Customer content: Role in Product Discovery and Selection

Judy Chevalier Yale and NBER

Definition

UGC : any form of digital content that is produced by (typically) volunteer end users of an online service or website and (typically) publicly available for consumption by other users of the service or website.

Examples: videos, blogs, discussion forum posts, photos, product reviews, seller reviews.

I am going to focus on IO and structured reputation systems, role of consumer feedback in algorithms, and product review systems, though other UGC and IO is worthy of study too (see Gans, Goldfarb, and Lederman (2019) on consumer tweets, for example).

Plan

- Tell you about the research, try to move from what HAS been done to what COULD be done.
- Like the bulk of the research, I am going to have too little to say about the role of Google and Facebook as purveyors of reviews.

Reputation systems

- Product\service reviews in those online marketplaces that plausibly couldn't exist without a reputation system.
 - Peer to peer markets: ebay, Airbnb, TaskRabbit, freelancer.com, upwork, etc.
 - Markets where reputation, branding, etc. weak.
 - Markets with both potential moral hazard and adverse selection problems.

Reputation systems literature

- Theoretical and empirical literature typically framed around vertical quality and opportunism.
- Inform future buyers about the outcomes of a seller's past behavior.
 - Sometimes muddiness about the seller's performance and the quality of the product.
- Tadelis (2016) focuses on the role of a reputation system in turning a (hopeless) anonymous one-shot game into a repeated game, thus expanding the available equilibria.

One-sided vs. Two-sided

- Covered well in Tadelis review (2016)
- Literature suggesting problematic retaliation/grade inflation issues in 2-sided systems. (Bolton et. al. 2013, Fradkin et. al. 2015, Horton and Golden 2015).
- Ebay eliminated the buyer side of the feedback system when Paypal became ubiquitous (and therefore buyer performance became trivial). Klein, Lambertz, Stahl (2016) document that the market becomes more transparent and seller behavior improves.
- If the buyer co-produces the output (Airbnb), buyer side non-trivial and two-sided rating inescapable. "Usual" solution is some kind of pseudo one-sided rating.

Research idea: impact/tradeoffs of platform design solutions in 2-sided systems.

Causal impact of reviews on sales

- Literature both in the context of the reputation systems created by online marketplaces and for online reviews more generally.
- Identification schemes:
 - cross-platform comparison (Chevalier and Mayzlin (2006)
 - Regression discontinuity((Luca (2011), Luca(2016)))
 - Field experiments (Resnick et. al. (2006), Godes and Mayzlin (2009))
 - Regional variation in the content of information sets (Helmers, Krishnan, and Patnam (2015)), Reimers and Waldfogel (2021))
 - Timing of product availability (Chintagunta et. al. (2010))

Chevalier and Mayzlin (2006)

- Our econometric analysis is designed to answer the following question: If a cranky consumer posts a negative review of a book on bn.com but not on Amazon.com, would the sales of that book at bn.com fall relative to the sales of that book at Amazon.com?
- We propose a "differences-in-differences" approach. For a sample of books, we measure reviews and a proxy for sales at Amazon.com and bn.com over three time points. We examine whether a change in the number and valence of reviews over time for a particular book at one site relative to the other site predicts a change in the *subsequent* sales of that book at one site relative to the other.
- By focusing on the differences between the relative sales of the book at the two sites, we are able to control for the possible effect of unobserved book characteristics on both reviews and sales. By focusing on the differences across sites over time, we control for the possibility that taste differences across the customer populations at the two sites differ in a way that affects both reviews and sales.

Cross-platform trick

- Chevalier and Mayzlin (JMR, 2006 Amazon vs. Barnes & Noble)
- Mayzlin, Dover, and Chevalier (AER,2014 Tripadvisor vs. Expedia)
- Chevalier, Dover, and Mayzlin (MS 2019, Tripadvisor/Expedia vs. Hotels.com, Priceline).
- Reimers and Waldfogel (2019, Amazon country sites)

Research idea: Use matched objects across platforms to isolate a platform design issue and/or causal effects.

Caveats

- Sloppiness in the use of sales ranks as a proxy for sales.
 - Leibowitz and Zenter (2021)
- New econometric findings make diff and diff strategies challenging.

– Econtwitter (2019, 2020,2022)

 Review impact largely a function of how reviews are displayed/incorporated into algorithms.

Reviews and Algorithms

Lee and Musolff (2021)

Paper investigates the Amazon Buybox in which one offer for a particular good is "recommended".

- Reverse-engineers characteristics that increase probability of recommendation
- Estimate a consumer choice model incorporating recommendation
- Consider impact of Buybox on new price competition, product entry and welfare.



(c) Shipping Time.

Figure 4: Non-Price Characteristics Give Recommendation (Dis-)Advantage.

Notes: These plots measure the extent to which a merchant could raise the price of her offer (relative to MSRP) if she were to move from the baseline value of a given feature to the value given on the x-axis. We see that both feedback count and the percentage of positive feedback matter slightly, but shipping time (measured here in hours) matters a lot more.

How important to welfare?

Reimers and Waldfogel (2021)

- Digitization has lowered the cost of production and distribution of new content.
- Explosion of new creative content –music, books.
- Lots of new content means discovery is a challenge
- Can we identify the welfare effects of product reviews? Ultimately the welfare impacts will come from-
 - consumers buying more overall when quality is known
 - AND redistribution of sales from lower-quality to higher quality products
 - (and better MATCH quality, but this is not in the paper).



Waldfogel JEP 2017

Strategy

- Find the causal impact of reviews on sales. Use these to simulate welfare effects
 - Professional reviews (~NYT) discrete events- an event study methodology can determine effects.
 - Consumer reviews tricky
 – reviews and sales co-determined.
 - Use a variant of Chevalier and Mayzlin (2006) exploiting the fact that Amazon Canada, UK, US have different reviews.
- By measuring impact of consumer reviews on sales, can think about the counterfactual that the book had the quality revealed from the reviews but that the quality was not known as precisely absent reviews.
 - Requires an estimate of the "expected" quality of the bookhere based on genre, publisher, author prior track record.

Welfare effect of reviews when reviews reveal true quality.

- Case with expected quality less than true quality
 - CS with ratings is A + B + C 1950
 - CS without ratings is A + B
 - Change CS is C
 - For overconsumption case, D



FIGURE 1. ILLUSTRATION: WELFARE ANALYSIS OF PRE-PURCHASE INFORMATION

Notes: This figure illustrates demand curves under full information about a product's quality (solid line) and with limited ex ante information about the quality when the expected quality is less than the true quality (dashed line). The corresponding consumer surplus under full information is areas A + B + C; under limited ex ante information, it is A + B.

Strategy: calibrate to a logit

- Using estimated causal impact of reviews on sales, estimate the elasticity of sales with respect to quality.
- Use estimated price elasticity to get the shape of the demand curve.
- Use an estimate of what quality would be expected absent reviews for each book (based on genre, author history, etc).
- Calibrate welfare impacts.

	Stars	Reviews	Ratio
ΔRevenue (net)	27.51	19.98	
	(11.79)	(2.10)	
$R > \hat{R}$	92.56		
	(17.00)		
$R < \hat{R}$	-65.05		
	(6.29)		
ΔCS (baseline)	35.83	3.18	11.27
	(6.98)	(0.41)	(2.28)
ΔCS (50 categories)	35.83	3.22	11.13
	(7.13)	(0.42)	(2.23)
ΔCS (reviewed books)	1.68	3.18	0.53
	(0.30)	(0.41)	(0.13)
$\Delta CS (\sigma = 0)$	36.58	3.27	11.18
	(7.43)	(0.42)	(2.34)
$\Delta CS (\sigma = 0.95)$	34.67	3.03	11.42
	(6.29)	(0.39)	(2.20)
ΔCS (Marshallian: unconstrained)	37.66	4.76	7.92
	(8.01)	(0.60)	(1.80)
ΔCS (Marshallian: $\Delta q = 0$)	40.52	3.89	10.41
	(7.83)	(0.48)	(2.24)

TABLE 4—WELFARE IMPACTS OF PROFESSIONAL REVIEWS AND AMAZON STAR RATINGS

Results

• Effects on revenue

- Reviews: +\$20m
- Stars: +28m, net

- Effects on CS
 - Stars: +\$35.8m
 - Reviews: +\$3.2m
 - Effect of stars is 11 times larger
 - Effects of stars for books that have professional reviews is only \$1.7m

Notes: All figures are in millions of dollars. Figures for stars are based on scaling sample results up to Amazon's share of 2018 book sales of 695 million units. Because we include all of the books reviewed at the *New York Times* and the other major papers in the sample, the model's direct measure of ΔCS from these reviews requires no scaling. Figures are based on estimates in column 5 of Table 2. Standard errors are based on 500 parametric bootstrap draws.

Research idea: Welfare effects of reviews, particularly exploring the MATCH QUALITY dimension.

Who posts reviews and why?

• See survey Berger (2015)

- impression management
- emotion regulation
- information acquisition
- social bonding
- persuasion

• Impact and ability to influence the audience.

- Chakraborty, Deb, Oery (2020)
- Wu and Huberman (2008)
- Moe and Trusov (2011)
- Zhang and Zhu (2011)
- Proserpio and Zervas (2017); Chevalier, Dover, and Mayzlin (2019)
- Negative feedback particularly difficult to elicit
 - Nosko and Tadelis (2014)

Research ideas: Dynamic quality vs. static quality environments. Experiments on making reviewing more/less satisfying.

Are reviews biased?

- Reciprocity bias: Cabral and Hortacsu (2010), Fradkin et. al. (2015)
- Promotional behavior bias: Mayzlin (2006), Mayzlin, Dover, and Chevalier (2014), Luca and Zervas(2015)
- Positive selection: Hitt and Li (2008), Horton (2014), Nosko and Tadelis (2015), Berger and Milkman (2012)
- Systematic review dynamics over time: Li and Hitt (2008), Godes and Silva (2011), Moe and Trusov(2011)

Promotional/Fake Reviews

Mayzlin, Dover, Chevalier (2014)

Exploit difference in posting rules across sites

TripAdvisor (anyone can post), Expedia (verified to have purchased).

Exploit differences across organizations in incentives to post fake reviews.

- Benefit of faking accrues to one hotel, cost of sanctions accrues to all in chain.
- Ex: Archon Hospitality, 41 hotels, different nameplates (Hampton Inn, Fairfield Inn, etc.)

Claim 1: A firm that is located close to a competitor will have more fake negative reviews than a firm with no close neighbors.

Claim 2: A firm that is part of a smaller entity will have more positive fake reviews.

Claim 3: A firm that is located close to a smaller entity competitor will have more fake negative reviews.

Idea: Examine the role of platform design tradeoffs in controlling fake reviews. How do fake reviews impact consumer selection?

Limitations

- Large number of studies of whether reviews impact sales
 - Little on the relationship between reviews/other
 UGC and more traditional IO/Policy concerns
 - In Econ & Marketing (but not CS): Substantial focus on quantitative review summaries (stars)
 - Challenge: bridge unstructured textual data created by UGC and quantitative inputs to IO/policy analysis.

Reviews and Regulation

Farronato and Zervas (2021)

- Uses the TEXT of the reviews to associate Yelp review words with regulator's hygiene violations.
- Shows some hygiene violations are predicted by review text and some are not.
- Shows that "Yelp visible" restaurants are less likely to have the "review visible" hygiene violations.
- Shows that these hygiene words are worse for probability of a restaurant selling out on OpenTable than other reviews of the same star values.

Possible research avenues

- How does existence/proliferation/nature of UGC *drive* variables of interest in IO/policy?
- How can UGC be used to *measure* variables of interest in IO/policy?

Possible research avenues

- How does existence/proliferation/nature of product reviews/UGC *drive* variables of interest in IO/policy?
- How can UGC be used to *measure* variables of interest in IO/policy?

Market Structure

- Intuition from Bar-Isaac, Caruano, Cunat (2012):
 - More product information (lower search costs)
 - Superstar effect
 - Consumers more likely to find the "best product"
 - Encourages more-advantaged firms to choose broad designs
 - Niche effect
 - Consumers are more likely to buy better-suited products
 - Induce more firms to choose more niche designs
 - Effects of more information/reduced search costs on competition and markups ambiguous.

Vellodi (2020)

- Product reviews and barriers to entry.
 - This paper examines the endogeneous entry and exit decisions in the presence of product reviews.
 - Intuitively, consumers will choose a highly-rated incumbent over an unrated entrant even if the entrant is higher quality.
 - This creates a barrier to entry. (This is still true with endogeneous pricing because the entrant has to price low to attract customers, still making entry unattractive.)

Role of UGC in entry/variety....

- We just discussed markets where we suspect they couldn't exist without review mechanisms (requiring trust of a random stranger).
- We discussed Reimers and Waldfogel (2020) which implicitly is suggests that reviews function to enable more varieties in books.
 - Luca (2011) another paper w this conclusion
 - Bar Isaac, Caruano, Cunat (2012) related
- Vellodi (2020) is identifying the opposite effect.

Research Idea: Explore this tension more theoretically/empirically

Impact of UGC on quality provision

- Quality improvements
 - Klein, Lambertz, Stahl (2016)
 - Ananthakrishnan, Proserpio, Sharma (2019)
 - Closely related: Gans, Goldfarb, Lederman (2017)

- Farronato and Zervas (2021)

- Reviews and market outcomes in a multitasking framework.
 - Example: patient reviews of doctors
 - Farronato and Zervas (2021)- hygiene

Astonishingly little research

- Reviews/UGC and the provision of *horizontal* quality information.
 - Render the undifferentiated differentiated.
 - Allow for creation/discovery of niche products.
 - To what extent do product reviews on sites like Tripadvisor allow the creation of niche products or to what extent do they just disintermediate an older information provision sector (travel agents, Zagat, Fodors, etc.)?
 - Role of UGC in improving match quality
 - See Donnelly, Kanodia, and Morozov (2022)

Research idea: This.

Competition among UGC platforms and competition between UGC platforms and other firms

- Does the existence of a UGC asset create competitive advantages in a transaction market?
 - The content is usually non-rival and non-excludable.
 - Freestanding UGC content providers entering the transaction role.
 - Yelp reservation system
 - TripAdvisor direct booking
 - The role of Google and Facebook in product review provision.

Possible research avenues

- How does existence/proliferation/nature of UGC *drive* variables of interest in IO/policy?
- How can reviews/UGC be used to measure variables of interest in IO/policy?

Measurement of characteristics

- Inputs to demand analysis
 - We typically use introspection/data availability to choose characteristics to include in models.
 - UGC contains rich information about characteristics that consumers evaluate.
 - Or at least a selected set of consumers.
 - Challenge: Bridge from unstructured linguistic
 UGC to characteristics, etc. that can be used in, for example, demand estimation.
 - Closest paper: Archak, Ghose, Ipeirotis (2012)
 - See also Timoshenko and Hauser (2019)

Second choice/Close competitor modelling

- Anything like second choice data helpful in demand system modeling (see BLP 1998)
- Textual analysis of UGC can be used for "brand associative network" (see Netzer et al 2012).
 - In reviews/blog posts about product x, what other products are referred to and how many times?

Possible research avenues

- How does existence/proliferation/nature of UGC *drive* variables of interest in IO/policy?
- How can UGC be used to *measure* variables of interest in IO/policy?