Dynamics of Online Reviews and Consumer Learning

Tommaso Bondi
NYU, Stern School of Business

May 31, 2018

Recent years have witnessed an exponential growth of peer generated recommendation systems in a variety of domains. A growing body of research has shown that such recommendations causally impact consumers' purchases. Most of the literature has taken a static approach to the problem: at a given point in time, how does demand depend on the aggregate score displayed by the website?

Here I propose a complementary approach, one involving studying the dynamics of online reviews over products' lifecycles. In my view, this is of first order importance for at least two reasons. First, it brings this literature much closer to the current dynamic models of demand and competition, and it provides insights on the effect of ratings on entry and exit dynamics - e.g., understanding how new products establish their reputation, and whether a declining one can be a driver of exit. Second, studying reviews' dynamics informs their very value. In an influential study, Hitt and Li (2008) found that reviews for a variety of goods tend to decrease over time, and they point to consumer self-selection as a main driver: die hard fans buy first, and rate enthusiastically; others naively follow. If this is the prevalent case, then it is *a priori* unclear whether recommender systems have a positive impact on welfare: in principle, they might simply create inefficient herds reflecting which products have a more vocal, if small, fan base.

Preliminary data analysis I have conducted on data scraped with a coauthor, Ryan Stevens (also at NYU), for a different project, presents a significantly different story: for the products we examine, beer and movies, the time series of reviews is often *increasing*. These early findings motivated my interest in understanding how consumers' opinions covary with (and possibly influence) products' lifecycles. Moreover, I'm interested in understanding how ratings affect market structure. Under which conditions do they help creating a long tail? When, instead, do they reinforce the initial popularity of stars?

When thinking about these questions, a key force to take into account, in my view, is horizontal differentiation, and the sorting role of reviews. Most existing papers consider vertical markets where reviews indicate quality. But in reality, consumers might be much more focused on what is *good for them* rather than what is simply *good*. This is particularly true in a variety of horizontally differentiated markets, such as movies, books, or restaurants, where reviews are the most prevalent. Indeed, Amazon and Netflix's business models heavily on matching consumers and products, especially niche ones.

It is easy to see that once we take these considerations into account, an increasing time series of ratings makes intuitive sense: initially, products (especially new ones) and consumers are poorly (near randomly) matched. As opinions accumulate, consumers are better able to identify products that match their taste. This is particularly important for niche products, which are likely to attract a majority of early bad ratings due to poor matches.[1]

That is, the quality of matches between products and consumers is increasing over time. Clearly, this offers a much more positive view of the welfare value of reviews. Not just do they reward quality in an objective sense, solving the moral hazard problem faced by buyers; but they also ease search for consumers, creating better matches and potentially helping sellers with polarising products find their intended customers.

Methodology

Very briefly, the study would include three main components:

- A **theoretical model** trying to illuminate the impact of (different) recommendation systems on entry, exit, market structure (e.g., long tail vs superstars) and consumers' satisfaction over time.

- An **empirical analysis**, leveraging a large dataset of (textual and numerical) ratings over time, as well as reviewers' characteristics. What does the time series of ratings look like for a given product? Does that depend on products' characteristics, e.g. niche vs mainstream, new vs known? Do reviews constitute a barrier to entry? What happens to consumers' self-selection over time? Do they become better matched with the product?

- A **ratings design** part, where in light of the evidence on the sorting role of reviews, I propose different ways to aggregate information that might increase their value to

[1] Perfectly rational consumers could internalise this, of course; but a majority of studies on the topic show that consumers take reviews at face value, and are unlikely to correct for their many, predictable biases.

consumers. This is, in essence, a prediction problem: what will make the consumer happiest? Therefore, a machine learning approach seems a promising way to go: one can, for instance, feed the algorithm both the individual numerical scores and the text reviews, as well as reviewers' characteristics, and let it predict satisfaction for future (heterogeneous) consumers. Their own reviews would provide the necessary feedback. Would this substantially change (and hopefully improve over?) the weighted average scores employed by a majority of the most prominent websites?

Estimated Costs

This grant would be of tremendous help in three different areas:

- **Data Acquisition: $7.500.** A variety of websites, including some of the ideal ones for this study, are starting to protect themselves from scraping. For this reason, it would be ideal to purchase the required data.
- **Travel Costs: $1.000.** I plan to attend at least two related events next year: the NBER Economics of AI Conference in Toronto in October and the NBER Economics of Digitization Tutorial at Stanford in March.[2]
- **RA Expenses $4.000.** I am currently an Economics Intern at Amazon Web Services in Seattle, and will be starting my 5th year in September. This leaves around ten months before the Marketing job market - one I am very interested in - begins. Hence, it would be key to delegate some of the most mundane tasks, from scraping to data cleaning, and potentially helping with ML - to ensure timely completion.

[2]The travel funding provided by my school would not cover both.