The Effect of Price on E-Commerce Platforms: Statistical Evaluation of Amazon's Pricing Strategy

Xinguo Luo[a]
Institute of Problem Solving, University of Toronto, 27 King's College Cir, Toronto, Canada

Keywords: Amazon, Online Retailing, Sales, Customer Rating, Pricing Strategy.

Abstract: This present study is meant to explain whether Amazon's pricing strategies, like providing discounts or not, have some benefits for product sales and consumer ratings (i.e., customers' satisfaction with the products). The original dataset was downloaded from the Kaggle platform, and it contains sales records of about 1.4 million products during September 2023. However, in order to fit this present study better, a simplified dataset with 166,481 samples was filtered from the 1.4 million samples of the original dataset according to the average monthly consumption of Americans. The method of this study was simple linear regression, which can be used to understand the association between two variables. Then, two linear models were generated with Excel. After all, the results showed that the p-values are statistically significant for both discounts and sales and discount and customer ratings. In more detail, discounted products have better sales than regular goods, and they have better customer evaluation.

1 INTRODUCTION

1.1 Background

With the growth of the Internet and electronic devices in recent years, online retail platforms have experienced significant development in many aspects. More and more people like to browse and buy what they want through e-commerce platforms. According to Jap et al. (2022), in 2020, some famous online marketplaces such as Amazon, Taobao, and T-mall earned $2.7 trillion in global sales, which held the majority of global online sales of that year. Besides selling things, Amazon has entered even more different areas these years, such as education. All of this shows that e-commerce platforms like Amazon have already become a substantial part of the global economy.

1.2 Literature Review

Studies indicate that many people visit shopping websites several times a week and shop at online retail stores at least once a month (Menon, 2023). Each platform has different advantages, and it is important to build suitable and customized sales strategies to develop an online sales platform (Kwak et al., 2019; Zhang et al., 2023). For example, some products sold on Amazon.com always remain at a low price or with a high discount rate (Reimers & Waldfogel, 2017).

Pricing is an important but complex part of retail, especially for online retailing. Unlike traditional offline stores, people can see the price of products more directly when shopping online and may be more sensitive to price. Online retail prices relate to many aspects, such as customer satisfaction, sales, and profit (Nataraja et al., 2017). In this case, although the price of goods may change frequently according to market conditions, many e-commerce platforms still implement computer algorithms to set prices automatically to increase profits (Aparicio et al., 2024). Some companies even rely on high sales to stay afloat at extremely low prices (Sussman, 2019). Consequently, this study will mainly focus on Amazon's pricing strategy and its potential influence. It seems that lower and discounted prices can help increase customer satisfaction and sales.

[a] https://orcid.org/0009-0001-3926-6151

Luo, X.
The Effect of Price on E-Commerce Platforms: Statistical Evaluation of Amazon's Pricing Strategy.
DOI: 10.5220/0012925300004508
Paper published under CC license (CC BY-NC-ND 4.0)
In *Proceedings of the 1st International Conference on Engineering Management, Information Technology and Intelligence (EMITI 2024)*, pages 239-242
ISBN: 978-989-758-713-9
Proceedings Copyright © 2024 by SCITEPRESS – Science and Technology Publications, Lda.

2 METHOD

2.1 Data

The original dataset found on the Kaggle website includes extensive information about Amazon's product sales in September 2023. It contains the records of 1,426,337 Amazon products, each with a product ID, a title, links to product images, official links to the products, rating, number of reviews, present price, primary price, category, bestseller status, and number of products sold on Amazon in September 2023 (Asaniczka, 2023). In these 1.4 million samples, the price of the products ranges widely from $0.01 to $19,700. However, according to the U.S. Bureau of Labour Reports (2023), the 2022 annual expenditures for all consumers in the United States averaged $72,967, and only 15.2% of that spending is likely to be related to online shopping. In other words, Amazon consumption will be at most $924.25 per month for most American consumers. The products sold for over $924.25 seem less familiar for daily purchases and were removed from the dataset. Also, products under $50 are deleted since the discount rate is too small. In addition, because some of the product information is incomplete, those products will not be analyzed.

Therefore, the simplified dataset has 166,481 products in 270 categories. The price interval becomes $50 to $924.25, with some discounts and some not. Besides, product ratings range from 0 to 5 stars, with an average rate of 3.8 (SD = 1.51), which shows that most customers are satisfied with the goods. The top seller in September 2023 was a refrigerator filter that sold 40,000 pieces in a month and had never been a best seller before.

2.2 Design

2.2.1 Pricing Strategy and Sales

A linear regression model will be generated through Excel. The independent variable is the pricing strategy, which determines whether the product is discounted. It is a dummy variable that was created for non-quantitative classification. If the present price is greater or equal to the original listed price, it is considered to have no discount (i.e., discount equals 0). Otherwise, if the current price is smaller than the primary one, a discount was implemented and could be represented by 1. The response variable is the number of sales in September 2023 collected from Amazon.com. This model aims to find the relationship between price discounting activity and sales volume.

The original formula for this linear regression model is:

$$y_1 = \beta_0 + \beta_1 * x_1 + \epsilon_1 \quad (1)$$

where $y1$ refers to the number of sales, $x1$ refers to whether there is a discount or not, $\beta 1$ shows how the sales will change with the discount, $\beta 0$where y_1 refers to the number of sales, x_1 refers to whether there is a discount or not, β_1 shows how the sales will change with the discount, β_0 is the number of sales when there is not any account, and $\epsilon_1$is the influence of other confounding factors.

2.2.2 Pricing Strategy and Customer Rating

This study analysed whether discounts impact customers' satisfaction with the products they buy. Whether there is a discount is an independent variable, and the response variable is the stars of the products. To be more specific, the stars are the scores the products got that aim to evaluate customers' satisfaction with the products. The statistical method used is also the linear regression model:

$$y_2 = \beta_0' + \beta_1' * x_2 + \epsilon_2 \quad (2)$$

$y_2$is the customer rating of products ranging from 0 to 5, x_2 refers to whether there is a discount or not, β_1 explains how the rating will change with the discount, β_0 means the scores the products got when there is not any account, and ϵ_2 is the potential random factors.

3 RESULTS

3.1 Pricing Strategy and Sales

Table 1: The linear model of pricing strategy and sales.

	Coefficients	*t Stat*	*P-value*
Intercept	48.46824613	44.03809732	0
Discount	56.7219767	26.77102313	1.5201E-157

Table 2: The regression statistics of pricing strategy and sales.

Multiple R	.06547152
R Square	.00428652
Adjusted R Square	.00428054
Standard Error	383.728542
Observations	166481

Table 3: ANOVA test of pricing strategy and sales.

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	105530536.3	105530536.3	716.6876787	1.5201E-157
Residual	166479	24513632187	147247.5939		
Total	166480	24619162723			

According to the regression model in Table 1, some statistical relationships between the price strategy and the sales quantity demonstrated that:

$$y_1 = 48.4682461 + 56.7219767x_1 + \epsilon_1 \quad (3)$$

Table 2 exhibits that the R-squared value for the linear regression is about 0.004, suggesting that the model could only explain a small portion of the sales variability and that there are many other factors affecting sales the model did not include. However, the p-value in Table 1 is much smaller than the significant level, $p < .001$, representing that the influence of discount for sales is statistically significant. Since the coefficient is greater than 0, there is a positive association between the pricing strategy and sales, which means the sales will increase as the discount gets bigger. Also, the model is a good fit for the data because Table 3 represents the significance of the F-statistic is less than 0.05.

3.2 Pricing Strategy and Customer Rating

Table 4: The regression statistics of pricing strategy and starts of products.

Multiple R	.06547152
R Square	.00428652
Adjusted R Square	.00428054
Standard Error	383.728542
Observations	166481

Table 5: ANOVA test of pricing strategy and starts of products.

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	105530536.3	105530536.3	716.6876787	1.5201E-157
Residual	166479	24513632187	147247.5939		
Total	166480	24619162723			

Table 6: The linear model of pricing strategy and starts of products.

	Coefficients	*t Stat*	*P-value*
Intercept	48.46824613	44.03809732	0
Discount	56.7219767	26.77102313	1.5201E-157

According to the tables above, the linear regression model of pricing strategy and customer rating shows a formula:

$$y_2 = 3.72731737 + 0.38655364x_2 + \epsilon_1 \quad (4)$$

The statistical data in Table 6 exhibits a statistical significance between the variables, $p < .001$. Because of the positive coefficient, the product rating would increase with the increase in the discount. Although the R-square in Table 4 indicates that only 1.299% of the variable could be explained by the model, the F-statistic result (i.e., significance F) of the ANOVA test in Table 5 is less than 0.05, indicating that the model can significantly fit the collected data.

Based on the analysis of two linear regression models above, because the two p-values in Table 1 and Table 6 are less than 0.01, there is a statistically significant association between the pricing strategy and the sales, as well as the pricing strategy and rating. According to the coefficients, the higher the discount, the higher the sales will increase significantly, and the customer's evaluation of the product will also increase. A suitable pricing strategy is vital for Amazon to operate better.

4 DISCUSSION

True to the original assumption, the discount and low-price strategy can improve product sales and ratings. Firstly, discounts can easily stimulate people's consumption. On the one hand, lower prices would make consumers more willing to order the target item. On the other hand, excitement may push consumers to make some purchases that are not in their primary plan, especially when discounts are high (Kim & Tanford, 2021). Secondly, because the price is low, the consumers' tolerance for the product has become higher. Although the goods may have some flaws, people will also feel satisfied because of the relatively lower price.

A proper promotion can bring many a lot of benefits to a company. Amazon understood this idea early on and provided a practical example. In the field of online book retailing, Amazon has beaten many competitors with the low prices they set (Reimers & Waldfogel, 2017). Amazon has never raised book

prices in the past two decades, even though few competitors have left (Reimers & Waldfogel, 2017). While most companies trying to compete with low prices have gone bankrupt, Amazon has already passed the deficit stage (Sussman, 2019) and increased revenue every year (Reimers & Waldfogel, 2017). Amazon has found a pricing strategy that works best for it.

However, although the results may match Amazon's actual situation, this present study has some limitations. Firstly, the dataset used in this study only includes Amazon's sales in the United States over a one-month period, which means it may not be a good interpretation for other countries or times. For example, some American holidays in September, like Labor Day, may become a factor that leads people to buy more items than usual for celebrating. Secondly, by refining the limited dataset, the study only discussed the influence of pricing strategies through the data collected but did not pay much attention to the products or the consumers themselves. Just as people are influenced by shopping values and price awareness, transactional tendencies, and coupons when shopping at the mall (Khare et al., 2014), so are they when shopping online. There are many environmental and personal factors that may affect the relationship between price, rating, and sales, but it is hard to collect and define them.

5 CONCLUSION

In conclusion, the preliminary analysis has been solved with the available variables: a promotion strategy positively affects Amazon's online retail store, both for sales and customer satisfaction with the products. Nevertheless, more potential variables are involved in the interaction in real life, and the pricing strategy may have a broader effect. As people continue to do more studies in wider fields and connect them with each other, more personal factors can be added to the discussion, allowing people to understand the pricing strategy and the customers better.

REFERENCES

Aparicio, D., Metzman, Z., & Rigobon, R. (2024). The pricing strategies of online grocery retailers. *Quantitative Marketing and Economics*, *22*(1), 1–21. https://doi.org/10.1007/s11129-023-09273-w

Asaniczka. (2023). *Amazon Products Dataset 2023 (1.4M Products)* [Data set]. Kaggle. https://doi.org/10.34740/KAGGLE/DS/3798081

Jap, S. D., Gibson, W., & Zmuda, D. (2022). Winning the new channel war on Amazon and third-party platforms. *Business Horizons*, *65*(3), 365–377. https://doi.org/10.1016/j.bushor.2021.04.003

Khare, A., Achtani, D., & Khattar, M. (2014). Influence of price perception and shopping motives on Indian consumers' attitude towards retailer promotions in malls. *Asia Pacific Journal of Marketing and Logistics*, *26*(2), 272–295. https://doi.org/10.1108/APJML-09-2013-0097

Kwak, J., Zhang, Y., & Yu, J. (2019). Legitimacy building and e-commerce platform development in China: The experience of Alibaba. *Technological Forecasting & Social Change*, *139*, 115–124. https://doi.org/10.1016/j.techfore.2018.06.038

Menon, S. (2023). *Amazon Consumer Behaviour Dataset* [Data set]. Kaggle. https://www.kaggle.com/datasets/swathiunnikrishnan/amazon-consumer-behaviour-dataset

Myers, S., Paulin, G. D., & Thiel, K. (2023). Consumer expenditures in 2022. U.S. *Bureau of Labor Statistics.* https://www.bls.gov/opub/reports/consumer-expenditures/2022/home.htm

Natarajan, T., Balasubramanian, S. A., & Kasilingam, D. L. (2017). Understanding the intention to use mobile shopping applications and its influence on price sensitivity. *Journal of Retailing and Consumer Services*, *37*, 8–22. https://doi.org/10.1016/j.jretconser.2017.02.010

Reimers, I., & Waldfogel, J. (2017). Throwing the Books at Them: Amazon's Puzzling Long Run Pricing Strategy. *Southern Economic Journal*, *83*(4), 869–885. https://doi.org/10.1002/soej.12205

Sussman, S. (2019). Prime predator: Amazon and the rationale of below average variable cost pricing strategies among negative-cash flow firms. *Journal of Antitrust Enforcement*, *7*(2), 203–219. https://doi.org/10.1093/jaenfo/jnz002

Zhang, T., Li, G., & Tayi, G. K. (2023). A strategic analysis of virtual showrooms deployment in online retail platforms. *Omega (Oxford)*, *117*, 102824-. https://doi.org/ 10.1016/j.omega.2022.102824