

経済実験を用いた評価システム別の他者の評価による自身の購買と評価への影響分析

高橋裕紀^{1,2} 西野成昭¹

1. 東京大学大学院工学系研究科技術経営戦略学専攻 2. 産業技術総合研究所

研究背景

評価プラットフォーム

基本的にサービスや小説、映画などのコンテンツは消費するまで価値がわからない。そこで多くの消費者は購買の意思決定をする前にそれらのサービス・製品についての情報を集める。特に近年、サービス・製品を経験した人が、その満足度、レビューを他者に共有する、評価プラットフォームが一般的に影響を持ってきている^[1]。

評価に関するバイアス

一方で、多くの研究で、
• 消費者のレビュー内容が直前の消費者のレビューに影響を受けている^{[2][3]}
• そもそも消費者がレビューを投稿するかどうか直前のレビューに影響する^{[4][5]}

など、製品の質自体とは関係ないバイアスを含んでいることが示されている。どのような評価提示の仕組みであれば、なるべくバイアスを少なく良い製品にたどり着けるのか分析することは重要である。

研究目的

「評価システム別の人の購買や評価に対する影響を実証検証すること」

手法

今回用いる手法

今回は、現実でのいろんな影響を排除して、人が他の人の評価を見ること、見られることによって、長期的に市場がどのような影響を受けるかに注目して検証したい。そこで購買活動における様々な要因を一定にして(統制して)、評価の見せ方だけを変えた影響検証することに適している経済実験を用いる。

経済実験について

経済実験では、実際の社会の経済活動を実験室内に仮想的な数字として再現し、人の振る舞いについて分析する方法である^{[6][7]}。経済実験では選好を統制することができる。つまり人々の好みや嗜好を現金報酬によって再現することができる。

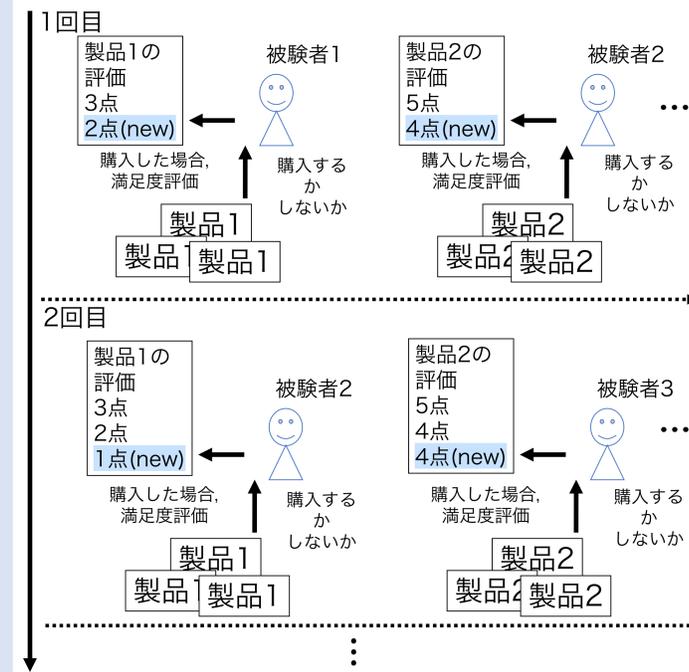
実験設定(1)

今回の実験について

被験者：東京大学所属の学生及び大学院生
実験日時：2020/10/12- 2020/10/14
実験回数：6回
被験者数：各回16人(合計96人)
方法：zoomと経済実験用のwebページで行った。

概要

実験は16人を1グループとして、同グループにおける他の被験者の製品レビュー(製品の評価スコア)の提示方法の違いから、4種類の設定で実験を行った。被験者は各設定において、レビュー情報を参考に、16回の仮想的な製品購入の意思決定を行なった。



全16回

図1. 16回の意思決定の流れ

各意思決定は、「(1) 仮想的な製品の購買意思決定」と「(2) 購入した製品の評価スコアの選択」の2つのステージから成る。(1)で購入した場合、その製品から得られる効用(嬉しさ)が判明し、それをもとに(2)で今回の製品の満足度を5段階で記入してもらった。ただし購入しなかった場合は評価をすることができない。所持金は毎回同じとした。

実験設定(2)

4つの設定

今回評価システムの設定によってどのように挙動が異なるかを検証するため、以下の四つの設定を試した。

- 設定A：レビュー情報なし
- 設定B：直前の評価のみ
- 設定C：全レビューの平均値を表示
- 設定D：全レビューの点数分布を表示



図2. 各実験の流れ

参加者の利得(報酬)

今回の利得設定はBesbes and Scarsini (2018)をもとに構成した^[8]。被験者の報酬は利得が大きくなるほど、大きくするように設定した。
(利得) =

$$(\text{所持金}) - (\text{製品価格}) + (\text{製品からの効用})$$

$$(\text{製品からの効用}) =$$

$$(\text{共通の基礎的な効用}) + (\text{個人の付加的な効用})$$

共通の基礎的な効用：誰にとっても同じ製品の好み。一様分布から製品ごとに決定。

個人の付加的な効用：人によって好みが変わるもの。一様分布から製品と個人の組み合わせごとに決定。

参加者は、評価を見つつ、なるべく共通の基礎的な効用が高い製品の購入を目指す。

実験結果

被験者は良い製品を購買できているか？

図3をみると、設定Cと設定Dでは、共通の基礎的な効用がよくなるほど、購買割合が比例して増えており、被験者は良い製品にたどり着ていることがわかる。一方で設定Bではあまり購買割合が増えていない。たまたま個人の付加的な効用分(好み)が悪かった時つけられた満足度を見て買い控えていると考えられる。

実験結果

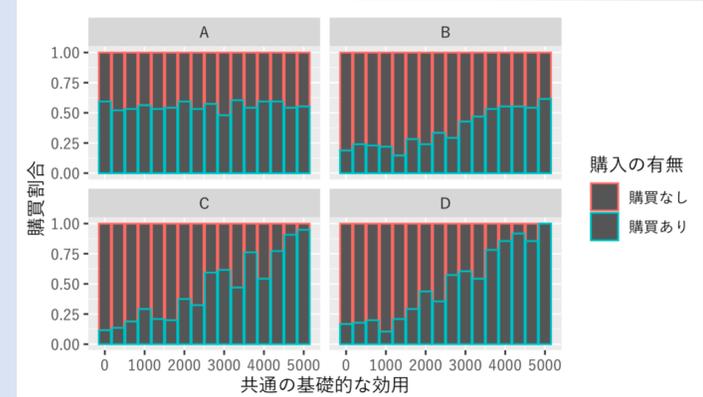


図3. 共通の基礎的な効用に対する購買意思決定の割合(設定別)

評価の付け方の違いは？

表1.をみると設定Bは高い効用になる程他の設定と比べて低くつけることが交互作用が有意な点からわかる。設定Bは自分の評価の影響が大きいので、自分がたまたま好み良かったのではないかと考え、他人に周知する時は少し控えめに評価するのではないと思われる。

表1. つけた満足度に対する順序ロジスティック回帰の結果

	満足度			
	(1)	(2)	(3)	(4)
効用	2.86*** (0.07)	3.00*** (0.11)	3.00*** (0.11)	3.03*** (0.11)
設定 B	0.11 (0.11)	0.08 (0.11)	0.09 (0.11)	0.09 (0.11)
設定 C	-0.04 (0.10)	-0.07 (0.10)	-0.06 (0.10)	-0.06 (0.10)
設定 D	-0.01 (0.10)	-0.04 (0.10)	-0.05 (0.10)	-0.05 (0.10)
効用 × 設定 B		-0.42*** (0.14)	-0.42*** (0.14)	-0.43*** (0.14)
効用 × 設定 C		-0.01 (0.14)	-0.01 (0.14)	-0.01 (0.14)
効用 × 設定 D		-0.20 (0.13)	-0.20 (0.13)	-0.20 (0.13)
コントロール変数	No	No	No	Yes
観測数	2,903	2,903	2,903	2,903
Log Likelihood	-3,152.73	-3,148.37	-3,147.34	-3,132.08

Note: コントロール変数には各実験のダミー変数を入れた
*p<0.1; **p<0.05; ***p<0.01

まとめ

- 評価の共有形態によって評価の仕方や学習が変わってくる。
- 確認できる評価が比較的少ないと、高評価が付きにくく、買い控えがおき、良いものにたどり着きにくい。

[1] Deloitte. (2007). Most Consumers Read and Rely on Online Reviews; Companies Must Adjust. Deloitte & Touche USA LLP. [2] Schlosser, A. E. (2005). Posting versus Lurking: Communicating in a Multiple Audience Context. Journal of Consumer Research, 32(2), 260-265. [3] Ho, Y. C. C., Wu, J., & Tan, Y. (2017). Disconfirmation effect on online rating behavior: A structural model. Information Systems Research, 28(3), 626-642. [4] Lafky, J. (2014). Why do people rate? Theory and evidence on online ratings. Games and Economic Behavior, 87(December 2013), 554-570. [5] Moe, W. W., & Schweidel, D. A. (2012). Online product opinions: Incidence, evaluation, and evolution. Marketing Science, 31(3), 372-386. [6] Friedman, D., & Sunder, S. (1994). Experimental Methods. Cambridge University Press. [7] 川越敏司. (2007). 実験経済学. 東京大学出版会. [8] Besbes, O., & Scarsini, M. (2018). On information distortions in online ratings. Operations Research, 66(3), 597-610.