Master IREN

M. Michelangelo Rossi

E-commerce

An Empirical Analysis of Algorithmic Pricing on Amazon Marketplace

Par Emma Bourdit et Lisa Ngo

Janvier 2022

ABSTRACT :

An Empirical Analysis of Algorithmic Pricing on Amazon Marketplace

Le Chen, Alan Mislove and Christo Wilson

CONTENTS

Objectives

In this study, We develop a methodology for detecting algorithmic pricing, and use it empirically to analyze their prevalence and behavior on Amazon Marketplace.
We wanted to analyze the dynamic pricing strategies being used by sellers on Amazon.

Indeed, while algorithmic pricing can help online-shops become more competitive and improve revenue, it also introduces new obstacles.
First, in complex systems inhabited by other algorithms, poorly built pricing algorithms might interact in unexpected ways and potentially create unexpected consequences. For example, two competing dynamic pricing algorithms mistakenly boosted the price of a secondhand textbook on Amazon to $23M.
Second, dynamic pricing algorithms can be used to carry out collusive techniques that are harmful to consumers.
Indeed, the general population is currently unaware of the prevalence and behavior of algorithmic pricing algorithms in the wild.

This paper addresses an economic issue since misuse of the price algorithm can lead to collusion and market failure and thus harm consumers.
Indeed, Traditional retailers do not have real-time competitor pricing and are forced to manually re-label the new price of each product

We wish to empirically analyze the algorithmic pricing strategies deployed on Amazon Marketplace. Specifically, the authors want to understand what algorithmic pricing strategies are being used by marketplace participants, how prevalent these strategies are, and ultimately how they impact the customer experience.

For this, we decided to choose Amazon for three reasons :

- Amazon is the largest e-commerce site in the US and Europe
- Amazon is a marketplace populated by 3rd-party sellers and Amazon itself
- API provides capability to algorithmic pricing

Data

We studied four months of data covering all merchants selling any of the 1641 top-selling products. We took data from the top 20 sellers of each product every 25 minutes, including prices, ratings and other seller attributes.

For their study, we also assumed that sellers using algorithmic pricing are likely to base their prices at least partially on the prices of other sellers. This makes intuitive sense: for example, a seller who always wants to offer the lowest price on a specific product must set his price relative to the competitor with the lowest price.

Nevertheless, there are two noteworthy limitations to the dataset.

1) The data is biased toward the most popular products.

To remove this bias, wes randomly picked 2,158 goods from a public listing of all Amazon products to quickly measure this bias. we compare the product price and the number of sellers and as predicted, best-sellers products have far more sellers than random goods, as well as slightly lower pricing.

2) They do not take into account the type of Amazon account used.

we crawled data from Amazon using browsers that were not logged-in to Prime accounts. Although the exact number of Prime members is unknown, estimates place it at around 20–40% of all Amazon's customers. Thus, the dataset should accurately reflect what the majority of Amazon users see. However, Amazon may alter pages for Prime users, typically to highlight sellers and products that are eligible for expedited Prime shipping. Thus, some of the analysis and conclusions may not extend to Prime users.

Methodology

The authors estimated a structural model. They want to analyze the algorithmic pricing strategy of sellers. And an important criteria to do sales is the Buy Box because 80% of the sales are made through it. So the authors started by understanding the Buy Box algorithm, how it deals with prices, to then analyze the algorithmic pricing strategies.

In order to conduct this study, the authors used a 3-step methodology supporting the hypothesis that sellers using algorithmic pricing are likely to base their prices at least partially on the prices of other sellers.

The authors report that they wanted to use the Amazon Marketplace Web Services (MWS) API to collect seller and product pricing information. However, the API did not meet their needs. This is because the API does not return the identity of 3P sellers (only the price they chose), and the API is severely rate-limited.

Thus, the authors have built a detection algorithm that attempts to locate sellers who behave like "bots" ie sellers whose prices they set and timing of changes suggest algorithmic control

1. <u>They first define several target prices that the seller could match against.</u>

The authors select these target prices by examining popular repricing software for Amazon Marketplace, and choose three target prices for each product:

- lowest price
- Amazon's price
- the second lowest price.

For each (s, r) = seller/product, the authors created 4 real time series. One for the seller prices in time, one for the lowest price, one for the second lowest price, and the least one for the amazon price, where r is the product, p is the price, and t is the time :

- Sr = (ti, pi) = time serie time/price of the seller on the product r

- LOWr = (ti, pi) = time serie time/price of the lowest price on the product r
- 2NDr = (ti, pi) = time serie time/price of the second lowest price on the product r
- AMZNr = (ti, pi) = time serie time/price of amazon price on the product r

2. Comparison with Spearman's Rank Correlation

They calculate the similarity between Sr and LOWr, 2NDr, and AMZNr (respectively) using Spearman's Rank Correlation. When ρ is large, it means that the price changes contained in the pair of time series occur at the same moments

They mark these pairs as algorithmic pricing candidates :
- with $\rho \geq 0.7$ (the empirical cutoff of a strong positive correlation)
- and p-value ≤ 0.05

3. They filter the candidates

They hypothesize that the more the seller has a high correlation with the target price and makes a large number of price changes then there is a greater chance that the seller will use algorithmic pricing. Nevertheless, if there is a correlation, but a small number of price changes, the correlation is possibly coincidental. Thus, the authors impose a threshold of 20 price changes required for the candidate to be screened as algorithmic pricing candidates.

In sum, the key variables in the empirical model are : the price of the item, the time, threshold, ρ, and the p-value. The variation of the row identifies the key components of the empirical model, because it shows the correlation between the seller's price and the other time series prices.

Results

The authors highlight many results about the behavior of the algorithmic sellers.

1. The number of algorithmic sellers and price variations

Table 2 shows a set of sellers that the authors found to be doing algorithmic pricing.

Strategy	Threshold = 10		Threshold = 20	
	Sellers	Products	Sellers	Products
Lowest Price	726	544	426	408
Amazon Price	297	277	176	183
2nd Lowest Price	721	494	425	370
Total	918	678	543	513

Table 2: Number of sellers and products with detected algorithmic pricing, based on two different change thresholds. We use a change threshold of 20 unless otherwise stated.

First, we observe that many more sellers appear to be using the overall lowest price or the second lowest price as the target for their algorithmic pricing than Amazon's price. We can remind here that there is a source of randomness in the empirical model, because Amazon does not sell all products in our dataset, thus only a subset of seller/product pairs include AMZNr. We can also say that sellers can do different strategies that are not necessarily mutually exclusive.

Then, we can see that for threshold = 20, 543 sellers use algorithmic pricing. It represents 2.4% of all sellers in the author's dataset, and 38% of all sellers that have ≥20 changes for at least one product they sell.

Figure 14 shows an example where a seller (in red) has a clear strategy to always match the lowest price across all other sellers. In this figure, we can see four other sellers that offer the lowest price over time, and the algorithmic seller (in red) always quickly matches their price. The authors also observe several cases where the seller offering the lowest price is able to sell the product well above their reserve price. As shown in Figure 15, the algorithmic seller always matches the lowest price from the other two sellers.

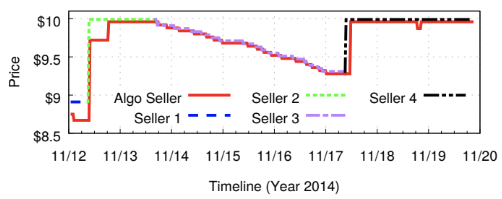

Figure 14: Example of 3P seller (in red) matching the lowest price of all other sellers.

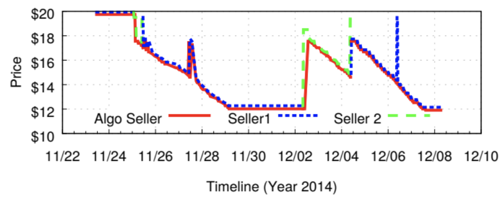

Figure 15: A second example of 3P seller (in red) matching the lowest prices offer by two other sellers.

Although the algorithmic seller is willing to sell the product for as low as $12, the majority of the time they sell at prices up to 40% higher.

2. <u>Amazon is an algorithmic seller</u>

Third, the authors show that Amazon itself appears to be employing algorithmic pricing.

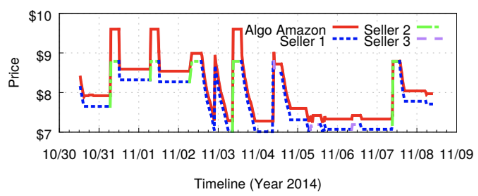

Figure 16: Example of Amazon (in red) setting a premium over the lowest price of all other sellers.

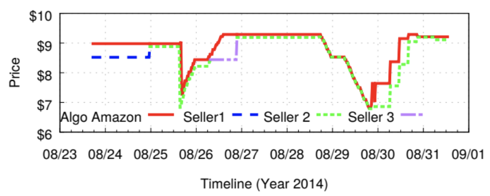

Figure 17: Example of Amazon (in red) matching to the lowest price over time.

In figure 16 we can see that Amazon (in red) is always slightly more expensive than the lowest price, positioning itself like a premium above the lowest price. But according to the authors, Amazon adopts a more complex pricing strategy than just matching lowest prices. In figure 17, Amazon appears positioning itself as a premium above 9$ and following the lowest price below 9$.

Then the authors analyzed the efficiency of the algorithmic strategy by comparing the characteristics of algorithmic and non-algorithmic sellers. They observed that algorithmic sellers exhibit significantly different characteristics than non-algorithmic sellers :

- they sell fewer unique products
- they participate in the marketplace for longer periods of time

- they acquire significantly larger amounts of positive feedback (suggesting they may have higher sales volumes)

Conclusions and Discussion

Based on the results, the authors does many conclusions :

1. Algorithmic pricing is powerful

We have shown that algorithmic sellers appear to be more successful than non-algorithmic sellers. They offer fewer products, but receive significantly higher amounts of feedback (suggesting they have much higher sales volumes). They also “win” the Buy Box more frequently (even when they do not offer the lowest price for a given product).

2. The Buy Box promotes algorithmic pricing adoption

The Buy Box algorithm exacerbates the disparity between algorithmic and non-algorithmic sellers, as it creates a largely winner-take-all marketplace where the Buy Box winner receives the vast majority of sales. We also think that it is challenging for non-algorithmic sellers to compete with algorithmic sellers. This can lead to a situation in which all serious sellers are algorithmic sellers.

Finally, algorithmic pricing can create market distortions according to the authors : “Although we do not observe any of these issues in our data, there are documented cases of algorithms pushing prices to unrealistic heights”. As a conclusion, we promote long-term monitoring of algorithms in markets for more transparency.

3. Discussions

We use the numerous price variations as an indication that sellers align their prices with those of other sellers. However, there are alternative interpretations of these results. It is possible that external reasons influence prices. For example, these variations can be due to a variation in demand. Indeed, other factors such as the time of year (day/night, winter/summer) or the location (rural/urban) affect demand and sellers know it!