビューが一定時期に大量投稿される

- 削除された高評価の怪しいレビューの投稿数が一定の閾値を超えたタイミングを不正操作開始の代理変数とする
- 初期に集中。怪しいレビュー自体は開始タイミングに集中



## ➤ 後に低評価が増加

- 怪しいレビューの数が閾値をこえて2週以降

	(1) ★ 1-3(非削除)	(2) ★ 1-3(非削除)	(3) ★ 1-3(非削除)
フェイク投稿開始 2 週間以降 (閾値 1)	0.09** (3.00)		
フェイク投稿開始 2 週間以降 (閾値 2)		0.15*** (3.68)	
フェイク投稿開始 2 週間以降 (閾値 3)			0.36*** (5.98)
★ 4-5(非削除・1 週ラグ)	-0.00 (-0.42)	-0.00 (-0.42)	-0.00 (-0.46)
★ 1-3(非削除・1 週ラグ)	0.14*** (6.51)	0.14*** (6.49)	0.14*** (6.40)
対数レビュー本数 (非削除・ラグ)	0.06*** (4.17)	0.06*** (3.90)	0.05*** (3.48)
取り扱い開始日ベース商品年齢	-0.00*** (-5.48)	-0.00*** (-5.30)	-0.00*** (-5.27)
切片	-0.03 (-0.58)	-0.09 (-1.41)	-0.28*** (-3.39)
<i>N</i>	50098	50098	50098

*t* statistics in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

- どのような企業がレビューの不正操作を行っているかは、ある程度判明。
  - まだまだエビデンス自体は必要
- 実際の被害がどれほどのものかについては未だ不明瞭
  - 限定的な分析しか存在しない
- 消費者サイドについてエビデンスが全くない
  - 公平性の観点からも問題
  - 精度の高いデータが必要
  - ラボ実験を行う例も存在(Ananthakrishnan et al.,2020)
- テキスト分析の可能性
  - 豊富なテキストデータが活かされていない
  - レビューへの騙されやすさが分かれば間接的に被害が推定可能
  - 商品同士の競合関係についても調査することができるかもしれない