Does the Distribution of Online Ratings affect Sales? Evidence from Amazon

Chinonso Etumnu[1]
Kenneth Foster
Nicole Olynk Widmar
Jayson Lusk
David Ortega

October 9, 2019

THIS IS A PRELIMINARY VERSION.

LINK TO THE MOST RECENT VERSION.

[1]Chinonso Etumnu is a Ph.D. Candidate and corresponding author (cetumnu@purdue.edu), Dr. Foster is a Professor and President's Fellow, Dr. Widmar is a Professor and Associate Head, Dr. Lusk is a Distinguished Professor and Head, all in the Department of Agricultural Economics at Purdue University. Dr. Ortega is an Associate Professor in the Department of Agricultural, Food and Resource Economics at Michigan State University. This manuscript greatly benefitted from comments from participants of a seminar held at Purdue University.

Abstract

This study examines how the distribution of online consumer ratings provides useful information about the quality of ground coffee products, which translates to sales. Using web scraped panel data from Amazon.com, the study quantifies the effect of not just the number of ratings, the average rating and the variance of rating, but also the skewness of the rating distribution. We find that the number of ratings and each of these moments of the distribution of online consumer ratings affect ground coffee sales. The size of the effect of the distribution of ratings was found to vary with respect to the sales level of the ground coffee products. Our results suggest that the distribution of online ratings plays an important informational role in e-commerce platforms. These results could aid decision making by producers of products sold online and online retailers/selling platforms both in providing consumers with important information as well as managing inventories and promotional activities.

Keywords: e-commerce; ground coffee; average rating; number of ratings; variance of ratings; skewness of ratings; distribution of ratings; online consumer ratings; web scraping.

1 Introduction

Traditionally, consumers rely on expert opinions and experience to resolve product quality uncertainty (Anderson and Magruder, 2012). However, with the growth in e-commerce and social media, online consumer ratings provide an alternative or supplementary source of product information as well as consumer shopping experiences. Online retailers and platforms typically allow consumers to provide a rating for a purchased product, with a range of 1 to 5 stars, and to post textual and even photographic evaluations. According to a recent survey, 92 percent of consumers who shop online read consumer ratings, and 40 percent of them form an opinion by reading about one to three ratings (Shrestha, 2017).

Many retailers and platforms with an online presence have recognized the role of online ratings and have responded by providing consumers with not just the summary of the number of ratings and the average rating, but also the percentage distribution of the ratings. For example, on Amazon.com, the beverage Folgers Classic Roast Ground Coffee, Medium Roast, 30.5 ounces, as of September 5, 2019, had received 584 consumer ratings, with an average rating of 4.5 stars, and the following distribution of ratings: 77% were 5 stars, 11% 4 stars, 6% 3 stars, 2%–2 stars and 4%–1 star.[2] But, does the distribution of online consumer ratings affect the sales of this ground coffee product? That is, do the moments of the distribution of online consumer ratings, namely average rating, variance of ratings and skewness of ratings, affect the sales of the ground coffee product? If yes, does the effect of the distribution of online consumer ratings vary in magnitude with the sales level of the product? Providing answers to these questions deepens our understanding of how consumers use and infer information from online consumer ratings as well as its relevance in sales predictions.

[2] https://www.amazon.com/Folgers-Classic-Ground-Coffee-Medium/dp/B010ULFOWC

The literature on the effect of the distribution of online ratings is sparse, especially in the context of online grocery markets. Hu et al. (2009) examined all products on amazon.com and found that online consumer ratings are overwhelmingly positive. They provide two explanations for this negatively skewed distribution: purchasing bias and under-reporting bias. On one hand, purchasing bias is because those who purchase a product are more likely to have a positive valuation of that product, hence the negative skewness of the distribution of the ratings. On the other hand, consumers with extreme experiences are more likely to provide a rating to "brag or moan" than consumers with moderate experiences, resulting in under-reporting bias. Schoenmueller et al. (2018) corroborate these findings. Schoenmueller et al. (2018) analyzed over 250 million reviews from 25 different platforms and found that the main driver of the extreme positive distribution of product reviews is the prevalence of reviews from consumers with extreme experiences. They also found that this polarity in self-selection is enhanced by cognitive dissonance to increase the negative skewness of product review distribution. To mitigate biases leading to the extreme distribution of product reviews, Marinescu et al. (2018) suggested the use of financial incentives and/or pro-social cues to encourage consumers with moderate experiences to provide product reviews. However, financial incentives and pro-social cues raise questions about the authenticity of the ratings.

While previous studies focus on the prevalence, drivers and mitigation of the extreme distribution of online ratings, this study focuses on quantifying the effect of the distribution of online consumer ratings. Since providing online product ratings is generally voluntary, this study is more concerned about quantifying the effect of the distribution of online consumer ratings than on ways to solicit ratings from consumers with moderate experiences. Furthermore, although the extant literature focuses on a wide range of review platforms and products, little attention has been given to online groceries, which deserves more attention given the projected growth in the sector.[3] In addition, given the breadth of the grocery category, it is expected that the effect of the distribution of ratings might vary in magnitude depending on the specific food category. For example, ground coffee may see different effects versus wine or chocolate or baby formula because of differences in consumer demand for each of these products.

Using web scraped data from Amazon.com, this study estimates the effect of the distribution of online consumer ratings on ground coffee sales. The focus on coffee is driven primarily by two reasons. First, according to the National Coffee Association of the United States, the coffee industry contributed $225.2 billion to the economy of the United States in 2015. Thus, the coffee industry accounted for approximately 1.6% of the total U.S. gross domestic product and is also responsible for about 1.7 million jobs and $28 billion in taxes. Consumers spent $74.2 billion on coffee in 2015 and a part of this consumer expense went to online purchases. Second, coffee is one of the highest-reviewed food products on Amazon, the largest online retailer in the United States (Heng et al., 2018). In addition, coffee is also a highly differentiated food product that serves as an experience good, which makes ratings even more valuable to consumers (Heng et al., 2018).

This paper contributes to the literature by quantifying the effect of the number of ratings, the average rating, the variance and skewness of the rating distribution. It isolates the effect of each of these moments of the distribution of online ratings and provides insightful results that could influence demand predictions in online grocery markets. Findings from this study contribute to the broader literature on the informational role of online consumer ratings. Many studies have investigated the effects of different kinds of information on consumer demand (Unnevehr et al.,

[3] The Food marketing institute and Nielsen project that online grocery markets in the United States will be worth $100 billion in 2022.

2010). For example, food labeling, advertisement and certification have all been found to influence the food choices of consumers. However, this paper emphasizes the effect of another kind of information that influences food choice in the digital age—online consumer ratings, with a particular focus on the entirety of its distribution.

2 Related Literature

This review of literature related to online consumer ratings of products is segmented in line with four attributes of the distribution of online consumer ratings: number of ratings, average rating, variance of ratings and skewness of ratings, and is summarized in Table 1. Chevalier and Mayzlin (2006) found that an increase in the difference in the number of book reviews on Amazon.com is associated with an increase in book sales on the website. Other studies such as Cui et al. (2012) and Dellarocas et al. (2007) corroborate this finding. Using movie ratings, Dellarocas et al. (2007) showed that the volume of online consumer ratings has a positive and statistically significant association with future box office sales. Cui et al. (2012) found a strong effect of the number of reviews on the sale of new experience and search products, although the effect is stronger for experience products. They also found that the effect of the number of reviews is stronger during the early life of a product and the effect lessens over time.

There have been studies that find insignificant and mixed effects of the number of reviews on sales, including Clemons et al. (2006) and Khare et al. (2011). Clemons et al. (2006) found that the association between the sales growth of beer and the number of beer ratings is statistically insignificant. Using experiments, Khare et al. (2011) found that the influence of the number of online consumer ratings on consumer behavior depends on the pre-commitment of the consumers. They also found that different types of consumers interpret the high number of ratings differently depending on their own characteristics. Thus, Khare et al. (2011) suggested that the effect of the number of reviews on sales may vary.

Many studies that investigate the effects of online consumer ratings use the average rating as an indicator of quality and try to tease out its effect on sales. For example, Anderson and Magruder (2012) employed a quasi-experimental design (regression discontinuity) to estimate the effect of average Yelp.com ratings on restaurant reservations in San Francisco. They found that a half-star increase in rating results in a 19 percentage points increase in the frequency of restaurant sales, and this effect is even larger for restaurants that have external accreditation. Their results provide evidence on the richness of aggregate consumer reviews in informing restaurant quality. In line with Anderson and Magruder (2012), Luca (2016) used Yelp.com ratings and data from the Washington State of Revenue to associate average ratings with restaurant revenue. Luca (2016) found that an increase of one-star in the average rating in Yelp leads to an increase in revenue of about 5-9 percent for restaurants.

Other studies corroborate these findings for many other products. For example, Chevalier and Mayzlin (2006) found a similar result for book sales on Amazon and Barnes and Noble, and Zhu and Zhang (2006) found that a 4 percent increase in sales results from a one-point increase in the average ratings of video games. Using data from food choice experiments on pork in China, Lin et al. (2018) estimated the impacts of different ex-ante methods of mitigating hypothetical bias on consumer willingness-to-pay (WTP) for pork attributes, including average rating and number of ratings. They found that Chinese respondents have a WTP of RMB 12.27 for a marginal increase in the average rating. They also find that their WTP for a marginal increase in the number of rating were RMB 0.12 and RMB 0.01 for 47 and 502 ratings respectively. However, in contrast, Duan et al. (2008) did not find a significant influence of the average rating on movies' box office revenues.

Few studies investigate the effect of the variance of ratings on product sales. For example, Sun (2012) investigated the effect of the variance of online consumer ratings on relative book sales on Amazon and Barnes and Noble. She found that the variance of online consumer ratings interacts with the average rating to influence product sales. Zimmermann et al. (2018) build on the work of Sun (2012) to show that, due to taste differences, the optimal price of a product increases in the variance of ratings while the optimal demand for a product decreases with the variance of ratings. They also found that, due to quality differences, optimal price and demand decrease in the variance of ratings. Finally, they showed that risk-averse consumers may prefer a high variance product if the average rating is held constant. Other researchers that investigate the effect of the variance of ratings include Park and Park (2013), Langan et al. (2017) and Wang et al. (2015). Langan et al. (2017) showed that considering the variance of ratings independent of the average rating may lead to erroneous decisions by consumers. They also showed that high variance of ratings interacts with average rating to affect purchase intentions, although these effects are moderated by intrinsic (product attributes) and extrinsic (brand and reviewer credibility) factors. Park and Park (2013) examined low and high variance of rating factor into consumer product evaluations. They found that when consumers have a positive prior expectation of a product, a high variance of ratings may improve or weaken how consumers evaluate a product based on the product type (experience versus search), the quality of review text and the number of ratings. Using both secondary data and experiments, Wang et al. (2015) show that the effect of the variance of ratings can be positive, negative or insignificant depending on the consumer.

With regard to the skewness of online consumer ratings, Hu et al. (2009) observed that online consumer ratings are overwhelmingly positive, which leads to a negatively skewed distribution of ratings. Other authors made similar observations. Schoenmueller et al. (2018) used

280 million ratings from 25 rating platforms to show that online consumer ratings are negatively skewed. Dalvi et al. (2013) found that the distribution of online consumer ratings for restaurants, movies, and other products resembles the log-normal distribution, which differs from the normal distribution that is expected and observed for critic/expert raters. In addition, Zervas et al. (2015) found that the ratings in Airbnb are typically above 4.5 stars, which leads to negative skewness of the rating distribution. However, they found that the average rating in Tripadvisor is much lower at 3.8 stars. All of these studies suggest that as the number of online consumer rating increase, the distribution of the ratings is negatively skewed. Hu et al. (2009), Dalvi et al. (2013) and Schoenmueller et al. (2018) suggested that self-selection, under-reporting, and purchasing biases are the main causes of this extreme distribution of ratings.

There are at least three reasons why this study differs from the existing literature. First, although the effects of the number, average, variance, and skewness of ratings on sales have been previously examined, studies that have considered their effects simultaneously are limited. Because it has been hypothesized that all these attributes of ratings could affect product sales, having all of them in the same regression model as explanatory variables control for potential bias due to omitted variables. There is also limited empirical evidence that associates the skewness of rating distribution with product sales, which this study addresses. Second, this study also quantifies the effect of the distribution of ratings at different sales levels. It provides evidence that the effect of the distribution of online consumer ratings varies by the sales level of the product, which has not been well-established in the literature. Third, the product category of interest differs from those of existing studies. Most of the literature on the effects of online consumer ratings focus on non-grocery items, but this study focuses on coffee. This study considers coffee as a gateway product to investigate the effect of online consumer ratings on grocery sales. As online grocery markets expand around the world, so will the use of online consumer ratings to evaluate the quality of grocery items.

3 Conceptual Framework

Consumers who want to purchase a product online may have some implicit distribution of prior belief about product quality and product quality affects the utility derived from a product. Because the purchase of products online is irreversible for the most part for food items, a consumer may delay purchase until they can resolve some of the uncertainty (update their prior). The distribution of online ratings provides a mechanism for partially resolving the uncertainty and a likelihood of creating a posterior distribution in the Bayesian context. The distribution of ratings provides consumer with information that would allow them to assess the true quality of a product. It is thus assumed to approximate the true quality of a product.

Let g represent a ground coffee product and Q_g be the quantity demanded of the ground coffee. The attributes of ground coffee g include price (P_g), true quality ($\tilde{X}_g$), vector of observed time-invariant characteristics such as country of origin, brand, roast-type, bean-type, etc. (Z_g), and other characteristics that are known to the consumer and not the researcher (ϵ_g). Consumers are assumed to be perfectly informed about P_g, Z_g and ϵ_g but they are imperfectly informed about $\tilde{X}_g$. This uncertainty about $\tilde{X}_g$ can be partially resolved by learning from the distribution of online consumer ratings. The demand function for ground coffee can be obtained by maximizing the utility from it subject to the consumer's budget for ground coffee.

$$\max_{Q_g} U(Q_g; \tilde{X}_g, Z_g) \tag{1}$$

subject to

$$P_g Q_g = I_g \tag{2}$$

where I_g is the part of the consumer's income that is spent on ground coffee. The Lagrangian function for the problem is given by

$$\mathcal{L} = U\left(Q_g; \tilde{X}_g, Z_g\right) + \lambda\left(I_g - P_g Q_g\right). \tag{3}$$

The first-order conditions of the problem are as follows:

$$\frac{\partial \mathcal{L}}{\partial Q_g} = \frac{U\left(Q_g; \tilde{X}_g, Z_g\right)}{\partial Q_g} - \lambda P_g = 0, \tag{4}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = I_g - P_g Q_g = 0. \tag{5}$$

Solving 4 and inputting it into 5 results in the following demand function:

$$Q_g^* = f\left(P_g, I_g, \tilde{X}_g, Z_g\right). \tag{6}$$

It is assumed that the quantity demanded is a function of the moments of the distribution of the true quality of the ground coffee product. By incorporating the moments of the true quality of the ground coffee into a demand function, this study builds directly on the works of Golec and Tamarkin (1998) and Garrett and Russell (1999), who rely on the third-order approximation of the Taylor series of the expected utility function to incorporate the first, second, and third moments of risky outcomes. According to Golec and Tamarkin (1998), an advantage of this approximation is that it allows for the direct test of an individual's preference for skewness which other risk models ignore. However, the Taylor series approximations are simply assuming that the individual's utility function is cubic, which may be sensible for small risks as suggested by Chiu (2010). Furthermore, Chiu (2010) has theoretically established that preferences of decision-makers could depend on the mean, variance and skewness of the returns only and tradeoffs are made among these three attributes. Following this precedent, this study captures the effect of the first, second and third moments of the true quality of the ground coffee on the quantity demanded. Because the true quality of ground coffee (X_g) is a random variable, let the quantity demanded of the ground coffee be

$$Q_g = \beta P_g + \alpha I_g + \gamma_1 \overline{X_g} + \gamma_2 \sigma^2_{X_g} + \gamma_3 s^3_{X_g} + \delta Z_g + \epsilon_g. \tag{7}$$

Equation 7 suggests that the quantity demanded of the ground coffee is a function of the mean ($\overline{X_g}$), the variance ($\sigma^2_{X_g}$) and the skewness ($s^3_{X_g}$) of the quality of the ground coffee, as well as the price (P_g), other characteristics of the ground coffee (Z_g), and the income consumers spent on ground coffee (I_g).

4 Compiling Datasets Via Web Scraping

Amazon Inc. is one of the largest e-commerce marketplaces and online retailers in the world in terms of revenue and market capitalization. Jeffrey Bezos founded Amazon.com in July 1995, and it subsequently expanded to other countries. For example, Amazon.co.uk was founded in October 1998, three years after the parent company started. Amazon sells millions of products to hundreds of millions of its customers every year. Amazon.com has over 40 product categories, including grocery and gourmet food, which has about 20 subcategories. Under each of these subcategories are other categories of grocery items. For example, the category chain to reach the ground coffee products on Amazon.com is as follows: Grocery & Gourmet Food—Beverages—Coffee, Tea & Cocoa—Coffee—Ground Coffee. However, consumers often use the search engine and other features of the platform, such as average rating, prices, sellers and brand, to narrow down their search results (Beck, 2017).

With written permission from Amazon Inc. (obtained via e-mail on November 21, 2018), data on ground coffee on Amazon.com and Amazon.co.uk were scraped using a web scraping software called Octoparse.[4] Using Octoparse, web data scraping was carried out multiple times on Amazon.com and Amazon.co.uk between December 2018 and July 2019. The first web scraping was conducted on December 16, 2018, on Amazon.co.uk rather than Amazon.com because Amazon.co.uk is more structured than Amazon.com. That is, the attributes of a product on Amzon.co.uk appear on the same part of a webpage for most of the products on the platform. To collect data from a specific product category, particularly ground coffee, the following procedures were followed on Amazon.co.uk. First, at the left corner of the main search engine of Amazon.co.uk is the "All" option. Clicking on the "All" option opens up the categories of products sold on the website. From the category list, "Grocery" was selected. Then within the "Grocery" category, the product "Ground Coffee" was searched. This produces search results with a statement similar to "1-24 of over 2,000 results for Grocery Store: "Ground Coffee"". To narrow the search result to only "Ground Coffee", the subcategory filter on the left-hand panel of the website is used. Clicking on the "Ground Coffee" subcategory narrows the search to only ground coffee products on the website. The search results now produce an outcome such as "1-24 of over 1,000 results for Grocery Store: Drinks: Coffee: Ground Coffee: "Ground Coffee"". It is this weblink that is inputted into Octoparse to scrape the web data. This process was repeated on June 17, 2019, to collect the second round of data from Amazon.co.uk. The data comprised of all ground coffee on Amazon.co.uk with available data.

For each ground coffee product on Amazon.co.uk, the price, title, the Best Sellers Rank as well as the online consumer rating information—number of ratings and the percentage distribution of the 1-star, 2-star, 3-star, 4-star and 5-star—were collected using Octoparse. In an ideal scenario, sales data from the ground coffee products would be available along with the rating data. In this

[4] https://www.octoparse.com/

study, a proxy for the sales data used for each ground coffee product is the Best Sellers Rank. The Best Sellers Rank of a given ground coffee product is a snapshot of its sales at a given time relative to other products in the same category.[5] The Best Sellers Rank is calculated based on Amazon's sales transactions of a product relative to other products in the same category.[6] At Amazon.co.uk or any other e-commerce platform owned by Amazon Inc., the highest-selling ground coffee product has a Best Sellers Rank of one, and the ground coffee products that have lower sales are given higher sequential ranks.[7] However, the Best Sellers Rank does not provide any information about the sales gap between two products.

Amazon Inc. lets users post reviews for individual products. Reviewers must provide a rating from a scale of 1 star to 5 stars for the product, and then choose whether to provide review text, photographs or even videos. Amazon Inc. also provides information about the date when the review was posted, review helpfulness vote, top reviewers, name of the reviewer and whether the purchase was verified. Amazon then aggregates consumer ratings to produce summary statistics such as number of ratings, average rating, and the percentage distribution of the rating from 5 stars to 1 star, which is presented in a bar chart. It is from this distribution of online consumer ratings that the average rating, the variance and the skewness of the rating distribution used in this study were calculated.

A similar process was followed to scrape data from Amazon.com. Once the XPath for the needed attributes of the product, such as price, is modified, then the steps followed in the Amazon.co.uk case were easily repeated. For example, searching for "Ground Coffee" within

---

[5] https://www.amazon.com/gp/help/customer/display.html?nodeId=525376

[6] https://www.amazon.com/gp/help/customer/display.html?nodeId=525376

[7] This unorthodox ranking is important to bear in mind when interpreting the coefficients of the regression models.

Amazon.com's "Grocery & Gourmet Food" category results in the following: "1-24 of over 10,000 results for Grocery & Gourmet Food: "Ground Coffee"". Then, using the "Ground Coffee" subcategory filter narrows the search result to something like: "1-24 of over 4,000 results for Grocery & Gourmet Food: Beverages: Coffee, Tea & Cocoa: Coffee: Ground Coffee: "Ground Coffee"". Furthermore, the category was narrowed by sellers, by ticking Amazon.com. Note that Amazon Inc. is both an e-commerce platform that allows retailers to sell their products to consumers as well as an online retailer that sells its products directly to consumers. The first round of data from Amazon.com was scraped on March 29, 2019. The data included only ground coffee products on Amazon.com (platform) that were sold by Amazon.com (retailer). The second round of data was collected on June 28, 2019. These data included only ground coffee products that were sold by Amazon.com (retailer). The third round of data was collected on July 5, 2019, which includes data from all ground coffee products with available data on Amazon.com.

For the following analysis, the data collected from Amazon.com (e-commerce platform) that are sold by Amazon.com (online retailer) are used. After the main analysis, the robustness checks allow for all products sold by Amazon.com and other third-party sellers as well as products sold on Amazon.co.uk. The decision to restrict the analysis to products sold by Amazon.com is because of the incidence of fake reviews on the e-commerce platform. For example, in February 2004, it was documented that book authors and publishers had promoted their own books through positive reviews (Mayzlin et. al., 2014). However, because Amazon.com has a good reputation in terms of resisting fake reviews, narrowing the data for the main analysis of this study down to those sold by Amazon.com avoids this potential pitfall.

5 Empirical Estimation

The empirical estimation relies on three sets of approximations and assumptions due to data limitations: (1) Because Amazon does not release its quantity (sales) data, this study relies on the sales rank as an approximation of the sales of the ground coffee products. (2) Although income and tastes vary between consumers, Amazon does not release consumer information to researchers. Therefore, the homogeneity of consumers is assumed. (3) The distribution of ratings is assumed to be a proxy for the true mean quality of the ground coffee products. It is assumed that the mean quality is approximated by the average rating, the variance of the quality is approximated by the variance of ratings, and the skewness of the quality is approximated by the skewness of ratings.

Based on these approximations and assumptions, let

$$K - SR_g = \beta P_g + \delta Z_g + \eta N_{R_g} + \theta \bar{R}_g + \vartheta R_g^2 + \varphi R_g^3 + \mu_g \tag{8}$$

Where K is a constant that is greater than the maximum sales rank, SR_g is the sales rank of g, N_{R_g} is the number of ratings of g, $\bar{R}_g$ is the average rating of g, R_g^2 is the variance of ratings of g, R_g^3 is the skewness of ratings of g, and $\mu_g$is the error term. It is hypothesized that $\beta > 0$ because an increase in price is expected to reduce the demand for ground coffee and thus increase the sales rank. It is also hypothesized that $\varphi > 0$ because an increase in skewness (that is a tail of the distribution towards the 1-star rating) is expected to increase the sales rank and reduce the demand for ground coffee. Conversely, it is hypothesized that $\theta < 0$ and $\vartheta < 0$ because an increase in mean of the distribution of ratings indicates an increase in the quality of the product while an increase in the variance of the ratings indicates that the products meets the needs of various types of consumers. Both are expected to reduce the sales rank. N_g is hypothesized to have a positive effect on the demand for ground coffee and a negative effect on the sales rank.

With two periods of data as is in this study, equation 11 is more precisely written as

$$K - SR_{gt} = \beta P_{gt} + \delta Z_g + \eta N_{R_{gt}} + \theta \bar{R}_{gt} + \vartheta R_{gt}^2 + \varphi R_{gt}^3 + \mu_{gt} \quad (9)$$

where each of the variables and the error term now have a time period t, except the time-invariant characteristics of the ground coffee products.[8] With the time variation, a fixed effects model can be estimated, which controls for all the time-invariant characteristics of the ground coffee and Amazon. An advantage of this method is that it controls for any potential bias due to unaccounted time-invariant characteristics. But, it also eliminates other useful time-invariant characteristics of coffee contained in Z_g that affect the sales rank. Therefore, this study relies on the more flexible random-effects models.[9]

To empirically estimate the effect of the distribution of online consumer ratings on different levels of ground coffee sales, this study relies on the quantile regression model.[10] The quantile regression models the relationship between the independent variables and the conditional quantiles of the sales rank. The quantile regression models are given as follows:

$$K - SR_{gt} = \beta_q P_{gt} + \delta_q Z_g + \eta_q N_{R_{gt}} + \theta_q \bar{R}_{gt} + \vartheta_q R_{gt}^2 + \varphi_q R_{gt}^3 + \mu_{gt} \quad (10)$$

---

[8] An alternative way to incorporate the distribution of online consumer ratings is to include the fractions of the 5-star, 4-star … 1-star ratings directly in the demand function. This framework builds on the Maximin framework of Engelmann and Strobel (2004). This framework suggests that a consumer would want to maximize the odds of a favorable outcome and minimize the odds of an unfavorable outcome. In this case, this entails maximizing the effects of the 5-star, 4-star and 3-star ratings, and minimizing the effects of the 2-star and 1-star ratings. Let R_{g5}, R_{g4}, … R_{g1} represent 5-star, 4-star … 1-star ratings. Then the demand function of the ground coffee can be represented as

$$K - SR_{gt} = \beta P_{gt} + \delta Z_g + \eta N_{R_{gt}} + \alpha_2 R_{gt5} + \alpha_3 R_{gt4} + \alpha_4 R_{gt3} + \alpha_5 R_{gt2} + \alpha_6 R_{gt1} + \mu_{gt}$$

where α_2, α_3 and α_4 are hypothesized to be positive, and α_5 and α_6 are hypothesized to be negative.

[9] In choosing between the pooled OLS and the random effects model, the Breusch and Pagan (1980) Lagrange multiplier test were conducted and the results favored the random effects models. Because of these test results, only the random effects regressions are presented in the following discussions.

[10] The quantile regression in this paper did not consider the time dimension of the data because of the difficulty in employing panel data approaches in quantile regressions (Koenker, 2005).

where the unknown parameters are now associated with the q^{th} quantile.

Using the sales rank as the dependent variable in the regression models raises some estimation concerns, which are addressed below. First, the relationship between price and sales is potentially endogenous because the price can be a part of sales if sales are measured as revenue from the ground coffee products. However, in the sales rank patent filed by Amazon, the price of products is not associated with the sales rank (Hanks and Spils, 2006). This addresses potential endogeneity concerns between sales, the sales rank and ground coffee prices because the sales rank is strongly correlated with sales and not with prices. Second, the relationship between online consumer ratings and Best Sellers Rank could be endogenous if the latter is used in calculating the former. However, again from the patent filed by Amazon, the best sellers rank does not have anything to do with online consumer ratings (Hanks and Spils, 2006). In the description of the best sellers rank, Amazon states that: "The Amazon Best Sellers calculation is based on Amazon.com sales and is updated hourly to reflect recent and historical sales of every item sold on Amazon.com. … Amazon Best Sellers list is a good indicator of how well a product is selling overall…"[11] This increases the confidence that the relationship between the Best Seller's Rank and online consumer ratings is not endogenous because, from the above description, online consumer ratings are not taken into consideration in determining the sales rank.

6 Results and Discussion

Figure 1 shows scatter plots of the pooled observations by the sales rank, price, number of ratings, average rating, standard deviation of ratings and skewness of ratings. The plots are categorized by year—March 2019 (2) and June 2019 (1). Table 2 presents summary statistics for the two periods,

[11] https://www.amazon.com/gp/help/customer/display.html?nodeId=525376

in March 2019 and June 2019, for ground coffee products sold on Amazon.com (platform) by Amazon.com (online retailer). The final sample of ground coffee for March 2019 comprises of 810 products while the final sample for June 2019 comprises of 1221 products. The uneven sample is due to the focus on products that are sold by Amazon.com, alongside Amazon's policy of culling under-performing products and adding new ones (Chevalier and Mayzlin, 2006). When the products in the two periods are matched, a total of 438 products appear in the two periods. Over the period on average, sales rank, price, and skewness slightly decreased while number of ratings, average rating and standard deviation of rating slightly increased. However, Sun (2012) found that on average the number of ratings of books in their sample slightly decreased while the average rating slightly increased from January to May 2009. Sun (2012) suggests that these changes may be a result of pruning carried out by Amazon, which tends to remove products with low ratings. Another set of statistics in Table 2 is the mean of the fraction of ratings, from 1-star to 5-star. Figure 2 shows the mean of the distribution of ratings from the 5-star to the 1-star. This figure highlights a key attribute of the rating distribution—skewness. The figure indicates that most of the ground coffee products have received a rating of 5-star, followed although quite far apart by 4-star. In third position is the 1-star rating, which indicates the heterogeneity in product valuation that consumers have. But both distributions are negatively skewed, that is they are skewed towards the 5-star. Finally, Table 2 shows the fraction of ground coffee products in the sample with no ratings: 14 percent in June and 15 percent in March.

The first set of regression results are presented in Table 3. Column 1 of Table 3 reports the results of random effects regression between sales rank and price. Column 2 of Table 4 adds the number ratings, average rating, standard deviation of ratings and skewness of ratings. Column 3 of Table 3 adds an interaction variable between number of rating and average rating to the results presented in column 2. Coulmn 4 of Table 3 adds an interaction variable between average rating and standard deviation to the regression presented in column 2. Each of these columns provides information on both the coefficients and their corresponding elasticities of the attributes of the distribution of online consumer ratings are also provided. The results presented in column 2 and 3 of Table 3 also control for the effect of having "no ratings" on sales. In column 1 of Table 3, the coefficient of price is negative and statistically significant, suggesting that when the price of a ground coffee product increases, sales of that product falls. More specifically, a one percent change in price is expected to result in a 0.03 percentage change in sales. This inelastic response of sales to price is consistent with previous research that finds price inelastic demand for retail food products.

In column 2 of Table 3, one can see that the coefficients of number of ratings, average rating and standard deviation of ratings are posiitve while the coefficient of skewness of ratings is negative. This result indicates that a one-unit change in the number of ratings, average rating and standard deviation of ratings is expected to increase sales. The number of ratings may increase the visibility and credibility of the product and the seller, respectively, while the average rating helps people better assess the odds that the ground coffee is good quality or bad quality. Both attributes of online ratings may encourage the risk-averse consumer to try a coffee product. The coefficient of the skewness of ratings suggests that an increase in skewness will lead to a reduction in sales. Recall that positive skewness means that online ratings have their tail towards the 1-star rating while negative skewness implies that the tail of the distribution is towards the 5-star (see figure 2 for reference). The absolute elasticities of the number ratings, the average rating, the standard deviation and the skewness of ratings suggest that a percentage change in each of these attributes will lead to 0.01, 0.25, 0.03 and 0.05 percentage change in sales, respectively. These results suggest that the effect of these attributes is inelastic, but the effect of the average rating is most elastic than the other attributes of the distribution of ratings. The specification reported in Table 3 increased the overall R-squared of the random effects regression, from about 2 percent in column 1 to 38 percent in column 2.

In column 3 of Table 3, the sign of the interaction variable between the number of rating and the average rating is negative and statistically significant. This implies that as the number of ratings increases, average rating has a negative effect on sales, and vise versa. Adding the interaction variable increased the magnificute of the elasiticty of the number of rating and average rating from 0.01 to 0.23 and from 0.25 to 0.26 respectively. This suggests that considering the number of rating alongside the average ratings provides important information that translates to sales. In column 4 of Table 3, the interaction variable between average ratinga and standard deviation of rating has a positive effect on sales. This suggest that the effect of the average rating on sales is positive as the standard deviation of ratings increases. Adding the interaction variable slightly changed the elasticities.

Table 4 reports an alternative way of estimating the effect of the distribution of online ratings on sales. It directly estimates the effect of the fractions of the 5-star, 4-star … 1-star ratings relative to "no ratings". In column 1 of Table 4, the coefficient of 5-star, 4-star and 3-star ratings are positive and statistically significant. This result suggests that a one-unit increase in each of these attributes is associated with an an increase in sales. In terms of elasticities, the fraction of 5-star ratings is more elastic relative to 4-star and 3-star ratings. The coefficients of the 2-star and 1-star ratings are negative but not statistically significant. This finding suggests that a unit increase in the 1-star and 2-star ratings is associated with a reduction in sales. Column 2 of Table 4 includes interaction variables between the number of ratings and the fractions of ratings. The main find from this column is that the informational role of the number of ratings becomes larger in terms of elasticities when the interactions are included.[12]

One problem with the regression results presented in Table 3 and 4 is that time-invariant characteristics that are also important to coffee sales are not included in the models. However, previous studies have linked coffee prices and sales to characteristics such as roast, country of origin, bean, specialty labels, caffeine content, and brand. To account for these features of ground coffee, which may provide valuable information to ground coffee retailers, Table 5 presents results of random effects regressions that add these control variables. Even with these new control variables, the results are similar to previous estimations. Specifically, the coefficients of price, number of ratings, average rating, standard deviation of ratings and skewness of ratings are statistically significant and have the expected signs and comparable sizes.

A few of the time-invariant characteristics are statistically significant with sizable effects while others are not. For example, the coefficient of medium roast is positive and statistically significant, with absolute elasticity of 0.01. The coefficient of Arabica, decaffeinated and some of the coffee brands are negative and statistically significant. With respect to the distribution of online consumer ratings and the time-invariant characteristics of ground coffee, the key points from Table 5 are (1) despite the use of additional controls, the effects of price, number of ratings, average ratings, and the skewness of ratings are with expected signs, sizable and statistically significant as previously shown in Table 3. (2) Some of the time-invariant characteristics drive sales in online grocery markets. Thus, in addition to online consumer ratings, individual time-invariant product characteristics also affect product sales in online grocery markets.

[12] The models presented in Tables 3 and 4 show the effects of the distribution of online consumer ratings on ground coffee sales using different frameworks. Their residual plots looks largely similar but the R-squared of the mean-variance-skewness model is higher than that of the Minimax framework. Subsequent models thus focused on the mean-variance-skewness framework.

While the regression models presented in Tables 3, 4 and 4 address the question "are the distribution of online consumer ratings important for ground coffee sales?", they cannot address another important question: "Does the distribution of online consumer ratings affect high sales versus low sales ground coffee products differently?" That is, does a ground coffee product with a low/high sales rank see a different effect from a ground coffee product with a median sales rank? To address this question, Figure 3 plots the quantiles of the sales rank and the fraction of the data. About 60 percent of the data has a sales rank of less than 5000 and about 20 percent of the data has sales rank of between 5000 and 10000.

Table 6 presents the summary statistics of pooled ground coffee data that are sorted by sales rank and divided into four different quantiles. Although the number of ground coffee products in each of the quantile is the same (219), the attributes of the distribution of online consumer ratings showed significant changes. The average sales rank increased from approximately 613 in the first quantile to 2914 in the second quantile to 5961 in the third quantile and finally to 11431 in the fourth quantile. On average, the number of ratings and the average rating decreased from 143 and 4.3 in the first quantile to 2 and 2.4 in the fourth quantile, respectively. The average skewness of the ground coffee products also increased from -1.8 in the first quantile to -0.21 in the fourth quantile. The changes in the standard deviation across the quantiles did not follow a particular pattern. These results indicate that, on average, products with higher sales are more likely to receive more ratings, higher average ratings and are more negatively skewed distribution of ratings.[13]

[13] The fact that ground coffee products with higher sales also have higher number of ratings, average rating and more negatively skewed distribution raises the question of whether there is a problem of reverse causality between sales and the distribution of online consumer ratings. This potential problem is dealt with by the fact that actual sales data are not used in this study. However, the best sellers rank is used. The best sellers rank appears to be a suitable proxy or instrumental

Table 7 reports the results of pooled quantile regressions specified at the 25th, 50th and 75th quantiles.[14] The results reported in Table 6 show that the effect of the distribution of online consumer ratings has similar signs but vary in sizes with respect to the quantile of the sales rank. The coefficients of the attributes of the distribution of the online consumer ratings are largely smaller at the 75th percentile relative to 50th and 25th percentile. This result indicates that the effect of the distribution of online consumer ratings increase with lower sales. In terms of elasticities, the absolute value of the elasticity of the number of ratings is 0.012 at the 75th quantile, 0.015 at the 50th quantile and 0.016 at the 25th quantile. Thus, the elasticity of the number of rating increases for lower sales ground coffee products. Similarly, the absolute value of the elasticity of the average rating and skewness of ratings increased from the 75th to the 25th quantiles. This shows that the effect of the distribution of ratings is higher for lower sales products.

7 Robustness Checks

Because the results presented in Tables 3-6 are specific to ground coffee products on Amazon.com (platform) that are sold by Amazon.com (online retailer) in the United States, a few other questions may be raised regarding the robustness of the findings. For example, do the main findings hold for the entire population of ground coffee products sold on Amazon.com? To address this question, both ground coffee products sold by Amazon.com and those that are sold by other third-party sellers on Amazon.com were collected and analyzed. Summary statistics of these ground coffee

---

variable for sales which eliminates potential issues about endogeneity. In addition, the study cautions against making strong causal claims with regards to the relationship between the best sellers rank and the distribution of online consumer rating.

[14] To justify the use of the quantile regression, the Breusch-Pagan heteroscedasticity test was carried out and the test statistic was significantly different from zero (Breusch and Pagan, 1979). This confirms that there is heteroscedasticity and justified the use of quantile regression.

products are presented under Amazon.com in Table 8. The Amazon.com column of Table 8 summarizes the statistics of 5354 ground coffee products that are available on Amazon.com as of June 2019. Table 9, column Amazon.com, reports OLS estimates of the effect of the distribution of online consumer ratings on ground coffee sales. The results indicate that the estimates have the expected signs but differ slightly in magnitude and in elasticities.

Another question that one can pose regarding the robustness of the main findings is whether the results hold in another country.[15] Given that Amazon operates in over 10 different countries and has different platforms in each of these countries, it is expected that the main results would hold in each of these countries and in any of its platforms, aside from Amazon.com. To investigate this, all ground coffee data on Amazon.co.uk, which was established three years after Amazon.com, were collected. Table 7 also presents the summary statistics for all the ground coffee on Amazon.com as well as the two rounds of ground coffee data collected in December 2018 and June 2019 from Amazon.co.uk. A total of 1031 ground coffee products are matched in the sample, which includes products sold by Amazon.co.uk and other sellers that operate on its platform. Column Amazon.co.uk of Table 8 presents random effects estimates of the effect of the distribution of online consumer ratings on ground coffee sales. The results are similar to those previously presented in terms of expected signs but the magnitudes and elasticities are slightly different.

Furthermore, an identifying assumption used in this study is that the relationship between sales rank and price is not endogenous. Assuming that it is, then the results presented would be

[15]In addition, both in-sample and out-of-sample predictions of the sales ranks were carried out and the results showed stability in the estimates. The entire analysis presented in this paper was also conducted with other grocery products: On Amazon.com, baby formula and chocolate gifts were used while on Amazon.co.uk, red wine and cereal bars were used. The results of this analysis showed similar findings irrespective of the differences in grocery products.

biased. One way to deal with an endogenous variable is to use a relevant and valid instrument for prices of ground coffee on Amazon.com. This was carried out by matching the data from Amazon.com and Amazon.co.uk and then using the prices from Amazon.co.uk as an instrument for prices from Amazon.com. This is justified because the prices of ground coffee from Amazon.co.uk is strongly correlated with the prices of ground coffee from Amazon.com, but it is uncorrelated with the sales rank of ground coffee from Amazon.com. Thus, the prices from Amazon.co.uk are exogenous and only affects the sales rank of the ground coffee from Amazon.com through their actual prices. The matched data between Amazon.com and Amazon.co.uk resulted in 179 ground coffee products, which are summarized in Table 7. An instrumental variables estimation was carried out via two-stage least squares and the second-stage results are presented in Table 8. These results are largely similar to previous findings in terms of signs and magnitudes of coefficients, as well as elasticities.

8 Conclusions and Directions for Future Work

This study examines how the distribution of online consumer ratings provide useful information about the quality of ground coffee products, which translates to coffee sales. The study estimates the effect of the three moments of the online consumer rating distribution—the average rating, the variance and the skewness of rating distribution, as well as the number of ratings, using panel data of ground coffee on Amazon.com. It was found that each of these attributes of the distribution of online consumer ratings affects coffee sales. Quantile regression results suggest that the effect of the distribution of ratings on ground coffee sales varies across the different quantiles of the sales rank, with products with lower sales seeing more effects than products with higher sales. Robustness checks for all the findings are provided by considering all ground coffee products sold on both Amazon.com, ground coffee products sold on Amazon.co.uk and using an instrumental variable approach to validate the findings reported.

With the expected growth in online grocery markets around the world, these results would be useful to producers of products sold online, e-commerce platforms and online retailers. In particular, these results provide retail managers an array of attributes of online consumer ratings to employ in their demand forecasts. For example, retail managers at Amazon, Walmart or eBay, and sellers that operate on these platforms can easily calculate the skewness of rating distribution and use this information to predict future demand for their grocery products. This can also be easily extended to other product categories.

This study builds on previous studies that investigate individual measures of online ratings such as the average rating and the variance of ratings. But there are still opportunities for future research in this area especially in the context of online grocery markets. For example, it may be interesting to replicate this study in other platforms such as eBay and Walmart. It may also be interesting to investigate interaction effects of the attributes of the distribution of online ratings as well as other components of the online consumer reviews that are becoming important over time, such as helpful votes, top reviews, review dates, answered questions, etc. and how these attributes affect grocery sales.

References

Anderson, M. and Magruder, J. (2012). Learning from the crowd: Regression discontinuity estimates of the effects of an online review database. The Economic Journal, 122(563):957–989.

Beck, B. (2017). Amazon is the new Google for Product Search. https://www.digitaldoughnut.com/articles/2017/november/amazon-is-the-new-google-for-product-search. Accessed 9/5/2019.

Breusch, T. S., and A. R. Pagan. 1980. The Lagrange multiplier test and its applications to model specification in econometrics. Review of Economic Studies. 47: 239-253.

Breusch, T. S. and Pagan, A. R. (1979). A Simple Test for Heteroskedasticity and Random Coefficient Variation. Econometrica. 47 (5): 1287–1294

Chevalier, J. A. and Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. Journal of marketing research, 43(3):345–354.

Clemons, E. K., Gao, G. G., and Hitt, L. M. (2006). When online reviews meet hyperdifferentiation: A study of the craft beer industry. Journal of management information systems, 23(2):149–171.

Chiu, H. 2010. Skewness preference, risk taking and expected utility maximisation. Geneva Risk Insurance Rev. 35(2) 108–129.

Cui, G., Lui, H.-K., and Guo, X. (2012). The effect of online consumer reviews on new product sales. International Journal of Electronic Commerce, 17(1):39–58.

Dalvi, N., Kumar, R., and Pang, B. (2013). Para'normal'activity: On the distribution of average ratings. In Seventh International AAAI Conference on Weblogs and Social Media.

Dellarocas, C., Zhang, X. M., and Awad, N. F. (2007). Exploring the value of online product reviews in forecasting sales: The case of motion pictures. Journal of Interactive marketing, 21(4):23–45.

Duan, W., Gu, B., and Whinston, A. B. (2008). Do online reviews matter?—an empirical investigation of panel data. Decision support systems, 45(4):1007–1016.

Garrett, T. A., and Russell S. S., (1999). Gamblers favor skewness, not risk: Further evidence from United States' lottery games, Economics Letters 63, 85-90.

Golec, J., and Tamarkin M. (1998). Bettors Love Skewness, Not Risk, at the Horse Track, Journal of Political Economy 106(1): 205-225.

Hanks, S. and Spils, D. (2006). Increases in sales rank as a measure of interest. US Patent 7,058,599. Google Patents.

Heng Y., Z. Gao, Y. Jiang, and X. Chen, (2018). Exploring hidden factors behind online food shopping from Amazon reviews: a topic mining approach," Journal of Retailing and Consumer Services, vol. 42, pp. 161–168.

Hu, N., Zhang, J., and Pavlou, P. A. (2009). Overcoming the j-shaped distribution of product reviews. Communications of the ACM, 52(10):144–147.

Ishihara, M. and Liu, Y. (2017), A Dynamic Structural Model of Endogenous Consumer Reviews in Durable Goods Markets. https://dx.doi.org/10.2139/ssrn.2728524

Khare, A., Labrecque, L. I., and Asare, A. K. (2011). The assimilative and contrastive effects of word-of-mouth volume: An experimental examination of online consumer ratings. Journal of Retailing, 87(1):111–126

Koenker, Roger. (2005). Quantile Regression. Cambridge: Cambridge University Press

Langan, R., Besharat, A., and Varki, S. (2017). The effect of review valence and variance on product evaluations: An examination of intrinsic and extrinsic cues. International Journal of Research in Marketing, 34(2):414–429.

Lin, W., Ortega, D., and Caputo, V. (2018). Are ex-ante hypothetical bias calibration methods context dependent? Evidence from online food shoppers in China. Journal of Consumer Affairs, 1-25.

Luca, M. (2016). Reviews, reputation, and revenue: The case of yelp. com. (March 15, 2016). Harvard Business School NOM Unit Working Paper, (12-016).

Marinescu, I., Klein, N., Chamberlain A., and Smart, M. (2018). Incentives can reduce bias in online employer reviews, Academy of Management Proceedings, vol 2018(1).

Mayzlin, D., Y. Dover, and J. Chevalier. (2014). Promotional Reviews: An Empirical Investigation of Online Review Manipulation. American Economic Review 104 (8): 2421–55.

Park, S.-B. and Park, D.-H. (2013). The effect of low-versus high-variance in product reviews on product evaluation. Psychology & Marketing, 30(7):543–554.

Schoenmueller, V., Netzer, O., and Stahl, F. (2018). The extreme distribution of online reviews: Prevalence, drivers and implications.

Shrestha, K. (2017). Stats you need to know about online reviews. Retrieved from Vendasta: https://www. vendasta. com/blog/50-stats-you-need-to-know-aboutonlinereviews.

Sun, M. (2012). How does the variance of product ratings matter? Management Science, 58(4):696–707.

Unnevehr L., James Eales, Helen Jensen, Jayson Lusk, Jill McCluskey, Jean Kinsey, (2010). Food and Consumer Economics, American Journal of Agricultural Economics, Volume 92, Issue 2, April 2010, Pages 506–521,

Wang, F., Liu, X., and Fang, E. E. (2015). User reviews variance, critic reviews variance, and product sales: An exploration of customer breadth and depth effects. Journal of Retailing, 91(3):372–389.

Zervas, G., Proserpio, D., and Byers, J. (2015). A first look at online reputation on airbnb, where every stay is above average. https://ssrn.com/abstract=2554500.

Zhu, F. and Zhang, X. (2006). The influence of online consumer reviews on the demand for experience goods: The case of video games. ICIS 2006 Proceedings, page 25.

Zimmermann, S., Herrmann, P., Kundisch, D., and Nault, B. R. (2018). Decomposing the variance of consumer ratings and the impact on price and demand. Information Systems Research, 29(4):984–1002.

Figures

Figure 1: Scatterplots of Observations and Variables

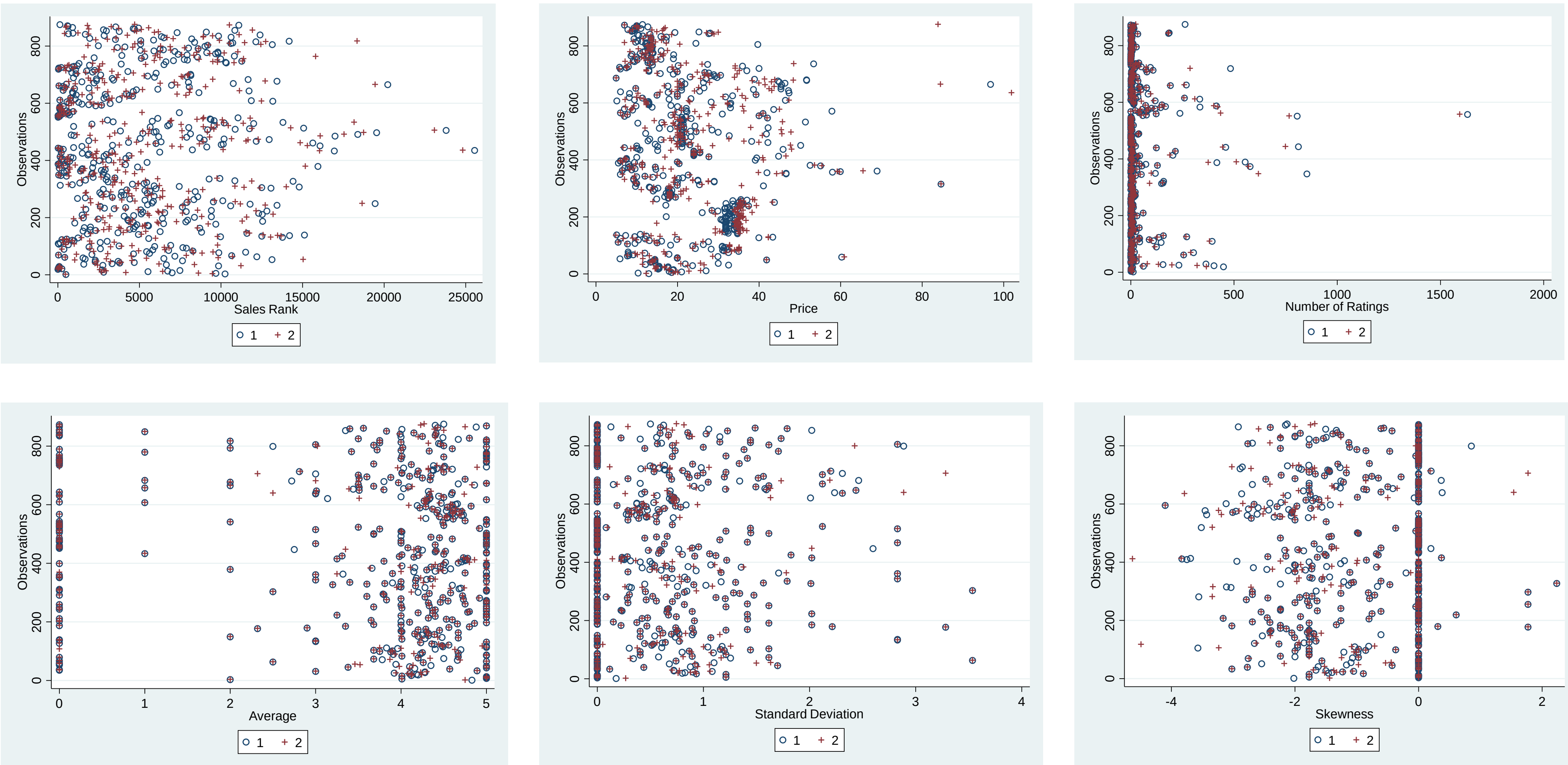


Figure 2: Means of the Fractions of the Distribution of Online Consumer Ratings

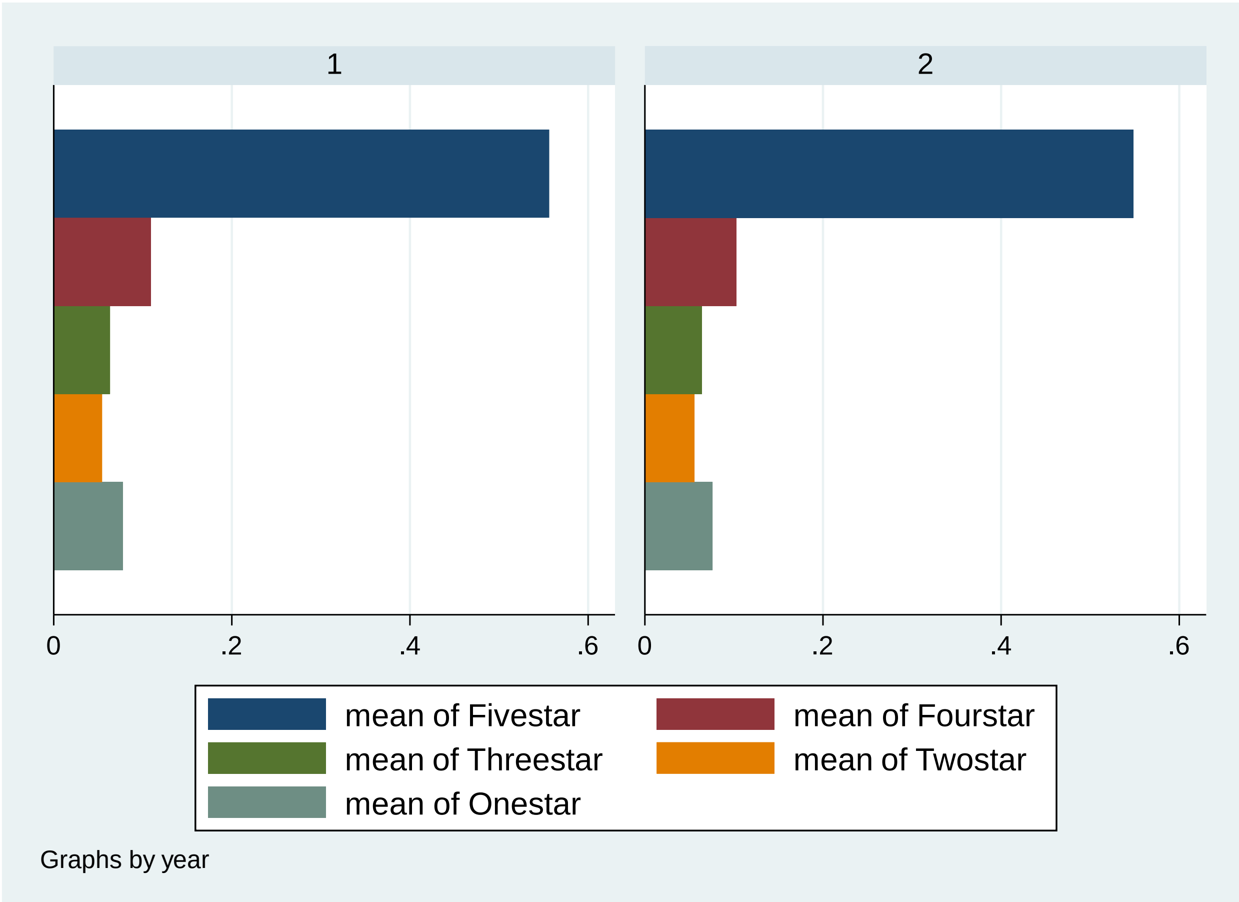
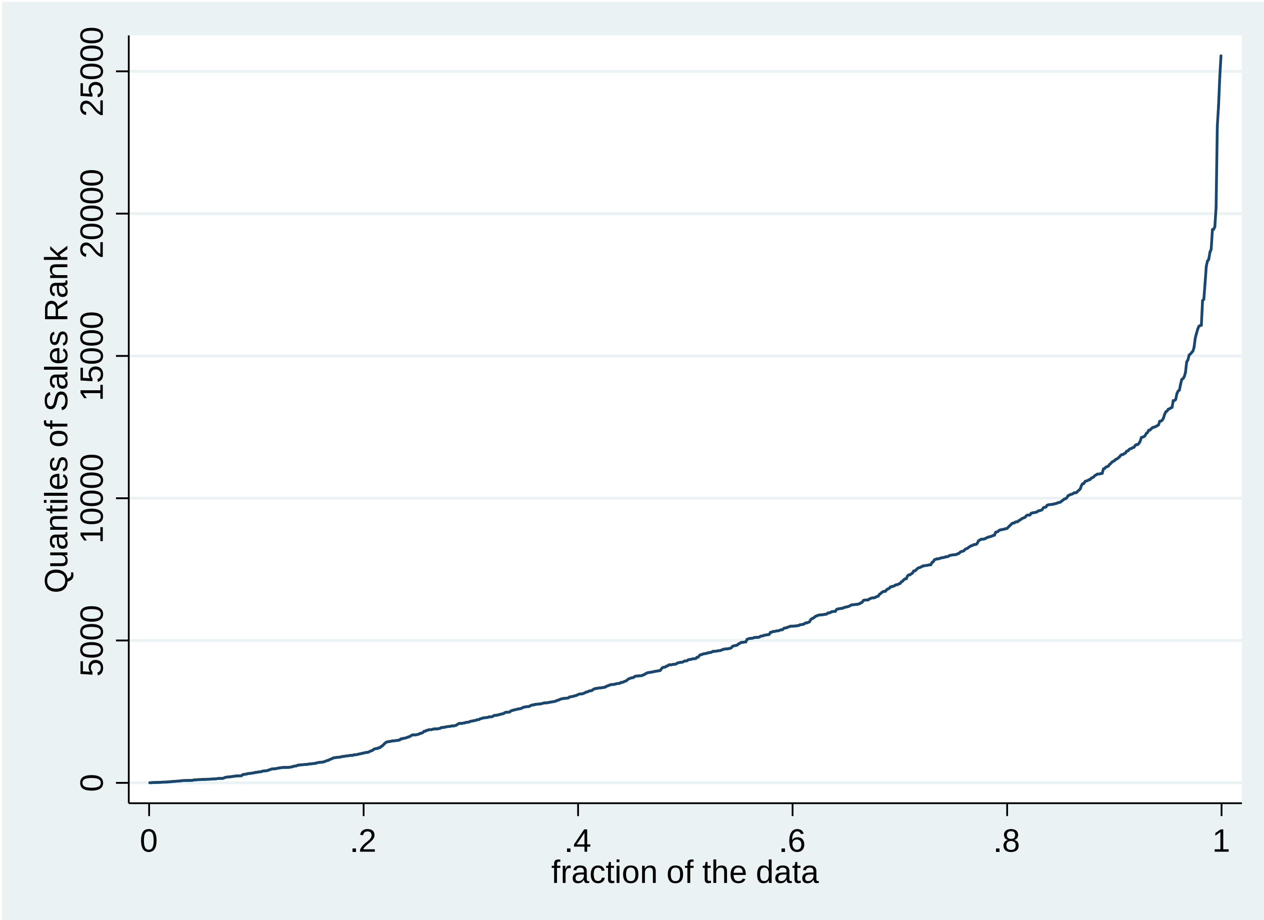


Figure 3: Sorted Sales Rank of Ground Coffee Products by Quantiles



Tables

Table 1: Summary of Related Literature

Authors	Research Summary	Product(s)	Platform(s)
	Number/Volume of Reviews		
Chevalier and Mayzlin (2006)	Positively associated with sales	Books	Amazon, Barnes & Noble
Clemons et al. (2006)	Insignificant association with sales growth	Beer	Ratebeer.com
Cui et al. (2012)	Positive association with new product sales	Video games, Electronics	Amazon
Dellarocas et al. (2007)	Positive association with future sales	Movies	Yahoo!Movies, BoxOfficeMojo, etc
Khare et al. (2011)	Effect differs by consumer	Movies	Experiment
	Average Rating		
Anderson and Magruder (2012)	Positive association with reservation	Restaurant	Yelp
Chevalier and Mayzlin (2006)	Positive association with sales	Books	Amazon, Barnes & Noble
Duan et al. (2008)	Insignificant association with sales	Movies	Yahoo!Movies, BoxOfficeMojo, etc
Luca (2016)	Positive association with revenue	Restaurant	Yelp
Zhu and Zhang (2006)	Positive association with sales	Video games	Gamespot, NPD Fun Group
	Variance of Ratings		
Langan et al. (2017)	Effect is moderated by intrinsic and extrinsic factors	Laptop, Camera	Consumer survey panel
Park and Park (2013)	Effect depends on prior expectation, product type	Unisex perfume, MP3 Player	Lab Experiment
Sun (2012)	High variance and low average increase sales	Books	Amazon, Barnes and Noble
Wang et al. (2015)	Effect may be positive, negative or insignificant	Movies, Cameras, Books	Experiment, Amazon, BoxOfficeMojo, etc.
Zimmermann et al. (2018)	High variance increases price and decreases demand	Theory	Theory
	Skewness of Ratings		
Dalvi et al. (2013)	Ratings are log-normally distributed	Restaurants, Movies	Yelp, Amazon
Hu et al. (2009)	Ratings are overwhelmingly positive	Books, DVDs, Videos	Amazon
Schoenmueller et al. (2018)	Ratings are extremely distributed	Broad	27 Platforms
Zervas et al. (2015)	Most ratings are above average	Housing, Travel	Airbnb, Tripadvisor

Table 2: Summary Statistics of Ground Coffee on Amazon.com

	June 2019	March 2019	June 2019	March 2019
	Cross-sectional data		Matched data	
	Mean (Std)	Mean (Std)	Mean (Std)	Mean (Std)
Sales rank	5782.24	4922.04	5229.59	5230.82
	(5048.49)	(4258.80)	(4472.16)	(4349.63)
Price ($)	22.55	23.85	22.54	23.58
	(17.59)	(15.02)	(12.79)	(13.18)
Number of ratings	43.12	51.02	43.44	39.12
	(135.04)	(142.86)	(130.50)	(121.12)
Average rating	3.21	3.51	3.59	3.53
	(1.87)	(1.70)	(1.61)	(1.66)
Standard deviation of ratings	0.58	0.61	0.65	0.64
	(0.73)	(0.70)	(.70)	(0.70)
Skewness of ratings	-0.86	-0.97	-1.00	-0.96
	(1.08)	(1.11)	(1.09)	(1.09)
Fraction of five-star rating	0.49	0.55	0.56	0.55
	(0.36)	(0.35)	(0.34)	(0.34)
Fraction of four-star rating	0.10	0.10	0.11	0.10
	(0.18)	(0.17)	(0.18)	(0.17)
Fraction of three-star rating	0.06	0.06	0.06	0.06
	(0.14)	(0.13)	(0.13)	(0.13)
Fraction of two-star rating	0.05	0.05	0.05	0.06
	(0.14)	(0.15)	(0.15)	(.15)
Fraction of one-star rating	0.07	0.07	0.08	0.08
	(0.16)	(1.41)	(0.16)	(0.16)
Fraction of no ratings	0.22	0.16	0.14	0.15
N	1,221	810	438	438

Table 3: Effect of the Distribution of Online Consumer Ratings on Ground Coffee Sales

	(1)	Elasticity	(2)	Elasticity	(3)	Elasticity	(4)	Elasticity
Price	-26.023**	-0.029	-17.843*	-0.012	-16.502*	-0.018	-17.057*	-0.019
	(10.608)		(9.636)		(9.558)		(9.627)	
Number			6.998***	0.014	115.382***	0.229	6.956***	0.014
			(1.246)		(28.321)		(1.242)	
Average			1437.637***	0.247	1514.960***	0.260	1392.437***	0.239
			(264.461)		(260.840)		(264.270)	
Std Dev			958.209***	0.030	819.798***	0.025	-976.208	-0.030
			(262.325)		(261.181)		(1017.051)	
Skewness			-1053.813***	0.050	-1135.536***	0.054	-933.035***	0.044
			(136.408)		(136.184)		(150.134)	
No ratings			4448.749***	0.031	4844.864***	0.034	4394.952***	0.031
			(1253.017)		(1237.267)		(1247.990)	
Number*Average					-24.429***			
					(6.377)			
Average*Std							619.326**	
							(314.564)	
Constant	21369.801***		13468.565***		13026.794***		13467.759***	
	(318.644)		(1238.390)		(1223.607)		(1233.148)	
N	876		876		876		876	
Overall R^2	0.019		0.382		0.407		0.389	

Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Note: Dependent variable is 26000-Sales Rank.

Table 4: Effect of the Distribution of Online Consumer Ratings on Ground Coffee Sales

	(1)	Elasticity	(2)	Elasticity
Price	-14.085	-0.016	-13.409	-0.015
	(10.090)		(10.114)	
Number	8.883***	0.018	72.173	0.143
	(1.309)		(70.974)	
Five Star	4027.810***	0.107	4044.996***	0.108
	(480.583)		(478.337)	
Four Star	2203.807***	0.011	2197.496***	0.011
	(747.862)		(754.145)	
Three Star	2754.584**	0.008	2288.468*	0.007
	(1342.020)		(1374.211)	
Two Star	-1865.759	-0.005	-1928.533	-0.005
	(1167.464)		(1176.069)	
One Star	-127.599	-0.000	-915.134	-0.003
	(1088.218)		(1118.350)	
Number*Five			-71.233	
			(73.145)	
Number*Four			-75.509	
			(89.399)	
Number*Two			-104.528	
			(109.246)	
Number*One			46.822	
			(97.312)	
Constant	18206.977***		18183.968***	
	(468.424)		(466.148)	
N	876		876	
Overall R^2	0.274		0.297	

Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Note: Dependent variable is 26000-Sales Rank

Table 5: Effect of the Distribution of Online Ratings with Time-Invariant Characteristics

	(1)		Elasticity
Price	-22.682**	(11.298)	-0.025
Number	4.174***	(1.185)	0008
Average	1558.207***	(244.719)	0.267
Stand Dev	944.893***	(242.742)	0.029
Skewness	-871.620***	(130.247)	0.041
No ratings	5444.977***	(1166.574)	0.038
Roast-type			
Medium	1438.316***	(536.192)	0.007
Dark	-70.413	(546.184)	-0.000
Light	-2052.950*	(1187.671)	-0.002
Country of Origin			
Colombian	-71.159	(989.888)	-0.000
Brazilian	665.299	(3073.420)	0.000
Ethiopian	-150.054	(3061.635)	-0.000
Guatemala	-3908.894*	(2169.442)	-0.000
Costa Rican	1246.224	(3093.794)	0.000
Bean-type			
Arabica	-1719.264*	(880.713)	-0.003
Specialty label			
Fair Trade	1518.760	(1395.155)	0.001
Organic	538.281	(612.292)	0.002
Caffeine-type			
Decaffeinated	-1664.783**	(670.345)	-0.005
Brand			
Amazon Fresh	1493.933	(1328.719)	0.001
Coffee Masters	-1319.213***	(451.571)	-0.012
Coffee Fool	-3731.334***	(517.515)	-0.023
Coffee Crema	-3560.059***	(508.724)	-0.021
Packaging			
Package Size	-10.782	(7.703)	-0.016
Constant	14768.789***	(1169.727)	
N	876		
Overall R^2	0.506		

Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: Dependent variable is 26000-Sales Rank

Table 6: Summary Statistics of Pooled Coffee Products by Quantiles: Sorted by Sales Rank

	Mean	Stand Dev
Sales rank	613.54	495.99
Price ($)	20.49	14.35
Number of ratings	143.22	220.09
Average rating	4.26	0.76
Standard deviation of ratings	0.64	0.44
Skewness of ratings	-1.77	0.64
Number of observations	219	
Sales rank	2914.30	731.45
Price ($)	23.40	12.51
Number of ratings	15.09	31.75
Average rating	3.95	1.26
Standard deviation of ratings	0.71	0.60
Skewness of ratings	-1.23	1.07
Number of observations	219	
Sales rank	5961.30	1083.46
Price ($)	24.43	10.55
Number of ratings	5.01	9.57
Average rating	3.63	1.60
Standard deviation of ratings	0.67	0.70
Skewness of ratings	-0.712	0.91
Number of observations	219	
Sales rank	11431.69	3176.69
Price ($)	23.91	13.93
Number of ratings	1.78	3.15
Average rating	2.4	2.01
Standard deviation of ratings	0.56	0.95
Skewness of ratings	-0.21	0.75
Number of observations	219	

Table 7: Effect of the Distribution of Online Ratings on Ground Coffee Sales: Quantile Regression

	Quantile Regression at 0.25 (19577)	Elasticity	Quantile Regression at 0.50 (22658)	Elasticity	Quantile Regression at 0.75 (24931)	Elasticity
Price	-51.898**	-0.063	-35.877***	-0.039	-19.850***	-0.019
	(21.118)		(10.901)		(7.604)	
Number	7.493***	0.016	7.940***	0.015	6.829***	0.012
	(2.008)		(1.818)		(1.096)	
Average	1563.486***	0.293	1327.811***	0.222	1447.007***	0.222
	(411.558)		(260.413)		(222.959)	
Stand Dev	728.926*	0.024	1105.056***	0.033	1324.935***	0.037
	(408.544)		(385.476)		(248.417)	
Skewness	-1599.328***	0.083	-1461.099***	0.067	-849.640***	0.036
	(142.531)		(132.439)		(114.405)	
No ratings	3519.512*	0.027	3396.828**	0.023	4545.435***	0.029
	(1964.254)		(1340.491)		(1112.851)	
Constant	11750.254***		14396.149***		15946.563***	
	(1911.574)		(1194.608)		(1015.785)	
N	876		876		876	
R^2	0.381		0.386		0.372	

Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Summary Statistics of all Ground Coffee on Amazon.com and Amazon.co.uk

	Amazon.com	Amazon.co.uk		Amazon.com*
	July 2019	June 2019	December 2019	July 2019
	Mean (Std Dev)	Mean (Std Dev)	Mean (Std Dev)	Mean (Std Dev)
Sales Rank	7419.26 (5750.09)	1781.01 (1278.19)	1584.88 (1138.58)	8212.85 (6818.542)
Price ($)	29.21 (29.62)	18.97 (21.09)	19.19 (21.51)	26.16 (21.35)
Number of ratings	18.30 (106.05)	7.54 (31.81)	6.33 (27.42)	32.77 (84.26)
Average rating	2.65 (2.26)	2.22 (2.28)	1.99 (2.27)	2.46 (2.29)
Standard deviation of ratings	0.27 (0.46)	0.23 (0.52)	0.19 (0.50)	0.22 (0.39)
Skewness of ratings	-0.66 (1.06)	-0.47 (0.99)	-0.41 (0.94)	-0.76 (1.18)
Fraction of five-star rating	0.44 (0.42)	0.36 (0.42)	0.33 (0.42)	0.41 (0.43)
Fraction of four-star rating	0.06 (0.16)	0.06 (0.18)	0.05 (0.16)	0.07 (0.19)
Fraction of three-star rating	0.04 (0.12)	0.03 (0.11)	0.03 (0.11)	0.02 (0.09)
Fraction of two-star rating	0.02 (0.05)	0.02 (0.09)	0.02 (0.09)	0.01 (0.03)
Fraction of one-star rating	0.02 (0.06)	0.03 (0.14)	0.03 (0.13)	0.02 (0.06)
Fraction of no ratings	0.41	0.49	0.55	0.46
N	5354	1027		179

*Note: Only products that appear both on Amazon.com and Amazon.co.uk

Table 9: Effect of the Distribution of Online Ratings on Ground Coffee Sales

	Amazon.com All ground coffee	Elasticity	Amazon.co.uk[a] All ground coffee	Elasticity	Amazon.com* All ground coffee	Elasticity
Price	-18.665***	-0.029	-4.584***	0.030	-36.967	-0.05
	(2.120)		(1.229)		(27.256)	
Number	3.298***	0.003	3.595***	0.009	8.332*	0.015
	(0.598)		(1.022)		(4.519)	
Average	759.456***	0.108	206.774***	0.149	3866.796***	0.535
	(274.568)		(41.021)		(1491.602)	
Stand Dev	1489.335***	0.021	180.190***	0.013	4871.971***	0.059
	(262.632)		(52.378)		(1511.520)	
Skewness	-1474.930***	0.052	-206.016***	0.031	-1160.352***	0.049
	(70.173)		(31.024)		(374.910)	
No rating	-297.593	-0.007	198.095	0.035	12039.248*	0.310
	(1340.684)		(189.752)		(7179.999)	
Constant	15809.833***		2312.251***		1515.816	
	(1336.261)		(190.662)		(7211.620)	
N	5354		2054		179	
(Overall) R^2	0.381		0.371		0.593	
Model Type	Ordinary Least Squares		Random Effects Model		Instrumental Variables	

t statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

[a]The dependent variable for Amazon.co.uk is 4600-Sales Rank.

*Note: Products from all ground coffee in the UK and the US were matched, and the prices for ground coffee in the UK were used as an instrument for prices of ground coffee in the US.