Vertical Integration and E-commerce Competition: Evidence from Amazon Marketplace

Luca Bennati *

November 25, 2024

CLICK HERE FOR THE MOST RECENT VERSION

Abstract

Many e-commerce platforms are vertically integrated and compete directly with third-party (3P) sellers. This raises potential competition concerns, as platforms may leverage their market power in ways that harm 3P-sellers and consumers. To analyze the effects of vertical integration on competition and welfare, I build a structural model that captures both the pricing and the product offering decisions of the platform and 3P-sellers. Using data from Amazon, I estimate the model and examine the impact of a policy prohibiting Amazon from selling products on its own marketplace. The counterfactual analysis reveals that while the policy decreases consumer welfare by 19.3%, subsequent entry by other sellers recovers up to 4.5 percentage points of this loss.

JEL Classification: L10, L50, C57

Keywords: Regulation, Digital Platforms, Entry Games

*Toulouse School of Economics, Université Toulouse 1 Capitole. E-mail: luca.bennati@tse-fr.eu

I am grateful to my advisors, Christian Bontemps, Daniel Ershov, and Bruno Jullien, for their support and guidance. I would like to thank Michele Bisceglia, Alexandre De Corniere, Alessandro Gavazza, Olivier De Groote, Isis Durrmeyer, Luise Eisfeld, Oscar Jara, Nicolas Martinez, and Patrick Rey for their useful comments and suggestions, as well as participants at the Compass Lexecon research seminar, ENTER Jamboree (2024), SIM Colloquium (2024), Ibeo Digital Economics Workshop (2024), CRESSE (2024), EARIE (2024), Jornadas de Economía Industrial (2024), and various workshops at Toulouse School of Economics and London School of Economics.

1 Introduction

E-commerce platforms play a crucial role in the retail sector, increasing accessibility and variety of goods and providing an alternative purchase channel to physical stores. In the first quarter of 2024, e-commerce sales in the U.S. accounted for 15.9% of total retail sales.[1] Many of these platforms—including Apple's and Google's App Stores, Walmart, Mercado Libre, and Amazon—are vertically integrated, meaning the platform owners also sell products or services that compete directly with the third-party (3P) sellers they host. This has raised concerns that platforms might exploit their position to the detriment of 3P-sellers and consumers. Practices under scrutiny include self-preferencing (favoring their own offerings over those of 3P-sellers), accessing private data from 3P sellers, and imposing excessive fees.

A possible policy to address this issue is structural separation, i.e., banning platforms from operating in markets where they compete directly with the 3P-sellers. This has been proposed by U.S. Congress[2], and it has been gradually implemented in India starting from 2019.[3] The effect of such a policy is ambiguous ex-ante. On the one hand, consumers have access to fewer products and might face higher prices. On the other hand, 3P-sellers can have an incentive to replace the platform's products with their own, thereby restoring product variety and reducing prices to the level pre-ban.

In this paper, I estimate the impact of structural separation in Amazon. In particular, I focus on quantifying the effect of a 3P-seller replacing Amazon in the products offered before the ban. To analyze this effect, it is crucial to understand which products the 3P-seller decides to offer after the ban. For this purpose, I develop a structural model incorporating entry and pricing. I estimate the model using data from the headphones market in Amazon.com[4] and I focus specifically on entry by sellers operating as retailers, buying products from upstream producers and reselling them on Amazon. In the estimation part, the main challenge is identifying the sellers' fixed costs for offering a product. I address it by using a revealed preference approach that allows me to identify a bound on the fixed costs. My findings show that structural separation is negative for consumers, but neglecting new entry overestimates consumers' loss: while consumer surplus decreases by 19.3%, additional entry by a 3P-seller can recover up to 4.5 percentage points (pp) of it.

[1] U.S. Census.

[2] "Investigation of Competition in Digital Markets", Subcommittee on Antitrust, Commercial and Administrative Law of the Committee on the Judiciary, House of Congress, U.S., 2020.

[3] Press release of the Ministry of Commerce and Industry, India, 2018.

[4] Amazon.com is the marketplace serving the U.S. market. This is the most important market for Amazon in terms of sales (Statista).

I collect data on listed products from Keepa and AmzScout, two market intelligence companies, and I use it to build a dataset of the headphones market on Amazon.com from March 2023 to September 2023. As is typical in e-commerce, there are many products offered and sellers active during this period. Most of the sellers operate as retailers: only 6% of the 3P-sellers sell their own products, while Amazon sells just two Amazon-branded products (both *Amazon Basics*). On average, 3P-sellers three products per week. Instead, Amazon offers around 221 products per week, making it the largest seller in terms of relative sales. Also, I find that Amazon is the only seller of around half of the products it offers; that is, there are no other 3P-sellers offering the same product simultaneously.

Next, I use this sample in order to estimate the structural model. The framework starts from a differentiated product market observed over a number of weeks. Sellers can use two logistics services: Fulfilled by Amazon (FBA), the default option for Amazon, or Fulfilled by Merchant (FBM).[5] 3P-sellers pay referral fees (the percentage of the final price) and FBA fulfillment fee (the fee paid to deliver the product). During each week, sellers choose which products to offer and prices. The fee level and the logistics choice by 3P-sellers are exogenous. Demand during a week is modeled using a random coefficient Logit, where the indirect utility depends on prices, whether the seller is Amazon, the fulfillment method used (Fulfilled by Amazon, FBA, or Fulfilled by Merchant, FBM), product ratings, and number of reviews. 3P-sellers' variable profit includes referral fees (the percentage of the final price paid to Amazon) and the FBA fulfillment fee, while Amazon's variable profit includes the fee revenues. All the sellers incur a fixed cost for every product offered during the week, including inventory and refilling costs. Fixed costs are homogeneous across products but vary on whether the seller is Amazon or a 3P-seller, and depending on the logistics service. Every week, 3P-sellers play a repeated static game with two stages: in the first stage, sellers decide which products to offer; in the second state, they choose prices by competing à la Bertrand.

By assuming demand and supply shocks are unobserved in the first stage and independent from fixed cost shocks, I can estimate the model using a two-step procedure. In the first step, I estimate demand using the methodology in Berry et al. (1995). I use the derived marginal costs to estimate a supply function where marginal cost depend on product, week, and whether the seller is Amazon. Demand and supply parameters are identified under the usual identification restrictions of standard demand/supply models commonly applied in empirical industrial organization (Berry and Haile, 2021). In the second step, I use a revealed preference approach

[5]When using FBA, the product is stored in an Amazon warehouse, and Amazon is in charge of delivery. Amazon also provides the customer service. When using FBM, a 3P-seller remains in charge of storage, delivery, and customer service.

(Kline et al., 2021) that allows me to identify a bound on the fixed cost using moment inequalities. Eizenberg (2014), Kuehn (2018), Wollmann (2018), Canay et al. (2023), and Bontemps et al. (2023) adopt a similar approach in their respective two-stage games.

Demand estimates show that demand is significantly higher for products sold by Amazon. This is the case also for products sold by 3P-sellers using FBA, although to a lower extent than Amazon. This is then reflected in larger market power when Amazon offers the same products simultaneously with other 3P-sellers. However, Amazon may not always be the most efficient seller: on average, Amazon's marginal costs are 5 US$ higher than 3P-sellers when offering the same product. This could be explained by 3P-sellers' ability to better source the product due to specialization in fewer products and markets.

Results on the fixed cost show that 3P-sellers using FBM have lower cost than 3P-sellers using FBA: the fixed cost for the FBM sellers is bounded between 9 US$ and 54 US$, while for FBA sellers is bounded between 11 US$ and 77$. This reflects the opportunity cost of not selling the products offered through FBA on other stores .[6] Hence, while FBA allows FBA to attract more customers, 3P-sellers may also incur higher fixed costs.

I can then use the estimated 3P-sellers fixed costs in order to evaluate the impact of structural separation, which is implemented as a ban on all Amazon's sales, both those realized as a retailer and those realized by selling its own brands. Estimation results suggest the following trade-off. On the one hand, consumers have a higher demand for products when they are sold by Amazon rather than a 3P-seller. Moreover, Amazon is the only seller of many products it offers, so these products are no longer provided after the ban. On the other hand, 3P-sellers face lower marginal costs and have lower market power than Amazon. Hence, lower prices might compensate for lower demand and double marginalization due to Amazon's fees. In addition to this, 3P-sellers might have an incentive to offer many products if they can attract more demand, as this increases profitability by reducing fixed cost per unit.

The policy experiment is carried out as follows. After the ban, an entrant 3P retailer chooses products among those offered by Amazon as a retailer. All incumbent 3P-sellers product offerings are kept fixed. I distinguish three counterfactual scenarios after the ban: in the first one, there is no additional entry; in the second one, the 3P entrant can offer or not the products using FBM; in the third one, the 3P entrant uses FBA.

In the first scenario without entry, consumer surplus falls by 19.3%, as consumers have access to fewer products and have to buy from less valued sellers at higher prices. Price increases

[6] When using FBA, 3P-sellers ship the products to an Amazon warehouse, where it is stored until the sale is realized. Therefore, the 3P-seller is effectively prevented from selling the product in other stores. This issue has also been highlighted in a recent antitrust complaint by the FTC against Amazon.

account for 1.2 pp of the loss in consumer surplus. Moreover, many consumers substitute for 3P-sellers' products. Overall, this results in a profit increase of 19.5%, which could attract more 3P-sellers to the market. Although Amazon gets more fee revenues, its profits are lower compared to pre-ban, suggesting that, at the current fee level, Amazon has an incentive to vertically integrate.

With an FBA entrant, consumer surplus is still lower compared to the pre-ban level, but it improves up to 4.5 pp. This is because the entrant offers up to 84% of the products previously offered by Amazon, restoring product availability and increasing pricing pressure. Results with an FBM entrant and an FBA entrant are quantitatively similar due to price changes: while the FBM entrant charges lower prices than Amazon, the FBA entrant charges higher prices, as it incurs in the FBA fee. Thus, additional entry only partially compensates for the welfare loss, as prices do not decrease enough to compensate for lower demand.

The results suggest that regulators should refrain from imposing structural separation, as platforms' products are highly valued by consumers. Nonetheless, the main benefit of vertical integration does not seem to come from increased product variety, as the same products could be offered by other 3P-sellers too. Consequently, in markets or products where platform and 3P-sellers are more substitutable, structural separation is going to have a small impact on consumers' welfare.

Literature. This paper contributes to the empirical literature on vertical integration in e-commerce platforms in two ways.

Firstly, papers have analyzed how the sellers' pricing in relation to the steering[7] (Tai-Lam, 2023; Lee and Musolff, 2023), platform's data advantage (Chen and Tsai, 2023) and two-sided network externalities (Gutierrez, 2022). Regarding entry, research has generally employed a reduced form approach to analyze the impact of platform's entry on 3P-sellers (Zhu and Liu, 2018; Wen and Zhu, 2019; Crawford et al., 2022). The only paper estimating a structural model is Lee and Musolff (2023), who study the entry and pricing of 3P-sellers assuming independent demand across products in the same market. However, in a differentiated product market, entry and pricing strategies create externalities not only between sellers of the same product but also across products. Therefore, assuming independent demand may not fully capture changes in prices and product differentiation. The estimation of the entry game becomes more challenging, though, because product entry decisions are not independent, and we cannot apply stan-

[7]Other papers have analyzed steering, but without employing a structural approach to analyze how prices change in equilibrium under different recommendation systems (Gomez-Losada and Duch-Brown, 2019; Chen and Tsai, 2021; Hunold et al., 2022; Farronato et al., 2023; Reimers and Waldfogel, 2023; Raval, 2023).

dard techniques to estimate entry games. Hence, I extend this literature by showing how we can estimate product entry with differentiated demand using moment inequalities and analyze the equilibrium effects in the whole market.

Secondly, Gutierrez (2022) and Tai-Lam (2023) show that structural separation is harmful to consumers. I complement this finding by incorporating product entry and showing that additional entry can partially compensate for the negative effect on consumer surplus.

This paper is also related to the theoretical literature studying platforms decisions to operate as a pure intermediary, as pure seller, or as vertically integrated platform (Condorelli et al., 2018; Etro, 2021; Jerath and Zhang, 2010; Jiang et al., 2011; Hagiu and Wright, 2015; Hagiu et al., 2020). Recent literature has suggested different reasons for vertical integration. The marketplace owner may compete with the 3P-sellers for efficiency reasons, to take advantage of its cost efficiencies (Etro, 2021), to reduce double marginalization (Etro, 2021), or to regulate 3P-sellers' prices (Jeon and Rey, 2022). In this paper, I show that, while 3P-sellers have an incentive to replace Amazon, the platform may vertically integrate to take advantage of larger demand and to reduce double-marginalization, but not to exploit cost efficiencies.

Roadmap. In Section 2, I describe the general structure of Amazon marketplace and the data I am using. In Section 3, I provide some descriptive evidence on the headphones market. Then, in Section 4, I present the structural model and, in Section 5, the estimation strategy. In Section 6, I present the estimation results and in section 7 I discuss the effect of structural separation.

2 Setting and Data

2.1 Setting: the Structure of Amazon Marketplace

Amazon marketplace works as an intermediary between consumers on one side and sellers on the other (Figure 1). Amazon operates different geographical marketplaces; for instance, Amazon.com serves mainly the US market, Amazon.mx the Mexican market, Amazon.fr the French market, etc. We can distinguish two types of sellers, retailers and producers. In the former case, we refer to downstream retailers buying products from upstream producers and reselling them on the marketplace (e.g. a 3P-seller or Amazon buys House of Marley headphones and resell them on the marketplace). In the latter case, we refer to upstream producers selling their product directly on the marketplace (e.g., House of Marley selling the 'House of Marley Positive Vibration 2' headphones on the platform). Amazon may operate either as a retailer selling its own products or as a producer selling its own private brands (e.g. Amazon Basics, Amazon Essentials, Solimo, Wag, Mama Bear). While generally, a product could be sold either by a retailer or by its own producer, Amazon's private brands can be sold only by Amazon itself.

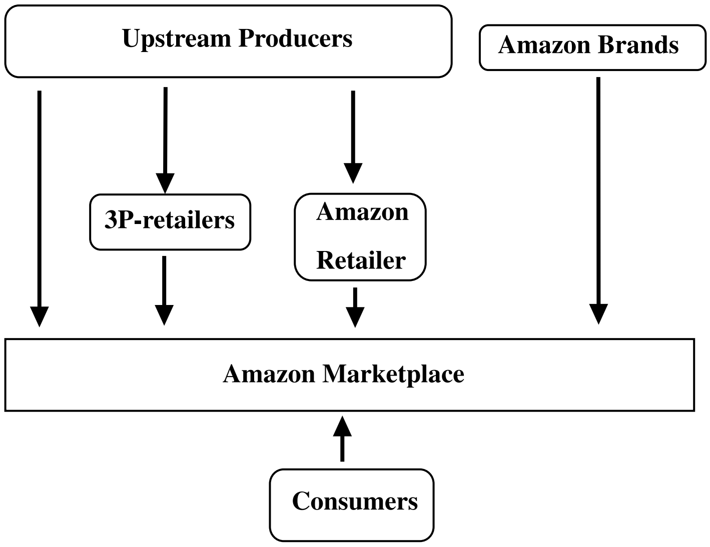


Figure 1: Marketplace Structure

Products can belong to different categories e.g. Consumer Electronics, Electronic Accessories, Home and Kitchen, Office Products. In each category, we can find several markets. For instance, in the Office Products category, a market could be staplers or in the Electronic Accessories headphones. A product is assigned a unique barcode, called ASIN, and each barcode can be sold by multiple sellers at the same time, with the only exception of Amazon's brand barcodes.[8]

There are three features that are central to the functioning of Amazon Marketplace. The first one is logistics, that is, how the delivery of the product is fulfilled. 3P-sellers can choose whether to deliver the product independently or to use FBA. When they use this logistic service, they will send the items to an Amazon warehouse and, upon purchase, Amazon will be in charge of delivering the item from its warehouse to the consumer. Hence, FBA is a way to outsource the storage and delivery of a product and, while Amazon may not be the only provider of this outsourcing service, it is the most commonly used by 3P-sellers. In order to use this service, 3P-sellers pay Amazon an FBA fee, which is made up of two components: FBA fulfillment cost, a cost per unit to fulfill the delivery (packing, shipping, customer service etc.), and FBA storage cost, an inventory cost charged monthly to store a product in Amazon's warehouse. During the sample period, for the Small standard category (median side up to 30 cm), the lowest FBA fulfillment cost was for products with shipping weight lower or equal to 113 gr (4 oz) and corresponded to \$ 3.22, while the highest FBA fulfillment cost was for products with shipping weight between 340 gr (12 oz) and 454 gr (16 oz) and corresponded to \$3.77. Given that there

[8]In Appendix B, I provide some examples of products and sellers.

is little difference in FBA fulfillment cost between the different weights, I will assume all FBA offers pay the lowest FBA fulfillment cost. When it comes to Amazon's logistic choice, it will use its own service, FBA, by default, and it will not pay the FBA fees. The second feature is the referral fees. These are an ad-valorem fee, and they amount to a percentage of the final price paid by consumers. These fees are typically 8% or 15% depending on the product category, although there are a few cases of non-linear fees. In this case, the fee may vary depending on the price (e.g. in Grocery and Gourmet in Amazon.com, the referral fee is 8% for a price below or equal to 15$, and 15% otherwise) or on the portion of the price (e.g. for Electronics Accessories in Amazon.com, the fee is 15% for the portion of the total sales price up to $100.00, and 8% for any portion of the total sales price greater than $100.00. Headphones are included in this category).

The last feature is the recommendation system, which is made up by three main components. The first is the page ranking. When costumers looks for an item, they will type a keyword in the search bar and a list of products will appear, the order depending by the relevance to the keyword, the product rating, the number of reviews and the available offers characteristics (prices, sellers' rating, product). The second is the Buy-Box. As mentioned before, multiple sellers may offer the same barcode simultaneously (Figure **??**) and there is not a limit on how many sellers can offer the same barcode. When this happens, all sellers' offers will be grouped in the same product page, but only one of them will be given more visibility to consumers. This offer will be place in a window on the top of the product page called the Buy-Box and its seller will be called the Buy-Box seller. Moreover, the product's price costumers observe in the page ranking will refer to the one of the Buy-Box offer. Finally, other products may be recommended in a product page under a page called Frequently Bought Together.

2.2 Data

I collect data from two sources. The first one is Keepa, a website scraping Amazon and providing information on product and offers characteristics of all currently listed products. Data from Keepa has also been recently used in Cabral and Xu (2021), Lee and Musolff (2023), Gutierrez (2022) and Chen and Tsai (2023). Product characteristics include title, brand, manufacturer, product description, in addition to real time changes in sales rank (a measure of aggregate product sales relative to other products in the same category), product rating, number of reviews and Buy-Box seller. Offer characteristics include seller's name, logistic method and real time changes in prices and shipping costs.

I complement the data from Keepa with data from AmzScout, a market intelligence company used by 3P-sellers. First, AmzScout sales estimator allows to estimate the aggregate quantity sold for a product in a given period from the sales rank. I provide more information about this in Appendix A.1.1. Secondly, I use data from number of keyword searches for headphones to compute the potential market (Appendix A.1.2). This information will also be useful to extract some relevant product characteristics.

Moreover, while I can estimate the aggregate quantity sold using data from Keepa and AmzScout, I still do not observe the quantity sold by each single seller. However, it is generally reported that most of the sales go through the Buy-Box seller and therefore, as in Gutierrez (2022), I will assumed that observed sales are realized only by the Buy-Box seller [9]. Finally, while in this paper I focus on the headphones market, I have collected data on other markets as well that I plan to use in future extensions.

3 Descriptive Evidence

In this empirical analysis, I am going to consider the headphones market in Amazon.com between March 2023 and September 2023. I focus on wireless Bluetooth headphones, whose demand can be assumed to be independent of wired headphones.[10] The choice of the headphones market is driven by the fact that this is a market where small-scale innovation (such as introducing headphones with a particular design or certain features), and therefore 3P-sellers' entry, is particularly important.

I describe the main features of the market here. In particular, I focus on the sample of offers with positive sales, which is the sample used to estimate demand and fixed costs and carry out the counterfactuals. In Table 1, I report the main statistics to describe the structure of the market. As is typical in digital markets, there are many products offered (where by-product, I refer to a distinct barcode) and many active sellers. Not all of these products are sold during every period, as the average number of products sold is around 68% of the total. Similarly, for the sellers, 38% of the total number of sellers make a positive sale during a week. The majority of these sellers operate as retailers, buying from upstream producers and reselling them on the platform on the platform. This is the case for Amazon, too, which sells just two Amazon-branded products (both Amazon Basics) during the sample period. 3P-sellers can choose between Fulfilled by Amazon (FBA) or fulfilling the sale independently (FBM). I observe that around 80% of the

[9] "Industry experts estimate that about 80% of Amazon sales go through the Buy Box, and the percentage is even higher for mobile purchases" ("Investigation of Competition in Digital Markets" Subcommittee on Antitrust, Commercial and Administrative Law, House of Representatives, 2020)

[10] I choose this type of headphones as they will likely be the most popular and recent products.

3P-sellers use FBA.[11]. Yet, in around 8% of the offers, the 3P-seller may change the fulfillment service for the same product across weeks, implying that some 3P-sellers change the fulfillment service or use multiple fulfillment services during the sample periods. Then, on average, every product is sold by almost one seller every week.

Regarding the number of products offered by a seller during a week, we can notice a significant difference between Amazon and 3P-sellers: while 3P-sellers sell around three products every week, Amazon sells almost 221 products, making Amazon the largest seller. Even when we look at the top sellers in terms of the number of products sold, Amazon distances itself significantly from the other 3P sellers. In Table 2, I report the top 10 sellers in terms of the average number of products sold weekly. The second largest seller, Sampell, sells less than half of the products every week, while the distance is even larger from the third largest seller onward. Looking at all the sellers, we can see that the distribution is skewed, with many sellers offering one product (Figure 2a). Moreover, I find little concentration of sellers across products, with most of the products offered by only one seller (Figure 2b). Finally, in around 58.9% of the products sold by Amazon during a week, Amazon is the only seller having positive sales. Moreover, Amazon does not compete with other 3P-sellers in many products it sells: in 49.3% of these products, Amazon is also the only seller offering the product.

Looking at the sales distribution (Figure 3), we can observe the long tail typical of e-commerce, where the majority of products receive a small proportion of the sales and a small fraction of products receive many sales. Within the distribution, Amazon is the seller obtaining more sales: each of the products sells on average 50 units per week, while its Amazon brands sell around 24 units per week, contrary to 3P-sellers products selling only 15 products per week.

However, differences in prices and quantities between Amazon and 3P-sellers may also mask a degree of heterogeneity across products offered. For instance, 3P-sellers may concentrate on lower quality products in the long tail, while Amazon in high-demand and high-quality products. In order to better disentangle the two effects, I compute the difference between 3P-sellers' prices and quantities and Amazon's when they offer the same product in the same week. Figure 7a reports the difference between a 3P-seller price and Amazon price, $\Delta\text{Price} = \text{Price}_{3P} - \text{Price}_{Amazon}$, and 7a reports the difference between a 3P-seller quantity sold and Amazon's quantity sold, $\Delta Q = Q_{3P} - Q_{Amazon}$. 3P-sellers may charge higher or lower prices than Amazon, but for most of the products, the difference is not too large (the median is 0 US$). Instead, regarding quantities, 3P-sellers sell almost the same quantity or less than

[11] Although very rare, a 3P-seller can offer the same product using two separate fulfillment services. In this case, two separate offers appear on the product page, competing for the Buy-Box independently. When this is occurs, I keep only the offer using FBA.

Amazon, while they never have much larger quantities (the median is -6).

To sum up, descriptive evidence shows that Amazon plays an important role in the headphones market, and it does so mainly as a retailer. Amazon is the seller offering the largest product variety, and it does not compete with other third-party retailers in many products it sells. Amazon also charges larger prices and obtains more sales than 3P-sellers, suggesting a certain degree of product heterogeneity: 3P-sellers may concentrate on products in the long tail, while Amazon tends to offer higher demand and high-quality products. However, when selling the same product simultaneously, Amazon often sells more quantities than 3P-sellers, even when 3P-sellers do not charge higher prices. This provides evidence of Amazon's competitive advantage vis-a-vis 3P-sellers.

The goal is then to investigate the possible sources of Amazon's advantage and to quantify the level of competition across products in the market. Since I only observe aggregate market shares, I need to take a structural approach. I estimate a Logit demand model with random coefficient, which allows me to find own- and cross-price elasticities across products. I then use the supply model to compare 3P-sellers' and Amazon's markups and market power. Finally, I use the estimated demand and supply model to compute sellers' fixed costs of offering a product in the market.

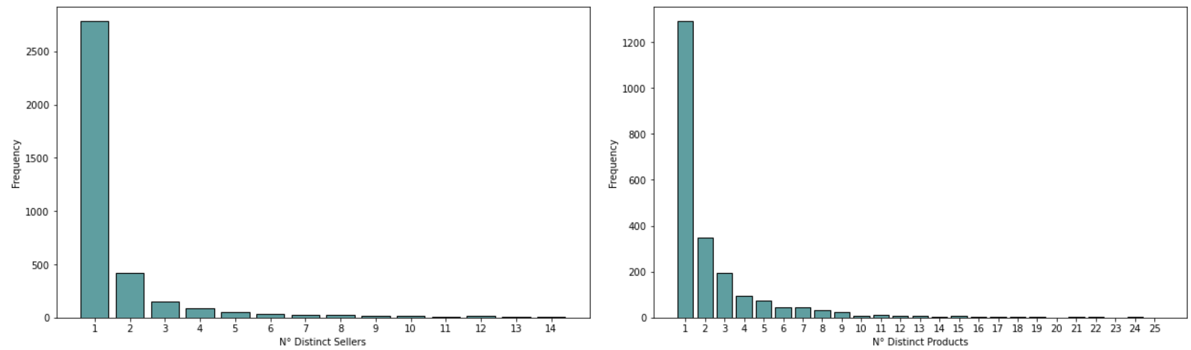
Total	
n° Products	3688
n° Sellers	2253
% Producers	5.73
n° Amazon Brands	2
% 3P-sellers FBA Offers	78.79
% Offers using both FBA/FBM	8.30

Weekly Mean	
n° Products	2345.07
n° Sellers	849
n° Sellers per Product	1.20
n° Products by a 3P-seller	3.05
n° Products by Amazon	220.77
Price Amazon	156.20 US$
Price 3P-seller	80.90 US$
Quantity Sold Amazon Retailer	49.45
Quantity Sold Amazon Brand	24.32
Quantity Sold Amazon Brand	14.55
Herfindahl-Hirschman Index	665

Table 1: **Summary Statistics**

Seller Name	Type Of Seller	Weekly N° Products
Amazon.com	Retailer/Producer	220.77
Sampell	Retailer	95.70
Zihnic	Producer	34.80
Thousandshores Inc	Retailer	30.30
Woot	Retailer	25.20
HEYE	Retailer	24.87
Avantree Direct USA	Producer	24.47
JLab Store	Producer	23.47
Trusted Kids Products	Retailer	23.33
Anker Direct	Producer	23

Table 2: **Largest Sellers**



(a) Number of Distinct Sellers per Product (b) Number of Distinct Products per Seller

Figure 2: **Level of Concentration Across Products and Sellers.** Figure (a): number of distinct sellers ever offering the product across all periods. Figure (b): number of distinct products offered by a seller in the sample period across all periods. Truncated distribution below the 99th percentile.

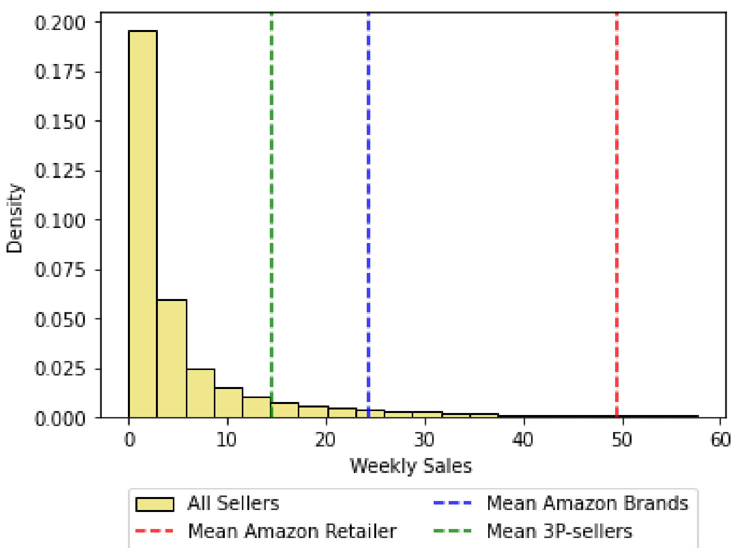
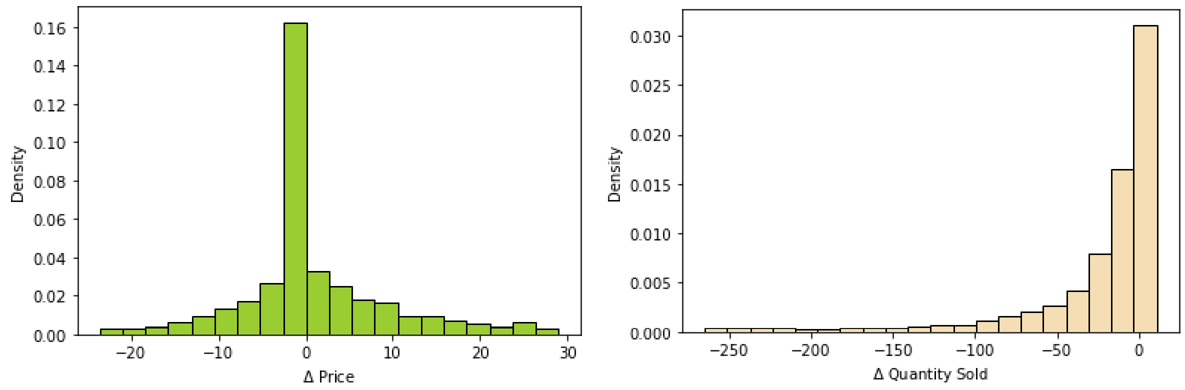


Figure 3: **Weekly sales.** Truncated distribution below 95th percentile.



(a) Price (b) Quantity

Figure 4: **Difference 3P-sellers' and Amazon's prices and quantity sold**. Figure (a): difference between 3P-seller's and Amazon's price when Amazon offers the same product in the same week. Figure (b): difference between 3P-seller quantity sold and Amazon's quantity sold when Amazon offers the same product in the same week. Truncated distribution above the 5th percentile and below the 95th percentile.

4 Model

4.1 Setup

Consider a differentiated product market observed over multiple weeks. Products are indexed by $j = 1, \ldots, J$, and each product j is identified by a unique barcode. Weeks are indexed by $t = 1, \ldots, T$ and months by $m = 1, \ldots, M$. Sellers are indexed by $d = 1, \ldots, D$ week. There are two types of sellers: 3P-sellers, indexed by $r = 1, \ldots, R$, and Amazon, indexed by a.

3P sellers pay a referral fee ϕ_{jrt}, which is dependent on the price they charge for each product, while if they use Fulfillment by Amazon (FBA), they also pay a constant unit fee τ that applies across all products, weeks, and 3P-sellers.

A 3P-seller r can offer product j through either FBA or Fulfillment by Merchant (FBM). The set of products sold by 3P-seller r using FBM is represented as $\mathcal{J}_{rt}$, while the set of products sold using FBA is represented as $\widetilde{\mathcal{J}}_{rt}$. Amazon always offers a set products $\mathcal{J}_{at}$ through FBA.

During each week t, each seller d makes two decisions: which set of products to offer; their prices, denoted by $\boldsymbol{p}_{rt} = \{p_{jrt} : j = 1, \ldots, J_{rt}\}$, $\widetilde{\boldsymbol{p}}_{rt} = \{\widetilde{p}_{jrt} : j = 1, \ldots, \widetilde{J}_{rt}\}$ and $\boldsymbol{p}_{at} = \{p_{jat} : j = 1, \ldots, J_{at}\}$. I denote by $\boldsymbol{\mathcal{J}}_{rt} = \{\mathcal{J}_{rt}, \widetilde{\mathcal{J}}_{rt}\}$ seller r product offered through FBM and FBA in week t.

Finally, I denote by $\boldsymbol{p}_t = \{p_{at}, p_{rt}, \widetilde{p}_{rt} \mid r = 1, \ldots, R\}$ the vector of all sellers' prices in week t and by $\boldsymbol{\mathcal{J}}_t = \{\mathcal{J}_{at}, \boldsymbol{\mathcal{J}}_{rt} | r = 1, \ldots, R\}$ the product assortment in week t.

4.2 Demand

I model demand as a logit with a random coefficient on prices.

Every week, $\mathcal{M}_t$ consumers enter the market on Amazon marketplace and decide whether to purchase one of the offered products.

The indirect utility of consumer i for product j sold by seller d at time t is:

$$U_{ijdt} = \boldsymbol{\beta}\boldsymbol{X}'_{jdt} - \alpha_i p_{jdt} + \gamma_j + \gamma_m + \xi_{jdt} + \varepsilon_{ijdt} \tag{1}$$

This is a function of $\boldsymbol{X}_{jdt}$, the products and sellers characteristics, p_{jdt}, the price, γ_m, the month fixed effect, γ_j, the product fixed effect, ξ_{jft}, the product-seller random shock, ε_{ijdt}, the consumer specific random shock distributed Type 1 EV. I introduce one random coefficient $\alpha_i = \alpha + \sigma\mu_i$, where $\mu_i \sim N(0,1)$.

The utility of the outside option is

$$U_{iot} = \varepsilon_{iot} \tag{2}$$

where the outside option includes both the choice of not buying any product at all or buying from another physical or digital store. Product and seller characteristics include average product rating, number of reviews for the product, whether the product is sold by Amazon, and whether the seller is using FBA (this is always the case for Amazon). I do not include whether the offer is assigned the Prime badge since this is the case for almost all the FBA offers.

Average product rating and number of reviews can be viewed as a proxy for product quality and are considered with a one-week lag to avoid potential endogeneity. Given the distribution of ε_{ijdt}, the market share predicted by the demand model is:

$$s_{jdt}(\boldsymbol{p}_t, \boldsymbol{X}_t) = \int \frac{\exp(\delta_{jdt} + \sigma\mu_i)}{1 + \sum_{f=1}^{F} \sum_{j \in \mathcal{J}_f} \exp(\delta_{jdt} + \sigma\mu_i)} dF(\mu) \tag{3}$$

where $\boldsymbol{p}_t$ is the vector of sellers' prices, $\boldsymbol{X}_t$ is the vector to products and sellers characteristics, and δ_{jdt} is the mean utility of consumers

$$\delta_{jdt} = \boldsymbol{\beta}\boldsymbol{X}'_{jdt} - \alpha p_{jdt} + \gamma_j + \gamma_m + \xi_{jdt} \tag{4}$$

I denote by $s_{jrt}(\boldsymbol{p}_t, \boldsymbol{X}_t)$ the market share of product j sold by 3P-seller r in week t using FBM and by $\widetilde{s}_{jrt}(\boldsymbol{p}_t, \boldsymbol{X}_t)$ the market share when the 3P-seller is using FBA.

4.3 Variable Profit

Every week, Amazon and 3P-sellers earn a variable profit from product sales.

A 3P-seller earns a variable profit in a week t, which is the sum of profits from products sold through FBM and those sold through FBA, as shown in Eq. (5). The seller's earnings are reduced by the referral fee ϕ_{jrt} paid to Amazon, in addition to the cost τ associated with products sold using FBA.

Marginal costs are represented as c_{jrt} when a seller uses FBM to fulfill product j in week t, and as $\widetilde{c}_{jrt}$ when the seller uses FBA. This means that the seller can encounter different marginal costs for selling the same product based on the fulfillment service utilized. The interpretation of these marginal costs varies depending on whether the 3P-seller acts as a downstream retailer or operates as a vertically integrated seller, producing products and selling them directly in the marketplace. For a 3P-seller functioning as a retailer, the marginal cost c_{jrt} includes both fulfillment costs and the wholesale price, while $\widetilde{c}_{jrt}$ represents only the wholesale price. Conversely, if the 3P-seller is vertically integrated, c_{jrt} encompasses both fulfillment and production costs, whereas $\widetilde{c}_{jrt}$ accounts for only the production cost.

Amazon's variable profit consists of both profit from product sales and fee revenues (Eq. (6)). The company collects a percentage of the price, represented by $\phi_{jrt} \cdot p_{jrt}$, from all third-party sellers (3P-sellers). Additionally, Amazon charges an FBA fee, τ, for all products sold by 3P-sellers using Fulfillment by Amazon (FBA).

Amazon marginal cost, c_{jat}, includes the wholesale price if Amazon is reselling product j, while it includes the production cost if Amazon is producing product j. In both cases, the marginal cost includes also the cost of fulfilling the product.[12]

$$\begin{aligned}\pi_{rt}(\boldsymbol{\mathcal{J}}_t, \boldsymbol{X}_t) = \mathcal{M}_t \cdot \Bigg(& \sum_{j \in \mathcal{J}_{rt}} s_{jrt}(\boldsymbol{p}_t, \boldsymbol{X}_t) \cdot \Big(\big(1 - \phi_{jrt}\big) \cdot p_{jrt} - c_{jrt} \Big) \\ & + \sum_{j \in \widetilde{\mathcal{J}}_{rt}} \widetilde{s}_{jrt}(\boldsymbol{p}_t, \boldsymbol{X}_t) \cdot \Big(\big(1 - \widetilde{\phi}_{jrt}\big) \cdot \widetilde{p}_{jrt} - \widetilde{c}_{jrt} - \tau \Big) \Bigg) \quad \forall\, r = 1, \dots, R \end{aligned} \tag{5}$$

$$\begin{aligned}\pi_{at}(\boldsymbol{\mathcal{J}}_t, \boldsymbol{X}_t) = \mathcal{M}_t \cdot \Bigg(& \sum_{j \in \mathcal{J}_{At}} s_{jAt}(\boldsymbol{p}_t, \boldsymbol{X}_t) \cdot (p_{jat} - c_{jat}) \\ & + \sum_r \Big(\sum_{j \in \mathcal{J}_{rt}} \phi_{jrt} \cdot p_{jrt} \cdot s_{jrt}(\boldsymbol{p}_t, \boldsymbol{X}_t) + \sum_{j \in \widetilde{\mathcal{J}}_{rt}} \widetilde{\phi}_{jrt} \cdot \widetilde{p}_{jrt} \cdot \widetilde{s}_{jrt}(\boldsymbol{p}_t, \boldsymbol{X}_t) \Big) \\ & + \sum_r \sum_{j \in \widetilde{\mathcal{J}}_{rt\{FBA\}}} \tau \cdot \widetilde{s}_{jrt}(\boldsymbol{p}_t, \boldsymbol{X}_t) \Bigg) \end{aligned} \tag{6}$$

[12] One variable not included in Amazon's variable profit is the cost of fulfilling the products sold by 3P-sellers using FBA. Since the marginal cost of fulfilling 3P-sellers products, c_L can be between 0 and τ, I can test the implication of this assumption by comparing c_{jta} when $c_L = 0$ and $c_L = \tau$. I describe the results in Appendix A.4.

4.4 Marginal Costs

Marginal costs depend on whether the 3P-seller is Amazon, on product j, week t and an random shock ω_{jft}

$$c_{jdt} = \bar{c}_{jdt} + \omega_{jdt} \tag{7}$$

$$\bar{c}_{jdt} = \lambda \mathbb{1}_{d=a} + \zeta_j + \zeta_t \tag{8}$$

For retailers, $\bar{c}_{jdt}$ represents the wholesale price, while ω_{jdt} changes in shipping cost. When a product is sold by multiple retailers in the same week, sellers' heterogeneity is reflected in whether the seller is Amazon and on the random shock ω_{jdt}. Since Amazon is more likely to be supplied directly by upstream producers and is the largest seller in the market, the most significant difference in wholesale prices exists between Amazon and 3P-sellers. Although there may be greater variability among the 3P-sellers, my estimation results demonstrate that this model can predict prices with a high degree of accuracy.[13]

4.5 Fixed Costs

During each week t, sellers incur a fixed cost f_{jdt} to offer product j. This fixed cost includes an average value, θ_d, which varies depending on whether the seller is Amazon and the fulfillment method used by the 3P-seller. Additionally, there is a random shock, V_{jdt}, that is specific to each product, seller, and week.

$$f_{jdt} = \begin{cases} f_{jrt} & = \theta + V_{jrt} \quad \textit{if } d = r \textit{ and } j \textit{ sold using FBM} \\ \\ \widetilde{f}_{jrt} & = \widetilde{\theta} + \widetilde{V}_{jrt} \quad \textit{if } d = r \textit{ and } j \textit{ sold using FBA} \\ \\ f_{jat} & = \theta_a + V_{jat} \quad \textit{if } d = a \end{cases} \tag{9}$$

The fixed cost for a product sold using FBM, θ, includes storage costs and the cost of refilling the stock of the product. The storage cost is related to the shelf space used for product storage in a warehouse; it reflects the opportunity cost of not storing other products. The cost of refilling the stock represents the 3P-sellers' cost of monitoring inventory levels and forecasting demand for the product, and, for 3P-sellers acting as retailers, the cost of placing a refilling order with the supplier. When the 3P-seller uses FBA, the fixed cost, $\widetilde{\theta}$, includes the opportunity cost of selling through other channels. In fact, since the product is stored in an Amazon warehouse, in

[13] Another source of heterogeneity could be found between 3P-sellers using FBM and 3P-sellers using FBA. In Appendix A.3, I provide an extension of the marginal cost function to account for this case.

practice, it can only be sold through Amazon (otherwise, the 3P-seller would need to withdraw the product and ship it by itself, which could be a lengthy process).

For Amazon, the fixed cost, θ_a, represents inventory costs and the opportunity cost of selling the product, which includes forgone revenues from 3P-sellers. Since introducing a product increases competition with 3P-sellers, reducing 3P-sellers' prices and diverting demand from their products, fee revenues decrease too.

The only source of heterogeneity comes from the random shock V_{jdt}, which represents variations in supplying a product for a specific seller, for instance, disruptions in the supply chain of a certain product preventing the seller from refilling the stock, or, for 3P-sellers, changes in the profitability of selling on other platforms.

The total fixed cost of offering a set of products is the linear sum of each fixed cost. Therefore, the total fixed cost for a 3P-seller r for offering a set of products $\mathcal{J}_{rt}$ through FBM and a set of products $\widetilde{\mathcal{J}}_{rt}$ through FBA is:

$$\dot{F}_{dt} = \begin{cases} F_{rt} + \widetilde{F}_{rt} = \sum_{j \in \mathcal{J}_{rt}} (\theta + V_{jrt}) + \sum_{j \in \widetilde{\mathcal{J}}_{rt}} (\widetilde{\theta} + \widetilde{V}_{jrt}) & \text{if } d = r \\ \sum_{j \in \mathcal{J}_{At}} (\theta_a + V_{jat}) & \text{if } d = a \end{cases} \tag{10}$$

The linearity assumption is a simplification for tractability. While it holds for certain parts of the fixed cost (e.g., storage cost), other components might be characterized by economies of scale (e.g., the more products the 3P-seller offers, the more likely it is to use an inventory management software, which reduces the refilling cost per product). In the conclusions, I discuss how the model can be extended to account for this.

4.6 Timing of the Game

At the outset of the game, Amazon sets the referral and FBA fees; after observing these choices, sellers pay the entry cost to operate in the headphones market on the Amazon marketplace. The level of platform fees and the set of active sellers is exogenously given. Moreover, I assume that for every potential product a 3P-seller can offer, it exogenously decides whether to use FBM or FBA to fulfill the sale.

Then, the game unravels as follows. During each week, sellers play a static (two-stage) entry and pricing game to maximize total profit.[14] The timing of the game is the following:

- **Stage 1**: Sellers learn the fixed cost shocks, $\boldsymbol{V}_t = \{V_{jdt} : j = 1, \ldots, J, d = 1, \ldots, D\}$. At this stage, sellers do not know about the demand shocks, $\boldsymbol{\xi}_t = \{\xi_{jdt} : j = 1, \ldots, J, d =$

[14]The static nature of the game is justified by the sellers having already paid the entry cost of operating in the market. Once they are present in the market, they have to decide which products in each week.

$1, \ldots, D\}$, and the marginal cost shocks, $\boldsymbol{\omega}_t = \{\omega_{jdt} : j = 1, \ldots, J, d = 1, \ldots, D\}$. However, they have information about the distribution of the shocks. Therefore, they choose which products to offer to maximize the expected total profits[15]:

$$\max_{\{\boldsymbol{\mathcal{J}}_{rt}, \widetilde{\boldsymbol{\mathcal{J}}}_{rt}\}} E_{\xi,\omega}[\Pi_{rt}] = E_{\xi,\omega}[\pi_{rt}(\boldsymbol{\mathcal{J}}_t, \boldsymbol{X}_t)] - \dot{F}_{rt} \quad \forall\, r = 1, \ldots, R \tag{11}$$

$$\max_{\boldsymbol{\mathcal{J}}_{at}} E_{\xi,\omega}[\Pi_{at}] = E_{\xi,\omega}[\pi_{at}(\boldsymbol{\mathcal{J}}_t, \boldsymbol{X}_t)] - \dot{F}_{at} \tag{12}$$

- **Stage 2**: Sellers learn the demand shocks, $\boldsymbol{\xi}_t$, and marginal cost shocks, $\boldsymbol{\omega}_t$, and compete à la Bertrand to maximize the expected variable profit:

$$\max_{\boldsymbol{p}_{rt}} \pi_{rt}(\boldsymbol{\mathcal{J}}_t, \boldsymbol{X}_t) \quad \forall\, r = 1, \ldots, R \tag{13}$$

$$\max_{\boldsymbol{p}_{at}} \pi_{at}(\boldsymbol{\mathcal{J}}_t, \boldsymbol{X}_t) \tag{14}$$

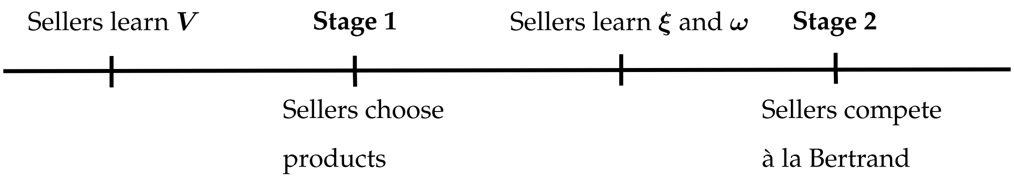


Figure 5: **Timing of the Game**

At Stage 1, sellers know the fixed cost shocks but not the demand and marginal cost shocks. Since consumers' utility and the marginal cost function include product and time fixed effects, sellers are likely to be unaware of these residual shocks before committing to a certain product offering.

Finally, the model is solved by backward induction to have a Bayesian Nash Equilibrium. While I assume an equilibrium exists, I do not assume its uniqueness. The multiplicity of equilibria stems from the fact that, given the primitives of the game, different product assortments could constitute an equilibrium of the game.

[15] Amazon's total profit does not account for the costs associated with fulfilling products sold by 3P-sellers using FBA. This is because the choice to use FBA is made solely by the 3P-sellers, and Amazon can only influence this decision through the fees associated with FBA. Since I am not including an extra stage where Amazon sets the fees, in the current model, Amazon's entry and pricing decisions are not affected by the cost of running the FBA program. Therefore, adding it would not change the results.

5 Identification and Estimation

5.1 Estimation Strategy

Note that, since sellers do not observe $\boldsymbol{e}_t = \{\boldsymbol{\xi}_t, \boldsymbol{\omega}_t\}$ before committing to a particular product offering, product choices are independent of the realizations of the error terms. Consequently, I can separately address the identification and estimation of parameters for each stage of the model.

First, for the second stage, consider the demand and marginal cost parameters collected in $\boldsymbol{\psi} = \{\boldsymbol{\beta}, \alpha, \sigma, \boldsymbol{\gamma}, \lambda, \boldsymbol{\zeta}\}$, where $\boldsymbol{\gamma} = \{\gamma_j, j = 1, \ldots, J\} \cup \{\gamma_m, m = 1, \ldots, M\}$ and $\boldsymbol{\zeta} = \{\zeta_j, j = 1, \ldots, J\} \cup \{\zeta_t, t = 1, \ldots, T\}$. $\boldsymbol{\psi}$ is point-identified under the usual identification restrictions of standard demand/supply models commonly applied in empirical IO (Berry and Haile, 2021).

For the fixed cost parameters $\boldsymbol{\theta} = \{\theta, \widetilde{\theta}, \theta_a\}$, I show that they are partially identified. First, due to the large number of products and sellers, listing all the options available to sellers and calculating their realization probabilities is virtually impossible. Second, it is well known that games of this type exhibit multiple equilibria, which generally prevents applied economists from using standard estimation techniques without additional assumptions (see de Paula in the Annual Review of Economics). For this reason, I employ a revealed preference approach (Kline et al., 2021) that allows me to bound $\boldsymbol{\theta}$ using moment inequalities. Eizenberg (2014), Wollmann (2018), Canay et al. (2023), and Bontemps et al. (2023) adopt a similar approach in their respective two-stage games.[16]

In Section 5.2 and Section 5.3, I describe in detail the estimation of demand, marginal costs, and fixed cost parameters.

5.2 Demand and Marginal Costs

Demand parameters are identified from the distribution of sales, prices, and products and sellers' characteristics. Marginal cost parameters are identified by the comovement between the identified marginal cost and products and sellers' characteristics.

As mentioned above, a selection bias may arise in the identification of $\boldsymbol{\psi}$ since the assortment $\mathcal{J}_t$ is the result of sellers' product choices. Therefore, we do not observe a random sample from the underlying distribution of products and sellers' characteristics. However, given the timing

[16] Other papers using moment inequalities to identify fixed costs in product variety games are Fan and Yang (2022) and Martinez (2023). They have some differences in terms of model and estimation. Fan and Yang (2022) has a similar model, but does not take a revealed preference approach to build the moments and rather use bounds on the entry probabilities. Martinez (2023) uses revealed preferences to construct moment inequalities in a dynamic game of spatial competition without a pricing stage.

of the game, $\boldsymbol{e}_t$ is unobserved when sellers choose products, and selection based on unobservables can be ignored. It is then possible to show that parameters in $\boldsymbol{\psi}$ are point identified.

Let $\mathcal{H}$ denote all potential product assortment $\boldsymbol{\mathcal{J}}$, where $\{jd\}$ denote a particular offer in a potential assortment and $\mathcal{X}$ the matrix of products and sellers' characteristics for the potential assortments. I define by $q_{jdt}(\mathcal{X}, \dot{\boldsymbol{F}}_t) = \{1, 0\}$ the decision of seller d to offer product j in week t. First, I assume that $e_{jdt} = \{\omega_{jdt}, \xi_{jdt}\}$ is independent of product and sellers' characteristics and on fixed costs $\dot{\boldsymbol{F}}_t$:

Assumption 1. $\mathbb{E}[e_{jdt}|\mathcal{X}, \dot{\boldsymbol{F}}_t] = 0 \quad \forall \{jd\} \in \mathcal{H}$

Let $z_{jdt}(\mathcal{X})$ be a vector of instrument for offer $\{jd\}$. Then, since $\boldsymbol{e}_t$ is unknown by sellers when choosing products, by using the Law of Iterated Expectations and by Ass. 1, we obtain that

$$\mathbb{E}[q_{jdt} e_{jdt} z_{jdt}] = 0 \tag{15}$$

Therefore, when product j is sold by seller d, we obtain

$$\mathbb{E}[e_{jdt} z_{jdt} | q_{jdt} = 1] = 0 \tag{16}$$

This is similar to the standard exclusion to estimate the demand and marginal cost parameters, with the addition of being conditional $q_{jdt} = 1$. Therefore, using the observed product assortments, $\boldsymbol{\psi}$ is point identified.

I use the moment conditions in Eq (16) to estimate the demand parameters $\{\boldsymbol{\beta}, \alpha, \sigma, \boldsymbol{\gamma}\}$ by GMM using the fixed point algorithm from Berry et al. (1995) (henceforth BLP).

The only endogenous variable in demand is price, which I instrument using product characteristics as (BLP) instruments. I build two groups of instruments. The first one is the average of other offers characteristics. Product characteristics describe different features of the product (e.g., whether the headphone is noise-canceling, designed for gaming or sports activities). Consider a product feature X_j^c where $X_j^c = 1$ if product j has this feature, and zero otherwise. Then, given N_t the cardinality of $\boldsymbol{\mathcal{J}}_t$ the instrument for offer $\{jd\}$ is

$$z_{jdt}^c = \frac{1}{N_t - 1} \sum_{k \neq \{jd\}} X_k^c \tag{17}$$

I consider two product characteristics, *sport* and *gaming*. Since sports and gaming are characteristics positively valued by consumers, the average number of sports headphones can be interpreted as a measure of how crowded the market is, so the higher the average, the higher the competitive pressure on sellers' prices.

The third instrument is the distance to the mean of other offers' characteristics:

$$z_{jdt}^c = |\frac{1}{N_t - 1} \sum_{k \neq \{jd\}} (X_{jd}^c - X_k^c)| \tag{18}$$

The argument is similar: the more similar the competitor's headphones, the larger the competitive pressure. In this case, the characteristic that I am considering is noise-canceling.

Once estimated the demand parameters, we can identify the marginal cost using the First Order Conditions from the supply models in Eq. (5) and Eq. (6). I describe the procedure in Appendix A.2. Finally, I use the identified marginal costs to estimate the marginal cost parameters in Eq. (8).

5.3 Fixed Cost

As mentioned before, characterizing the probability distribution of all potential equilibria is untractable. However, by revealed preferences, we know that any deviation from the observed equilibrium would be unprofitable. Take seller d offering a set of products $\mathcal{J}_{dt}$ in week t and let $\boldsymbol{\mathcal{J}}_{-dt}$ be the vector of sets of products offered by all its competitors. Then, keeping $\boldsymbol{\mathcal{J}}_{-rt}$ fixed, adding or removing a product from $\mathcal{J}_{dt}$ would be unprofitable for seller d.

Let π be the expected variable profit and consider the case of a 3P-seller selling product j through FBM. If $j \notin \mathcal{J}_{rt}$, then the change in profit from selling j, $\Delta\pi^{+}_{jrt}$, is lower then the product fixed cost

$$\pi(\mathcal{J}_{rt} + \{j\}, \boldsymbol{\mathcal{J}}_{-rt}, \boldsymbol{X}) - \pi(\mathcal{J}_{rt}, \boldsymbol{\mathcal{J}}_{-rt}, \boldsymbol{X}) \leq \theta + V_{jrt} \quad \text{if } q_{jrt} = 0 \tag{19}$$

If $j \in \mathcal{J}_{dt}$, then the difference in profit between selling and not selling j, denoted $\Delta\pi^{-}_{jrt}$, is larger then the product fixed cost

$$\pi(\mathcal{J}_{rt}, \boldsymbol{\mathcal{J}}_{-rt}, \boldsymbol{X}) - \pi(\mathcal{J}_{rt} - \{j\}, \boldsymbol{\mathcal{J}}_{-rt}, \boldsymbol{X}) \geq \theta + V_{jrt} \quad \text{if } q_{jrt} = 1 \tag{20}$$

The same argument holds for adding and removing products using FBA, and for Amazon.

In this way, I obtain two inequalities (one for added products and one for removed products) that I can use to build a set of moment inequalities. Using one-product deviations to identify fixed costs has been used often in the literature on product variety (Eizenberg, 2014; Wollmann, 2018; Canay et al., 2023; Bontemps et al., 2023). Bontemps et al. (2023) show that considering multiple product deviations does not refine the identified set under the usual economic restrictions on the profit function.

5.3.1 Deriving the moment inequalities

Next, I discuss how to derive the moment inequalities and estimate the fixed cost parameter for 3P-sellers using FBM. The same procedure applies to Amazon and 3P-sellers using FBA.

Equations (19) and (20) need to be transformed into moment inequalities. However, if I assume $\mathbb{E}[V_{jrt}] = 0$, this is no longer the case for $\mathbb{E}[V_{jrt}|q_{jrt}]$ because the sellers observe fixed cost shocks when deciding whether to offer their products. To address this selection bias, I follow Canay et al. (2023) and make the following assumption:

Assumption 2. $|\mathbb{E}[V_{jrt}|q_{jrt}]| \leq \bar{V}$

Here, $\bar{V}$ is an ad-hoc value chosen by the econometrician. While this assumption is relatively strong, it is weaker than assuming that $V_{jrt} = 0$, or $\bar{V} = 0$ (Canay et al., 2023). Moreover, it allows for $\mathbb{E}[V_{jrt}|q_{jrt} = 0]$ and $\mathbb{E}[V_{jrt}|q_{jrt} = 1]$ to differ. It is an alternative to Eizenberg (2014), who assumes that $\theta + V_{jrt}$ is bounded and uses, for the upper bound, the maximum value of $\Delta\pi^-_{jrt} = \pi(\boldsymbol{\mathcal{J}}_{rt}, \boldsymbol{\mathcal{J}}_{-rt}, \boldsymbol{X}) - \pi(\boldsymbol{\mathcal{J}}_{rt} - \{j\}, \boldsymbol{\mathcal{J}}_{-rt}, \boldsymbol{X})$, which can give large bounds.

I show how this assumption leads to two moment inequalities. First, consider Equation (19), which bounds the realized fixed cost from below. Take a set of products and sellers $\{j, r\}$ over a number of weeks t such that q_{jrt} is always equal to zero. Then, by taking the conditional expectation of (19) with respect to the set of products not offered by sellers in some weeks, i.e, those for which $q_{jrt} = 0$ and using Assumption 2, I obtain the following:

$$\mathbb{E}[\Delta\pi^+_{jrt} - \theta - V_{jrt}|q_{jrt} = 0] \leq 0 \tag{21}$$

$$\iff \mathbb{E}[\Delta\pi^+_{jrt}|q_{jrt} = 0] - \theta \underbrace{-\mathbb{E}[V_{jt}|q_{jrt} = 0]}_{\geq -\bar{V}} \leq 0 \tag{22}$$

$$\iff \mathbb{E}[\Delta\pi^+_{jrt}|q_{jrt} = 0] - \theta - \bar{V} \leq 0 \tag{23}$$

Thus, I define a "lower" moment[17] $m^L(\theta) = \Delta\pi^+_{jrt} - \theta - \bar{V}$ such that $\mathbb{E}[m^L(\theta)|q_{jrt} = 0] \leq 0$.

Similarly, I derive an "upper" moment $m^U(\theta) = \Delta\pi^-_{jrt} + \theta - \bar{V}$, where $\Delta\pi^-_{jrt}$ is defined from Equation (20) by $\pi(\boldsymbol{\mathcal{J}}_{rt}, \boldsymbol{\mathcal{J}}_{-rt}, \boldsymbol{X}) - \pi(\boldsymbol{\mathcal{J}}_{rt} - \{j\}, \boldsymbol{\mathcal{J}}_{-rt}, \boldsymbol{X})$, for a product j which is offered by r at week t. We have:

$$\mathbb{E}[m^U(\theta)|q_{jrt} = 1] \leq 0.$$

As a result the identified set for the fixed cost parameter θ is the interval

$$I_\theta = \left[\mathbb{E}[\Delta\pi^+_{jrt}|q_{jrt} = 0] - \bar{V}; \mathbb{E}[\Delta\pi^-_{jrt}|q_{jrt} = 1] + \bar{V}\right].$$

It is obvious that the choice of $\bar{V}$ is critical (although small variations in $\bar{V}$ do not substantially affect our counterfactual policy outcomes). Increasing $\bar{V}$ by one euro mechanically raises the upper bound by one euro and lowers the lower bound by the same amount. I set $\bar{V}$ equal to the mean of the expected values of $V_{jrt}|q_{jrt} = 0$ and $V_{jrt}|q_{jrt} = 1$ assuming normality for these fixed cost shocks. This is the usual Tobit correction known from Heckman (1976). I provide more details in Appendix A.5 about the calibration of the different quantities involved in this calculation. Note that I do not assume a specific distribution for V_{jrt}; instead, I use the Tobit correction to provide a suitable candidate for $\bar{V}$.

[17] This is called lower moment as it allows me to identify the lower bound on θ.

5.3.2 Estimation and Inference

The identified set being an interval, I can estimate it by replacing each expectation by its empirical analogue. Let $\mathcal{J}^+$ and $\mathcal{J}^-$, the set of deviations used to compute $\Delta\pi^+_{jrt}$ and $\Delta\pi^-_{jrt}$, respectively and let N^+ and N^- be their cardinality. The estimated identified set $\widehat{I_\theta}$ is the interval $\left[\hat{\theta}^L; \hat{\theta}^U\right]$ where,

$$\hat{\theta}^L = \frac{1}{N^+} \sum_{\{j,r,t\}\in\mathcal{J}^+} (\Delta\pi^+_{jrt} - \bar{V}), \tag{24}$$

$$\hat{\theta}^U = \frac{1}{N^-} \sum_{\{j,r,t\}\in\mathcal{J}^-} (\Delta\pi^-_{jrt} + \bar{V}). \tag{25}$$

Under standard assumptions, these estimators tend to the true values when both N^+ and N^- tend to infinity. As a result, I obtain a consistent estimator of my identified set.

The last step is to compute the confidence region of the true (unknown) parameter θ, denoted θ^0. In my case, the identified set being an interval, I can follow Imbens and Manski (2004) and I propose a confidence interval of (asymptotic) coverage rate $1-\alpha$:

$$C_{1-\alpha}(\theta^0) = \left[\hat{\theta}^L - q_{1-\alpha}\frac{\hat{\sigma}_L}{\sqrt{N^+}}; \hat{\theta}^U + q_{1-\alpha}\frac{\hat{\sigma}_U}{\sqrt{N^-}}\right],$$

where $q_{1-\alpha}$ is the $1-\alpha$ quantile of the standard normal distribution and $\hat{\sigma}_L$, resp. $\hat{\sigma}_U$, are the standard deviations of the $\Delta\pi^+_{jrt}$s, resp. the $\Delta\pi^-_{jrt}$s. We have:

$$\lim_{N^+,N^-\to\infty} P(\theta^0 \in C_{1-\alpha}(\theta^0)) = 1-\alpha.$$

Here, $\hat{\theta}^L$ and $\hat{\theta}^U$ are sufficiently far enough to consider that the identified set is not reduced to a singleton. Otherwise, I could have used the quantile correction proposed by Stoye (2009) to guarantee the right coverage rate for $C_{1-\alpha}(\theta^0)$.

6 Results

6.1 Demand

In Table 3, I present the results from the demand estimation. In column (1), I show the estimates of the Logit demand model without instruments. This is to show the importance of controlling for price endogeneity: without instruments, demand is almost inelastic, with a median own price elasticity of -0.402. In column (2), I report the estimates of the Logit demand with price instruments, and in column (3) the estimates of the Random Coefficient (RC) Logit demand with price instruments. I now discuss the demand estimates of the RC Logit, which is the demand specification in the paper.

The median own price elasticity is -4.87, a result in line with other papers estimating demand in Amazon like Lee and Musolff (2023), Gutierrez (2022) and Chen and Tsai (2023). In Tai-Lam (2023), demand elasticity for products in the Home & Kitchen category is lower, but it is still quite elastic (around -2). Then, consumers prefer products with higher rating and number of reviews in the previous week, which can be interpreted as two proxies for products' quality. Finally, I find that demand is increasing both for products sold by Amazon and for products sold using Fulfilled by Amazon, although the preference for Amazon is larger. This trend is also observed in Lee and Musolff (2023) and Chen and Tsai (2023), where consumers consistently favor Amazon over third-party sellers, even when controlling for factors such as the Buy Box, as highlighted by Lee and Musolff (2023).

	Estimates		
	(1)	(2)	(3)
Intercept	-8.739 ***	-4.195 ***	-2.586 *
	(0.317)	(0.977)	(1.344)
Price	-0.009 ***	-0.080 ***	-0.108 ***
	(0.000)	(0.014)	(0.020)
Amazon	1.483 ***	1.409 ***	1.374 ***
	(0.021)	(0.026)	(0.035)
Fulfilled by Amazon (FBA)	0.438 ***	0.378 ***	0.395 ***
	(0.013)	(0.018)	(0.025)
Product Rating $_{week-1}$	0.167 ***	0.229 ***	0.299 ***
	(0.024)	(0.027)	(0.038)
Log Reviews $_{week-1}$	0.079 ***	0.069 ***	0.083 ***
	(0.008)	(0.008)	(0.011)
σ			0.018 **
			(0.007)
Product FE	Yes	Yes	Yes
Month FE	Yes	Yes	Yes
IV	No	Yes	Yes
Model	Logit	Logit	RC Logit
Median η_{own}:	-0.402	-3.653	-4.631
Mean η_{own}:	-0.762	-6.924	-7.292

Table 3: **Demand Estimates.** Standard errors in parentheses. * p-value < 0.1, ** p-value < 0.05, *** p-value < 0.01. η_{own}: own price elasticity. N=84149.

In order to evaluate the average effect of being Amazon and of using FBA, I assess how much 3P-sellers should reduce prices in order to have the same demand as Amazon and as a 3P-seller using FBA. First, I compute the market share of a product having price, rating, number of reviews and fixed effects equal to the mean. Then, I compute how the market share changes when the product is sold by Amazon or using FBA, and, finally, the price change necessary to compensate demand for being a 3P-seller and for using FBM. I find that, to have the same demand as Amazon, a 3P-seller using FBM should decrease the price by 24%, while an FBA 3P-seller by 19%. Then, an FBM 3P-seller should reduce prices by 4% in order to have the same demand as an FBA 3P-seller. This larger price cut 3P-seller must make to gain as much market share as Amazon reflects the consumers' preference estimated in the demand model.

Finally, to get an insight into substitution patterns across products offered by Amazon or 3P-sellers, I compute the own- and cross-price elasticities for a random sample of product offerings (Table 4). Overall, I find the highest cross-price elasticity between Beats and Kef products (two brands in the high-end of the markets) offered by Amazon. However, consumers do not seem to substitute more for Amazon products. For instance, when Amazon increases the price of the Beats product, consumers substitute both towards other Amazon's offers and 3P-sellers' offers (after Kef, many consumers substitute Jabra and PowerLocus products offered by 3P-sellers).[18]

	Seller – Product – Brand							
	r1-j1-Sony	r2-j2-PowerLocus	r3-j3-QearFun	r4-j4-Perytong	r5-j5-Jabra	a-j6-Beats	a-j7-Belkin	a-j8-Kef
r1-j1-Sony	**-10.837**	**0.001**	**0.001**	**0.001**	**0.003**	**0.003**	**0.001**	**0.003**
r2-j2-PowerLocus	0	**-3.178**	0	0	0	0	0	0
r3-j3-QearFun	0	0	**-2.497**	0	0	0	0	0
r4-j4-Perytong	0	**0.001**	**0.001**	**-2.099**	0	0	**0.001**	0
r5-j5-Jabra	**0.010**	**0.002**	**0.002**	**0.002**	**-14.772**	**0.025**	**0.002**	**0.026**
a-j6-Beats	**0.030**	**0.003**	**0.003**	**0.002**	**0.067**	**-20.348**	**0.004**	**0.131**
a-j7-Belkin	0	0	0	0	0	0	**-3.429**	0
a-j8-Kef	**0.001**	0	0	0	**0.003**	**0.006**	0	**-26.371**

Table 4: **Own- and Cross- Price Elasticities for a Sample of Offers**. Random sample of 5 products offered by 3P-sellers and 3 products offered by Amazon in the first week of May 2023. Each index represents a product (j) sold by a 3P-seller (r) or Amazon (a) belonging to a specific brand. The (k,i) cell reports the percentage change in the market share of i (column) given a one percent increase in the price of offer k (row).

[18]This result could also driven by the demand specification, which does not include product nests. An alternative model specification could include Amazon and 3P-sellers nests to analyze differences in substitution patterns between the two groups.

6.2 Marginal Costs

I use the supply model to identify the marginal costs and then estimate the marginal cost function in Eq. (8).

	Estimates
Amazon	4.917*** (0.211)
Product FE	Yes
Week FE	Yes

Table 5: **Marginal Cost Estimates**. Standard errors in parenthesis. * p-value < 0.1, ** p-value < 0.05, *** p-value < 0.01. R^2 =0.986.N=84149.

Amazon appears to have higher marginal costs than 3P-sellers: on average, during the same week, Amazon's marginal cost is almost 5 US$ larger than 3P-sellers.

I also compare the marginal cost by computing the difference between 3P-sellers' and Amazon's marginal costs ($\Delta_{jrt} = c_{jrt} - c_{jat}$) for the same product j and week t. As shown in Figure 6, 3P-sellers may not always have the lowest marginal costs.

The reason 3P-sellers have lower marginal costs may arise from the complexity of operating in several markets. As 3P-sellers are more specialized in fewer markets and products compared to Amazon, they might be able to buy products at a lower wholesale price than Amazon. However, the existing literature has provided mixed evidence on the difference between Amazon and 3P-sellers when they offer the same product in the same week. While Tai-Lam (2023) finds that Amazon's marginal costs are lower, Chen and Tsai (2023) show that 3P-sellers using FBA have lower marginal costs.[19]

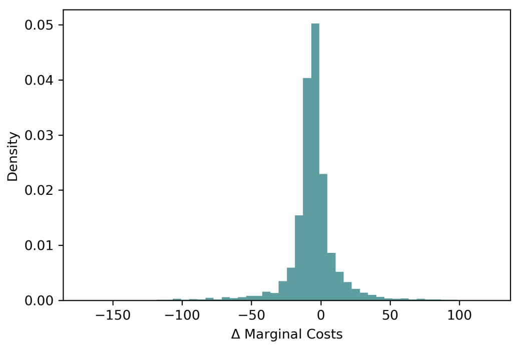


Figure 6: **Difference 3P-sellers' and Amazon's marginal costs for the same product and week**.

[19]In Appendix A.3, I include FBA in the marginal cost function and show that 3P-sellers using FBA face lower marginal costs than Amazon and 3P-sellers using FBM.

6.3 Markups and Lerner Index

Using the identified marginal costs, we can then compare 3P-sellers' and Amazon's markups and Lerner Index (Table 6).[20]
I find that 3P-sellers' markups are similar when FBM or FBM is used to fulfill the product but are much lower than Amazon's markups: on average, Amazon gains 14 US$ more than 3P-sellers. However, 3P-sellers using FBA and Amazon enjoy higher market power than 3P-sellers using FBM. Moreover, 3P-sellers using FBA enjoy more market power than Amazon.

As expected, markups are strongly correlated with prices for all sellers and fulfillment methods, with a correlation coefficient of around 0.95.

	Average Markups	Average Lerner Index	Corr(Markups, Prices)
3P-sellers using FBM	10 US$	17%	0.95
3P-sellers using FBA	9 US$	26%	0.96
Amazon	24 US$	22%	0.96

Table 6: **Average Markups and Lerner Index**. Averages across sellers, products, and weeks.

To gain further insight into the difference in markups and the Lerner Index, I compute the difference when 3P-sellers and Amazon simultaneously offer the same product in the same week. In Figure 7a, I plot the difference in markups between 3P-sellers using FBM and Amazon for the same product j and week t, $\Delta_{jrt} = m_{jrt} - m_{jat}$, while in Figure 7b, I plot the difference in Lerner Index $\Delta_{jrt} = LI_{jrt} - LI_{jat}$. I then compute the same differences between 3P-sellers using FBA and Amazon.

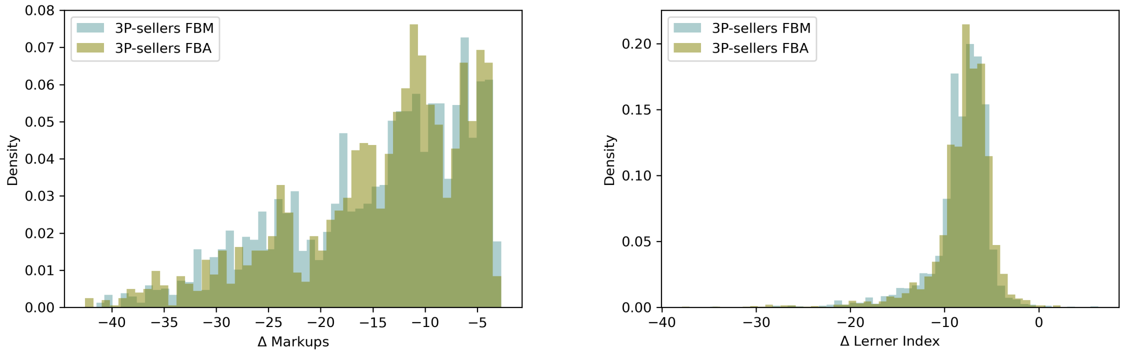
We can see that when competing with Amazon for the same product, 3P-sellers indeed gain a lower markup: on average, 14 US$ less when they use FBA and 15 US$ less when they use FBM. In addition to this, 3P-sellers enjoy less market power than Amazon, regardless of the fulfillment service used: on average, the Lerner Index is 8 pp lower than Amazon for both fulfillment services.

[20] The markups for Amazon, 3P-sellers using FBM, and 3P-sellers using FBA are respectively:

$$m_{jat} = p_{jat} - c_{jat} \quad m_{jrt} = p_{jrt}(1 - \phi_{jrt}) - c_{jrt} \quad \widetilde{m}_{jrt} = p_{jrt}(1 - \phi_{jrt}) - c_{jrt} - \tau$$

Then, the Lerner Index (LI) for Amazon, 3P-sellers using FBM, and 3P-sellers using FBA are respectively:

$$LI_{jat} = \frac{p_{jat} - c_{jat}}{p_{jat}} \quad LI_{jrt} = \frac{p_{jrt}(1 - \phi_{jrt}) - c_{jrt}}{p_{jrt}(1 - \phi_{jrt})} \quad \widetilde{LI}_{jrt} = \frac{p_{jrt}(1 - \phi_{jrt}) - c_{jrt} - \tau}{p_{jrt}(1 - \phi_{jrt})}$$



(a) Markups (b) Lerner Index

Figure 7: **Difference 3P-sellers' and Amazon's markups and Lerner Index for the same product and week**.

Overall, we can see that 3P-sellers using FBM enjoy lower market power than 3P-sellers using FBA and Amazon. This is explained by the lower demand for these offers. 3P-sellers using FBA have larger market power than Amazon when we look at aggregate sales, but this could hide a certain degree of heterogeneity in the products sold: 3P-sellers using FBA might be offering products with larger demand, thereby extracting a larger surplus from consumers. In fact, when offering the same products during the same week, 3P-sellers have lower market power, regardless of the fulfillment method used.

One question then is how to distinguish between product heterogeneity and pricing pressure imposed by Amazon. It could be that the lower market power is competing from competing directly with Amazon and so these 3P-sellers, especially those using FBA, could have an incentive to raise prices more and exert more market power. In the policy analysis in Section 7, I verify this by comparing the changes in prices across all 3P-sellers and across 3P-sellers competing directly with Amazon.

6.4 Fixed Costs

I discuss the results of 3P-sellers' and Amazon's fixed cost parameters for a product each week.[21] I focus on products sold by 3P-sellers and Amazon operating extra retailers, while I exclude products sold directly by their producers. Since Amazon mainly operates as a retailer in this market, in the policy analysis, I study replacement in products for which Amazon is a retailer. To do so, I need the fixed cost parameter of 3P-sellers operating as retailers. For comparison, I also compute Amazon's fixed cost parameter for products sold as a retailer.

[21] I provide more details on the estimation steps in Appendix A.5

Two sources of variation contribute to the identification of the bounds. First, regarding the 3P-sellers, I have assumed the fixed cost parameter is homogeneous across 3P-sellers using the same fulfillment method. Since each 3P-seller offers few products, on average, this helps identify the upper bound. Secondly, sellers operating as retailers can offer any product in the market, except for Amazon private labels, which are only provided by Amazon. This allows us to identify the lower bound.[22]

In the first column of Table 23, I provide the estimated fixed cost parameters computed by Eq (24) and Eq (25). In the second column, I provide the inference on a 95% confidence region.[23]

Comparing the estimates, we can notice that 3P-sellers face higher fixed costs compared when they use FBA compared to using FBM. While using FBA avoids the cost of storage, 3P-sellers also face an opportunity cost from being unable to sell the product through other channels. Regarding Amazon, while the lower bound is comparable to 3P-sellers', the upper bound is much larger. The reason for that is that Amazon's fixed cost includes the opportunity cost in terms of lost fee revenues from 3P-sellers, and this could vary largely across different products.[24]

$\hat{\theta}$	Estimated Bounds (US$)	95% Confidence Region (US$)
θ	$[\,11,\ 50\,]$	$[\,9\ ,\ 54\,]$
$\widetilde{\theta}$	$[\,14,\ 73\,]$	$[\,11,\ 77\,]$
θ_a	$[\,20,\ 261\,]$	$[\,8,\ 284\,]$

Table 7: **Estimated Bounds and Confidence Region for Fixed Cost Parameter for Retailers.** θ: fixed cost parameter for 3P-sellers using FBM; $\widetilde{\theta}$: fixed cost parameter for 3P-sellers using FBM; θ_a: fixed cost parameter Amazon. $\bar{V}$: 19 US$ for 3P-sellers using FBM; $\bar{V}$: 21 US$ for 3P-sellers using FBA; $\bar{V}$: 119 US$ for Amazon. $N^+ = 140$ and $N^- = 118$ for θ; $N^+ = 210$ and $N^- = 227$ for $\widetilde{\theta}$; $N^+ = 186$ and $N^- = 111$ for θ_a.

To gain further insight into the magnitude of the fixed costs, I compare the median fixed cost per unit sold and the median markup for Amazon, 3P-sellers using FBM, and 3P-sellers using

[22] The problem is different from other product variety games, such as Canay et al. (2023) with soft drinks. In this case, each seller can offer only the products it produces. Therefore, in order to compute $\Delta^+\hat{\pi}$, changes in product offerings across geographical markets should be observed. Whereas, if the seller were offering all its products in every market, we would not be able to compute $\Delta^+\hat{\pi}$ and so identify the lower bound.

[23] In Appendix A.6, I compare the confidence region obtained using Eizenberg (2014) and show that the results are quantitatively similar.

[24] Another reason why Amazon's fixed costs are so high is heterogeneity across products. If vertical differentiation were higher for Amazon than for 3P-sellers', this would generate larger bounds on the fixed costs, too.

FBA.[25] As we can see in Table 8, the profit per unit can vary substantially for the sellers due to the large bound on the fixed cost per unit. Moreover, the difference in profits for 3P-sellers using FBM and 3P-sellers using FBA is not too large. This is because 3P-sellers using FBA also have to pay the FBA fee. Although the FBA fee is not too big in absolute terms, in relative terms, it can have a large impact on 3P-sellers' profit. Without FBA fee, the median markup would be 11.53$, so the 3P-sellers have to pay around 30% of the markup to Amazon. This implies that, while using FBA attracts more consumers, the overall profit for 3P-sellers does not increase accordingly. For Amazon, we can see that the profit per unit varies more than 3P-sellers, in line with the estimated larger bound on the fixed cost. Therefore, on the one hand, larger markups give Amazon more incentive to enter, reflected in its strong presence in the market. On the other hand, Amazon may also face large fixed costs, preventing it from entering more products. Once again, this might be explained by Amazon's opportunity cost in terms of lost revenues from 3P-sellers' fees, implying a lower benefit from offering certain products, even in cases were markups were larger than 3P-sellers.

In conclusion, estimation of the fixed cost has shown that for both Amazon and 3P-sellers, the fixed cost can correspond to a substantial portion of sellers' markup. Therefore, accounting for fixed costs becomes crucial to understanding which products sellers choose to offer and the impact of different policies.

	Fixed Cost per Unit (US$)	Markup (US$)	Profit Per Unit (US$)
3P-sellers using FBM	$[\,1.26,\ 7.58\,]$	9.73	$[\,2.35,\ 8.49\,]$
3P-sellers using FBA	$[\,1.1,\ 7.67\,]$	8.31	$[\,0.99,\ 7.41\,]$
Amazon	$[\,1.7,\ 18.02\,]$	22.77	$[\,1.72,\ 21.59\,]$

Table 8: **Median Fixed Cost per unit, Median Markup and Median Profit per unit for the Offered Products.** Average fixed cost per unit sold computed assuming $V_{jdt} = 0$. Both average fixed cost and average markup are computed using the entire sample. For 3P-sellers, products with a weekly profit below the 30th percentile are excluded. For Amazon, products with a weekly profit below the 40th percentile are excluded. truncation justified by presence of 3P-sellers on the fringe $N = 3453$, $\widetilde{N} = 28229$, $N_a = 4454$.

[25]To compute the average fixed cost per unit sold, I use the bounds from the confidence region in Table 23 and assume that $V_{jdt} = 0$. The lower and the upper bound of the average fixed are $\bar{F}^L_{jrt} = \theta^L/Q_{jrt}$ and $\bar{F}^U_{jrt} = \theta^U/Q_{jrt}$ if r uses FBM; $\bar{F}^L_{jrt} = \widetilde{\theta}^L/Q_{jrt}$ and $\bar{F}^U_{jrt} = \widetilde{\theta}^U/Q_{jrt}$ if r uses FBA; $\bar{F}^L_{jat} = \theta^L_a Q_{jat}$ and $\bar{F}^U_{jat} = \theta^U_a/Q_{jat}$ if Amazon.

7 Policy Analysis

7.1 Setup

In this section, I discuss the effect of structural separation, where Amazon is banned from offering any product, both as a retailer and with its brands. In particular, I focus on the effect of replacing Amazon in the products offered before the ban.

Results from the estimation of demand and fixed cost have highlighted the following trade-off. On the one hand, consumers have a higher demand for products when they are sold by Amazon rather than 3P-sellers: once Amazon is banned, they might decide not to buy any product or substitute to less valued sellers. On the other, when offering the same product as Amazon, 3P-sellers pay a lower wholesale price and exert lower market power than Amazon. This implies that the 3P-sellers replacing Amazon might actually charge lower prices. This price effect depends on how much the price decrease compensates for lower consumers' preferences and on the other 3P-sellers' incentive to keep prices low once Amazon is banned. Moreover, 3P-sellers can have a high incentive to replace Amazon in many products after the ban: this is because more consumers may substitute for 3P-sellers, thereby decreasing the fixed cost per unit sold and increasing the total profit.

The policy analysis unfolds as follows. First, I compute a counterfactual where all Amazon's offers are banned, and there is no entry. Secondly, I compute a counterfactual where an entrant 3P-sellers can decide whether to offer or not the products previously offered by Amazon. The entrant 3P-seller is a new (*fictitious*) 3P-seller that operates as a retailer and enters the headphones market on Amazon with zero offers. The set of products it can offer every week corresponds to the products Amazon was offering as a retailer before the ban. Hence, every week the entrant 3P-seller can offer from zero up to the number of products previously offered by Amazon. As mentioned before, this set of products constitutes the majority of the products offered by Amazon, accounting for almost all of Amazon's sales. Then, incumbent 3P-seller offers are kept fixed. Therefore, in both counterfactual scenarios, incumbent 3P-sellers can adjust prices but not the choice of products. The motivation is twofold. First of all, I am interested in evaluating the effect of 3P-sellers replacing Amazon in the products it was offering. The first-order effect will then come from having one 3P-seller choosing to enter this set of products. The second reason is reducing the computational burden. If we were incorporating every 3P-seller decision, we would need to calculate the expected profit of each 3P-seller under any possible product assortment. Given the large number of potential product assortments, this is unpractical. Instead, by incorporating one entrant, I limit the number of potential product assortments, making computation more tractable.

In order to distinguish between an entrant using FBM and an entrant using FBM, in the first counterfactual with entry, I assume the entrant can offer products only through FBM. In contrast, in the second counterfactual I assume the entrant can offer products only through FBA. I keep the incumbent 3P-sellers fulfillment service fixed.

I explain now procedure to estimate the counterfactual with a 3P-seller entrant here below.[26] Consider the set of potential product offerings for the entrant. The goal is to find which of them are offered in equilibrium, keeping the incumbent products offerings fixed.

I denote by $\hat{\pi}_{Et}$ the expected variable profit of the entrant and by $\mathcal{J}_{Et}$ one of the potential product offerings in $\boldsymbol{\mathcal{J}}_{Et}$ the entrant can choose from. The procedure unfolds as follows:

1. **Expected Profits**. In the first step, I compute the expected variable profit of the entrant 3P-seller for each potential product offerings.[27]

2. **Dominated Strategies**. The second step eliminates all the dominated strategies. Let θ^L denote the lowest possible value of θ from the estimated confidence region and assume $\bar{V} = 0$. For each product portfolio $\mathcal{J}_{Et} \in \boldsymbol{\mathcal{J}}_{Et}$, consider the inequality

$$\hat{\pi}_{Et}(\mathcal{J}_{Et}, \boldsymbol{\mathcal{J}}_{-Et}) - J_{Et} \cdot \theta^L \geq 0 \tag{26}$$

Denote by $\boldsymbol{\mathcal{J}}^2_{Et} \subset \boldsymbol{\mathcal{J}}_{Et}$ the set of all product offerings $\mathcal{J}_{Et}$ for which Eq. 26. This is then the set of not dominated strategies which can be a candidate equilibrium.

3. **Equilibrium Strategies**. In the third step, I find the equilibrium strategies. For each $\mathcal{J}_{Et} \in \boldsymbol{\mathcal{J}}^2_{Et}$, I check the condition

$$\hat{\pi}_{ft}(\mathcal{J}_{Et}, \boldsymbol{\mathcal{J}}_{-Et}) - \hat{\pi}_{ft}(\mathcal{J}'_{Et}, \boldsymbol{\mathcal{J}}_{-Et}) - J^{IN}_{Et} \cdot \theta^L + J^{OUT}_{Et} \cdot \theta^U \geq 0 \tag{27}$$

where J^{IN}_{Et} is the number of products in $\mathcal{J}_{Et}$ but not in $\mathcal{J}'_{Et}$, J^{OUT}_{Et} is the number of products in $\mathcal{J}'_{Et}$ but not in $\mathcal{J}_{Et}$, and θ^U is the highest possible value of θ from the fixed costs. Let $\boldsymbol{\mathcal{J}}^3_{Et} \subset \boldsymbol{\mathcal{J}}^2_{Et}$ denote the set of product portfolios J_{Et} for which Eq. 27 holds for all $J'_{Et} \in \boldsymbol{\mathcal{J}}^2_{Et}$. If $\boldsymbol{\mathcal{J}}^3_{Et}$ is a singleton, then there is a unique equilibrium. If it contains more than one product offerings, we have multiple equilibria.

4. **Summary of the Results**. Construct bounds on the objects of interest (average prices, profits, CS etc.) by finding max and min values for those objects in $\boldsymbol{\mathcal{J}}^3_{Et}$.

[26] The procedure follows Canay et al. (2023). I only depart by assuming $\bar{V} = 0$. This follows the common practice in counterfactual analysis not to take into account the error on the estimated parameter.

[27] Given the large number of potential offerings $(2^J - 1)$, I split the products in four clusters according to the level of price and assume that the entrant offers products in the same group together. Therefore, the number of possible combinations of product clusters is $2^4 - 1 = 15$.

7.2 Results

In Table 10, I present the results of the policy experiment. With no additional entry, consumer surplus decreases by 19.29% after the ban. This can be explained by three reasons. First, consumers like buying from Amazon, whereas after the ban, they can only access less-valued sellers. Second, many products are lost: since Amazon was the only active seller in around half of the products offered, these are lost after the ban. Third, there is a small, not irrelevant increase in prices: incumbent 3P-sellers raise prices by 0.37%, accounting for 1.2 pp of the welfare loss. In particular, we can see that the increase in prices is even smaller for 3P-sellers offering products where Amazon was present. This implies that, compared 3P-sellers in other products, 3P-sellers competing in the same product with Amazon do not have a larger incentive to increase markups. Then, Amazon profits fall, too, as Amazon gets lower profits, although the revenues from fees increase by 25.15%. Indeed, 3P-sellers get greater demand and see their profit increase by 19.48%, mainly due to consumers substituting Amazon for their products.

With an FBA entrant, consumer surplus increases by 4.54 pp compared to the case with no entry. This is because the FBA entrant enters up to 84% of the products that were offered by Amazon before the ban, restoring the product offerings which was previously provided by Amazon.[28] Moreover, the FBA entrant prices are higher, as the entrant pays both the referral and the FBA fee to Amazon. This makes the results with a FBM entrant quantitatively similar. The FBM entrant enters up to 84% of the products previously offered by Amazon, and while consumers have a larger preference for FBA sellers, the FBM entrant decreases compared to the prices charged by Amazon before the ban. Moreover, in both scenarios with entry, there is a small increase in pricing pressure.

Overall, in both scenarios, we see that an entrant 3P-seller improves consumers' welfare by restoring many product offerings. However, this is not enough to compensate for the loss in consumer surplus due to higher demand for products sold by Amazon and because prices do not decrease enough to compensate for this.

Compared to the existing literature, the consumer surplus loss is larger than previously found in Gutierrez (2022) and Tai-Lam (2023). In the former, consumer surplus decreases by 3%, while in the latter by 7%. These differences mainly come from the demand. Gutierrez (2022) does not include consumers' preference for Amazon in the demand model so the loss comes from the lower number of offers and larger 3P-sellers' markups. In Tai-Lam (2023), the loss on consumer surplus is larger since demand increases for products sold by Amazon.

[28]In Appendix A.7, I show which combinations of clusters of products were chosen in equilibrium.

	Basis	No Entry	FBM Entrant	FBA Entrant
		Ban on Amazon Sales		
	(US$)	(%Δ)	[min %Δ, max %Δ]	[min %Δ, max %Δ]
Consumers				
Consumer Surplus	28′570′188	-19.29	$[-16.62, -14.95]$	$[-16.58, -14.75]$
Amazon				
Amazon Variable Profit	20′470′192	-26.25	$[-23.67, -22.28]$	$[-22.25, -20.85]$
Amazon Fees Revenues	12′062′924	25.15	$[29.15, 31.49]$	$[31.91, 34.28]$
Incumbent 3P-sellers				
Incumbent Variable Profit	10′337′264	19.48	$[13.66, 16.13]$	$[13.42, 16.05]$
Incumbent Prices	81	0.37	$[0.28, 0.3]$	$[0.27, 0.3]$
Entrant 3P-seller				
Incumbent Prices	168	0.28	$[0.22, 0.23]$	$[0.21, 0.22]$
Share Products Entered	-	-	$[0.72, 0.84]$	$[0.67, 0.84]$
Entrant Prices (vs. Amazon)	-	-	$[-3.67, -2.15]$	$[1.96, 3.1]$
Products Lost	-	49.27	$[6.69, 12.7]$	$[6.74, 15.39]$

Table 9: **Structural Separation: Ban on Amazon Sales**. Products Lost: share of the products offered by Amazon and not offered by other sellers after the ban. Entrant Prices (vs. Amazon): difference with Amazon's prices for the same product before the ban.

	Basis	No Entry	FBM Entrant	FBA Entrant
		Ban on Amazon Sales		
			[min , max]	[min , max]
Market Share Outside Option	60	62.87	$[62.18, 62.79]$	$[62.18, 62.8]$

Table 10: **Market Share of the Outside Option**.

8 Conclusion

In this paper, I evaluate the effect of structural separation on Amazon.com using data from the headphones market. I find that Amazon has a significant market presence, offering many products as a retailer, which consumers particularly value. Moreover, many products are offered exclusively by Amazon. This suggests that a ban on Amazon's sales would likely harm consumers. However, I also find that 3P-sellers incur lower marginal costs and hold less market power than Amazon. Therefore, if 3P-sellers replaced Amazon's offers, they could partially compensate for consumer losses. To evaluate this, I estimate a structural model of entry and pricing and use the model to assess the impact of an entrant 3P-seller replacing Amazon after the ban. I find that the increased profits of incumbents strongly incentivize the 3P-seller to offer many of the products previously sold by Amazon. However, the 3P-seller does not charge prices low enough to compensate for the lower demand, so overall, consumers are worse off.

Therefore, this study demonstrates that policies enforcing structural separation on vertically integrated platforms may harm consumers. However, I also show that non-price effects, such as 3P-sellers' entry, play an important role in mitigating the negative effects of the policy.

The model can be extended in several directions. First of all, we could evaluate incumbents' incentives to offer more products. As these results suggest, this could compensate even more for the loss for consumers by increasing the number of products they have access to. Then, sellers gaining zero market shares are not part of the current model but could be included by analyzing how their sales change after the ban. Finally, the model assumes high search costs for non-Buy-Box offers, meaning all sales go to the Buy-Box seller. The goal, then, is to quantify the percentage of sales typically occurring out of the Buy-Box in this specific market and consider this in the model.

The paper also provides many avenues for future research. Firstly, it would be interesting to analyze how the ban impact other dimensions, such as fees and logistics choices. The model could be extended to include an extra stage where the platform sets the fees and where 3P sellers choose the logistics. Then, an important aspect of competition policy is whether the Buy-Box algorithm is biased towards Amazon and how much this affects consumers. I can evaluate this by incorporating the Buy-Box algorithm using a demand model with search costs (Moraga-Gonzalez et al., 2023).[29]. Then, the model can be extended to account for upstream firms and analyze how entry and wholesale prices change as a consequence of structural separation.

[29]I provide a framework of this model in Appendix C.

References

S. V. Berg. *Annals of Economic and Social Measurement, Volume 5, number 4*. NBER, 1976.

S. Berry, J. Levinsohn, and A. Pakes. Automobile prices in market equilibrium. *Econometrica*, 63(4):841–890, 1995.

S. T. Berry and P. A. Haile. Chapter 1 - foundations of demand estimation. In K. Ho, A. Hortaçsu, and A. Lizzeri, editors, *Handbook of Industrial Organization, Volume 4*, volume 4 of *Handbook of Industrial Organization*, pages 1–62. Elsevier, 2021.

C. Bontemps, C. Gualdani, , and K. Remmy. Price competition and endogenous product choice in networks: Evidence from the us airline industry. TSE Working Paper, N°1415, 2023.

L. Cabral and L. Xu. Seller reputation and price gouging: Evidence from the covid-19 pandemic. *Economic Inquiry*, 59(3):867–879, 2021.

I. A. Canay, G. Illanes, and A. Velez. A user's guide for inference in models defined by moment inequalities. *Journal of Econometrics*, page 105558, 2023. ISSN 0304-4076.

N. Chen and H.-T. Tsai. Steering via algorithmic recommendations. SSRN Working Paper, 2021.

N. Chen and H.-T. Tsai. Price competition under information (dis)advantage. SSRN Working Paper, 2023.

D. Condorelli, A. Galeotti, and V. Skreta. Selling through referrals. *Journal of Economics & Management Strategy*, 27(4):669–685, 2018.

C. Conlon and J. Gortmaker. Best practices for differentiated products demand estimation with pyblp. *The RAND Journal of Economics*, 51(4):1108–1161, 2020.

G. Crawford, M. Courthoud, R. Seibel, and S. Zuzek. Amazon entry on amazon marketplace. CEPR Discussion Paper DP17531, 2022.

A. Eizenberg. Upstream innovation and product variety in the u.s. home pc market. *The Review of Economic Studies*, 81(3 (288)):1003–1045, 2014.

F. Etro. Product selection in online marketplaces. *Journal of Economics & Management Strategy*, 30(3):614–637, 2021.

Y. Fan and C. Yang. Estimating discrete games with many firms and many decisions: An application to merger and product variety. NBER Working Paper 30146, 2022.

C. Farronato, A. Fradkin, and A. MacKay. Self-preferencing at amazon: Evidence from search rankings. *AEA Papers and Proceedings*, 113:239–43, May 2023.

M. S. Goeree. Limited information and advertising in the u.s. personal computer industry. *Econometrica*, 76(5):1017–1074, 2008.

A. Gomez-Losada and N. Duch-Brown. Competing for amazon's buy box: A machine-learning approach. *Lecture Notes in Business Information Processing*, 373:445–456, 2019. ISSN 1865-1348 (online).

A. Goolsbee and J. Chevalier. Measuring prices and price competition online: Amazon and barnes and noble. *Quantitative Marketing and Economics*, 1, 08 2002.

G. Gutierrez. The welfare consequences of regulating amazon. SSRN Working Paper, 2022.

A. Hagiu and J. Wright. Marketplace or reseller? *Management Science*, 61(1):184–203, 2015.

A. Hagiu, B. Jullien, and J. Wright. Creating platforms by hosting rivals. *Management Science*, 66(7):3234–3248, 2020.

S. He and B. Hollenbeck. Sales and rank on amazon.com. SSRN Working Paper, 2020.

J. J. Heckman. *The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models*, pages 475–492. In Berg (1976), October 1976.

M. Hunold, U. Laitenberger, and G. Thébaudin. Bye-box: An analysis of non-promotion on the amazon marketplace. CRED Working Paper N° 2022-4, 2022.

G. W. Imbens and C. F. Manski. Confidence intervals for partially identified parameters. *Econometrica*, 72(6):1845–1857, 2004.

D.-S. Jeon and P. Rey. Platforms, hold up and innovation. mimeo, 2022.

K. Jerath and Z. J. Zhang. Store within a store. *Journal of Marketing Research*, 47(4):748–763, 2010.

B. Jiang, K. Jerath, and K. Srinivasan. Firm strategies in the "mid tail" of platform-based retailing. *Marketing Science*, 30(5):757–775, 2011.

B. Kline, A. Pakes, and E. Tamer. Moment inequalities and partial identification in industrial organization. In *Handbook of Industrial Organization*, volume 4, pages 345–431. Elsevier, 2021.

J. Kuehn. Spillovers from entry: the impact of bank branch network expansion. *The RAND Journal of Economics*, 49(4):964–994, 2018.

K. H. Lee and L. Musolff. Entry into two-sided markets shaped by platform-guided search. Working Paper, 2023.

N. Martinez. Market coverage and network competition: Evidence from shared electric scooters. Working Paper, 2023.

J. Moraga-Gonzalez, Z. Sandor, and M. Wildenbeest. A framework for the estimation of demand for differentiated products with simultaneous search. TI 2023-015/VII Tinbergen Institute Discussion Paper, 2023.

A. Pakes. Alternative models for moment inequalities. *Econometrica*, 78(6):1783–1822, 2010.

D. Raval. Steering in one click: Platform self-preferencing in the amazon buy box. Working Paper, 2023.

I. Reimers and J. Waldfogel. A framework for detection, measurement, and welfare analysis of platform bias. NBER Working Paper 31766, 2023.

J. Stoye. More on confidence intervals for partially identified parameters. *Econometrica*, 77(4): 1299–1315, 2009.

H. Tai-Lam. Platform search design and market power. Working Paper, 2023.

W. Wen and F. Zhu. Threat of platform-owner entry and complementor responses: Evidence from the mobile app market. *Strategic Management Journal*, 40(9):1336–1367, 2019.

T. G. Wollmann. Trucks without bailouts: Equilibrium product characteristics for commercial vehicles. *American Economic Review*, 108(6):1364–1406, June 2018.

F. Zhu and Q. Liu. Competing with complementors: An empirical look at amazon.com. *Strategic Management Journal*, 39(10):2618–2642, 2018.

A Appendix

A.1 Data Appendix

In this section, I describe how the dataset is constructed and the variable in the dataset. I collect data on all listed products in a market from Keepa, where each product is distinguished by a unique number called ASIN.

For each product, I observe the all the real-time changes in sales rank, number of reviews, sellers' logistics, prices and shipping costs since the product was tracked by Keepa. Other information include the changes in Buybox (Buybox seller and price), product's title and description.

I rearrange the data in order to create a weekly panel of offers for the products. An offer is defined as a seller-day-ASIN combination and it contains the price, shipping cost, sales rank, number of reviews and rating for a seller of an ASIN in a certain day. Then, I will add whether the seller was in the Buybox during that day and compute the percentage of time spent in the Buy-Box.

One issue which might occur is that the Buybox seller may have more than one offer for the same ASIN varying only by the logistic method, that is, there will be an offer using FBA and one using FBM. Since I do not know which of the two offers is contained in the Buybox, I will assume that the FBA offer is the Buybox one.

Finally, I will add the product characteristics. Since these are contained in the product description and the title, I use the following strategy. First, I download all the keywords associated with "headphones" from AmzScout, another market intelligence company; this set includes approximately 190 keywords. In addition to the word headphones, these keywords are informative of the most salient characteristics as usually consumers search for products having particular characteristics e.g. noise-cancelling, sleeping, kids, iphone, gaming. Therefore, while these characteristics are not exhaustive (for instance, they do not capture more technical characteristics or aesthetic features), they are useful as a starting cream-skim of product differentiation. To extract these characteristics, I first select the first 100 keywords in terms of average number of monthly searches during the sample period. After some text cleaning and after having removed the brands name, I am left with about 40 words. Then, I add each word the the panel dataset as a dummy equal to 1 if the word is contained in the text or in the title of the ASIN, and 0 otherwise.

A.1.1 Estimated Quantity sold

Different methods have been proposed to approximate the quantity sold from the sales rank. Goolsbee and Chevalier (2002) and He and Hollenbeck (2020) estimate a Pareto distribution model using the category sales rank and actual sales [30]. By taking the logs, the Pareto distribution model can be transformed into a relationship between the log of sales rank and the log of sales.

$$\log(Quantity_t) \approx \alpha - \beta \log(SalesRank_t) \tag{28}$$

In the case of He and Hollenbeck (2020), the model is estimated using the average sales data at the weekly level and the weekly observations of the category sales rank. Chen and Tsai (2021) estimate the same model using the daily sales rank. However, since they do not have actual data on sales, they assume β in the regression is equal to one, while, given their model specification, they do not need to estimate α.

Finally, Gutierrez (2022) collects sales estimates from two leading market intelligence companies for Amazon sellers, AmzScout and JungleScout. These companies use data on actual sales and SalesRank in given period to estimate the relationship between them. Since it appears that AmzScout employs a power test model, Gutierrez (2022) uses a sample of estimated quantities and SalesRank from this website in order to retrieve the estimated parameters; then, he repeats the same procedure for a sample from JungleScout, but this time using a spline, which is the model JungleScout seems to employ. He uses the estimated parameters in order to find estimates of quantities sold in his dataset.

Here, I take a similar approach to Gutierrez (2022) and, using data from AmzScout, I estimate a model which approximates the one used by AmzScout. To collect the sample, I start from a electronics sales rank [31] equal to 1 and then double the sales rank until the estimated quantity sold remains constant.

As Gutierrez (2022), we do not know the precise model used by AmzScout to estimate the quantity sold. Therefore, I start from the simplest model found in Chevalier and Goolsbee (2003) and He and Hollenback (2020).

$$log(\hat{Q})\_AmzScout = \alpha - \beta log(sales\_rank\_electronics) \tag{29}$$

[30] Goolsbee and Chevalier (2002) use data from a seller and own experiments for the book category. He and Hollenbeck (2020) compute sales using changes in inventory reported by Amazon.com

[31] Electronics is the rootcategory in case for headphones

	(1)
VARIABLES	$log(\hat{Q})\_AmzScout$
$log(sales\_rank\_electronics)$	-0.874***
	(0.0661)
Constant	12.59***
	(0.419)
Observations	24
R-squared	0.888

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 11: Log-log regression estimation

Finally, I use $\hat{\alpha}$ and $\hat{\beta}$ to compute the estimated quantity sold in my data.

A.1.2 Market Size

I use data from the estimated number of keyword search from AmzScout. First, I collect information on all the keywords with two to five words containing the word *headphones*; according to AmzScout, this corresponds to about 700 keywords. Together with the keywords, AmzScout provides also the estimated number of searches in a month.

For the above application, I assume that the headphones market can be divided in different subgroups and I start considering the "wireless bluetooth" subgroup. So, I select all the keywords containing the word *wireless*, *bluetooth*, or both, and I assume that every single search corresponds to a consumer. Therefore, the market size for the "wireless bluetooth" subgroup is built as the sum of the market sizes for each separate keyword.

Wireless Bluetooth Keywords		
Headphones Wireless Bluetooth	Wireless Headphones	Bluetooth Headphones
Beats Headphones Wireless Bluetooth	Headphones Bluetooth	Headphones Wireless Bluetooth Noise Cancelling
Beats Solo3 Wireless On-Ear Headphones	Wireless Headphones Over Ear	Sleep Headphones Wireless Bluetooth
Wireless Headphones Bluetooth	Bluetooth Headphones Wireless	Sennheiser Headphones Wireless
Bluetooth Headphones with Mic	Noise Cancelling Headphones Bluetooth	Dual Wireless TV Headphones
Tribit Xfree Tune Bluetooth Headphones	Bluetooth Headphones Over the Ear	JBL Headphones Wireless Bluetooth
TV Headphones Wireless	Bluetooth Eyemask Sleep Headphones	P47 Wireless Headphones
Kids Bluetooth Headphones	Mpow Headphones Wireless Bluetooth	White Headphones Wireless
Wireless Headphones for TV	White Wireless Headphones	Sony Wireless Headphones
Bose Headphones Wireless Bluetooth	iPhone Headphones Bluetooth	Sony WH-CH500/B Stamina Wireless Headphones
Bluetooth Headphones Wireless Earbuds	Open Ear Headphones Wireless Bluetooth	Wireless Kids Headphones
Sleep Mask with Bluetooth Headphones	Tagry Bluetooth Headphones	Bluetooth Headphones for Kids
Spiderman Kids Volume-Limiting Bluetooth Headphones	Neckband Bluetooth Headphones	Wireless Headphones with Microphone
Headphones Wireless	Wireless Headphones Gaming	Bluetooth Running Headphones
Lenovo TH30 Wireless Bluetooth Headphones	Sony Headphones Wireless Bluetooth	Wireless Gaming Headphones
Beribes Headphones Wireless Bluetooth	Veatool Bluetooth Headphones	Kids Headphones Bluetooth
Sleep Headphones Bluetooth Headband		

Table 12: Keywords containing *headphones* + *wireless* and/or *bluetooth*.

Month	monthly market size
March 2023	246’000
April 2023	319’340
May 2023	1’271’830
June 2023	804’670
July 2023	919’990
August 2023	1’010’870
September 2023	441’970

Table 13: Potential Market by Month for Bluetooth-wireless headphones

A.2 Marginal Costs

In Stage 2, Amazon and 3P-sellers play a Bertrand pricing game and choose prices to maximize the per-period variable profits from Eq. **??** and Eq. **??**:

$$\max_{p_{jR}} \pi_{Rt} = \sum_{j \in \mathcal{J}_{Rt} \cap FBM_t} \mathcal{M}_t \cdot s_{jRt}(\boldsymbol{p}_t, \boldsymbol{X}_t) \cdot \big((1-\phi_{jRt}) \cdot p_{jRt} - c_{jRt}\big)$$

$$+ \sum_{j \in \mathcal{J}_{Rt} \cap FBA_t} \mathcal{M}_t \cdot s_{jRt}(\boldsymbol{p}_t, \boldsymbol{X}_t) \cdot \big((1-\phi_{jRt}) \cdot p_{jRt} - c_{jRt} - \tau\big) \tag{30}$$

$$\max_{p_{jA}} \pi_{At} = \sum_{j \in \mathcal{J}_{At}} \mathcal{M}_t \cdot s_{jAt}(\boldsymbol{p}_t, \boldsymbol{X}_t) \cdot (p_{jAt} - c_{jAt}) + \sum_{R} \sum_{j \in \mathcal{J}_{Rt}} \phi_{jRt} \cdot p_{jRt} \cdot \mathcal{M}_t \cdot s_{jRt}(\boldsymbol{p}_t, \boldsymbol{X}_t)$$

$$+ \sum_{R} \sum_{j \in \mathcal{J}_{Rt} \cap FBA_t} \tau \cdot \mathcal{M}_t \cdot s_{jRt}(\boldsymbol{p}_t, \boldsymbol{X}_t) \tag{31}$$

Given the competition assumption, we can derive the sellers' marginal costs.

To better display the marginal cost function, I rewrite the variable profit functions in Eq. 30 and Eq. 31 into one function:

$$\Pi_{jft} = \sum_{j \in \mathcal{J}_{ft}} \mathcal{M}_t \cdot s_{jft}(\boldsymbol{p}_t, \boldsymbol{X}_t) \cdot \big((1-\phi_{jft} \cdot \mathbb{1}_{f=R}) \cdot p_{jft} - c_{jft} - \tau \mathbb{1}_{f=R \,\cap\, j \in FBA_t}\big) +$$

$$\mathbb{1}_{f=A} \cdot \sum_{R} \sum_{j \in \mathcal{J}_{Rt}} \phi_{jft} \cdot p_{jRt} \cdot \mathcal{M}_t \cdot s_{jRt}(\boldsymbol{p}_t, \boldsymbol{X}_t) + \mathbb{1}_{f=A} \cdot \sum_{R} \sum_{j \in \mathcal{J}_{Rt} \cap FBA_t} \tau \cdot \mathcal{M}_t \cdot s_{jRt}(\boldsymbol{p}_t, \boldsymbol{X}_t) \tag{32}$$

The first order condition is then

$$\text{FOC}_{p_{jft}} : (1 - \mathbb{1}_{f=R} \cdot \phi_{jft}) \cdot s_{jft} + \sum_{k \in \mathcal{J}_f} [(1 - \mathbb{1}_{f=R} \cdot \phi_{jft}) \cdot p_{kst} - c_{kst} - \tau \mathbb{1}_{f=R \,\cap\, k \in FBA_t}] \cdot \frac{\partial s_{kst}}{\partial p_{jft}} +$$

$$\mathbb{1}_{f=A} \sum_{R} \sum_{k \in \mathcal{J}_{Rt}} \phi_{jft} \cdot p_{kRt} \cdot \frac{\partial s_{kRt}}{\partial p_{jft}} + \mathbb{1}_{f=A} \sum_{R} \sum_{j \in \mathcal{J}_{Rt} \cap FBA_t} \tau \cdot \frac{\partial s_{kRt}}{\partial p_{jft}} = 0 \tag{33}$$

Let JFt be the total number of offers in period t. Then, in matrix notation, Eq. (33) is equivalent to

$$(\boldsymbol{1} - \mathbb{1}_{f=R} \odot \boldsymbol{\phi}) \odot S + \widetilde{\Omega}[(\boldsymbol{1} - \mathbb{1}_{f=R} \odot \boldsymbol{\phi}) \odot P - C - \tau \mathbb{1}_{f=R \,\cap\, k \in FBA_t}]$$

$$+ \mathbb{1}_{f=A} \widetilde{\Omega}' \boldsymbol{\phi} \odot P + \mathbb{1}_{f=A} \widetilde{\Omega}' \boldsymbol{\phi} \odot \tau \mathbb{1}_{f=R \,\cap\, k \in FBA_t} = 0 \tag{34}$$

where $\boldsymbol{1}$, $\mathbb{1}_{f=R}$, $\mathbb{1}_{f=A}$, $\boldsymbol{\phi}$, P, C are a column vector of length JFt, and $\widetilde{\Omega}$ and $\widetilde{\Omega}'$ are square matrices of length JFt.

$\widetilde{\Omega}$ is equal to $O \odot \Omega$: the first term, O, is the ownership matrix, whose term is equal to 1 when

the row-index offer and the column-index offer belong to the same seller, and 0 otherwise; the second term, Ω, is the matrix of market shares derivatives with respect to prices.
Then, $\widetilde{\Omega}'$ is equal to $O' \odot \Omega$: the first term, O', is the non-ownership matrix, whose term is equal to 1 when the row-index offer and the column-index offer do not belong to the same seller, and 0 otherwise; the second term, Ω, is the same as defined before [32].
From Eq. (34), I can then derive the vector of marginal costs C:

$$C = (1 - \mathbb{1}_{s=R} \odot \boldsymbol{\phi}) \odot [\tilde{\Omega}^{-1} S + P] - \tau \mathbb{1}_{f=R \,\cap\, k \in FBA_t}$$

$$+ \tilde{\Omega}^{-1} \tilde{\Omega}' \mathbb{1}_{f=A} \odot \boldsymbol{\phi} \odot P + \tilde{\Omega}^{-1} \tilde{\Omega}' \mathbb{1}_{f=A} \odot \tau \mathbb{1}_{f=R \,\cap\, k \in FBA_t} \tag{35}$$

Finally, I model the marginal cost as a function of seller identity (Amazon or not), product, and weekly fixed effects.

$$c_{jft} = \beta_1 Amazon + \gamma_j + \gamma_t + \omega_{jft} \tag{36}$$

A.3 Marginal Costs FBA Sellers

In the first column of Table 14, I estimate again the marginal costs by including whether the seller is using FBA. As we can see, 3P-sellers using FBA have lower marginal costs than Amazon and 3P-sellers using FBM, while all 3P-sellers have lower marginal costs than Amazon. Since Amazon is always fulfilling the sale with FBA, its marginal costs are around 3 US$ higher than a 3P-seller using FBM and 7$ higher than a 3P-seller using FBA, on average. Instead, the marginal cost of 3P-sellers using FBM is 4$ higher than 3P-sellers using FBA. The difference between 3P-sellers can be explained by the fact that since using FBM incurs higher fixed costs, only the most efficient sellers can afford to pay for FBA.

[32] I provide here an example for illustration:

$$\Omega = \begin{pmatrix} \frac{\partial s_{11t}}{\partial p_{11t}} & \cdots & \frac{\partial s_{J1t}}{\partial p_{11t}} & \cdots & \frac{\partial s_{SJ_t}}{\partial p_{11t}} \\ \vdots & & & & \vdots \\ \frac{\partial s_{11t}}{\partial p_{JFt}} & \cdots & \frac{\partial s_{J1t}}{\partial p_{JFt}} & \cdots & \frac{\partial s_{JFt}}{\partial p_{JFt}} \end{pmatrix}$$

$$O = \begin{pmatrix} 1 & \cdots & 1 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & \cdots & 0 & \cdots & 1 \end{pmatrix}$$

$$O' = \begin{pmatrix} 0 & \cdots & 0 & \cdots & 1 \\ \vdots & & & & \vdots \\ 1 & \cdots & 1 & \cdots & 0 \end{pmatrix}$$

Another aspect we are interested in verifying is whether producers have lower marginal costs than retailers when retailers sell the same product. In the second column of Table 14, I add a dummy on whether the seller is a producer or a retailer of the product. We can see that producers incur a lower marginal cost, which corresponds to the production cost, whereas the wholesale price is higher due to double marginalization. While it can still occur that a 3P-seller using FBA competes with a producer using FBA, from Table 15 we can see that in most of the cases producers use FBA. Hence, the possibility that a 3P-seller has lower marginal costs than a producer for the same product is very low.

	(1)	(2)
Amazon	7.131^{***}	7.02^{***}
	(0.225)	(0.226)
FBA	-4.01^{***}	-3.89^{***}
	(0.145)	(0.147)
Producer		-2.78^{***}
		(0.436)
Product FE	Yes	Yes
Week FE	Yes	Yes

Table 14: Alternative marginal costs specifications. Standard errors in parenthesis.

	Producer	
FBA	0	1
0	15287	1160
1	55431	12271

Table 15: Frequency Table Producer - Fulfillment Method

A.4 Amazon Marginal Costs

In the current model, I do not compute the Amazon's marginal cost of fulfilling 3P-sellers' offers which use FBA, c_L. Hence, I am implicitly setting c_L equal to zero in the current model. To test the implications of this result, I compare the computation of Amazon's marginal costs when $c_L = 0$, like in the current model, and when $c_L = \tau$.

In Table 8, I plot the $\Delta_{jat} = c_{jta}(c_L = 0) - c_{jta}(c_L = \tau)$ for all products offered by Amazon. We can see that, when $c_L = \tau$, the estimated marginal costs are lower.

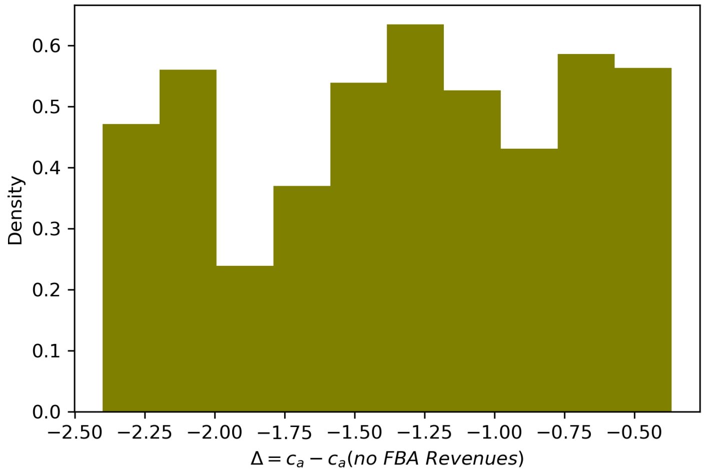


Figure 8: Difference Amazon's Marginal Costs: $\Delta_{jat} = c_{jta}(c_L = 0) - c_{jta}(c_L = \tau)$

To assess the impact of this difference on the results, I compute show the descriptive statistics of three variables: Δ to prices, corresponding to Δ_{jat}/p_{jta}; the difference in markups, $\Delta_{jta}(markups)_{jta} = markup(c_L = 0)_{jta} - markup(c_L = \tau)_{jta}$; the difference in Lerner Index, $\Delta_{jta}(LI) = LI(c_L = 0)_{jta} - LI(c_L = \tau)_{jta}$. We can see that, overall, Amazon markups and market power would be lower if $c_L = \tau$. Thus, we can claim that the higher the cost of providing FBA, the lower Amazon's markups and market power. This is because, the larger the revenue from FBA, the larger Amazon's incentive to soften competition and increase prices: while higher prices shift demand to 3P-sellers, Amazon also gains from this by getting more revenues.

However, the overall effect is not large for most products Amazon offers. Given that $0 < c_{<}\tau$, we can see that not accounting for the marginal cost of delivering the product does not have a substantial impact on the final results.

	Δ_{jat}/p_{jta}	$\Delta_{jta}(markups)$	$\Delta_{jta}(LI)$ (pp)
count	6623	6623	6623
mean	-0.06	1.33	1.48
std	0.03	0.59	1.46
min	-0.13	0.37	0.23
25%	-0.08	0.81	0.60
50%	-0.05	1.27	0.94
75%	-0.04	1.83	1.67
max	-0.02	2.40	10.73

Table 16: Ratio difference in marginal costs to price: Δ_{jat}/p_{jta}

A.5 Fixed Cost Estimation

Given that $\bar{V}$ is chosen ad-hoc, a possibility is to set $\bar{V}$ equal to the standard deviation of $\Delta\pi$. Procedure to compute the fixed costs for the different parameters. I select a random sample of 75 products and compute $\Delta^{-}\pi$ and a random sample of 25 products to compute $\Delta^{+}\pi$.

- For Amazon and FBA 3P-sellers, I select a random sample of 25 products and compute $\Delta^{-}\pi$ and $\Delta^{+}\pi$
- For FBM 3P-sellers, since there are few 3P-sellers with positive sales, I increase the sample of products to compute $\Delta^{-}\pi$ from 25 to 75

Then, given the computed $\Delta^{-}\pi$ and $\Delta^{+}\pi$, I cut the tails of the distribution in order to reduce the risk of small sample variance.

- FBM: $q_{0.3} < \Delta^{+}\pi < q_{0.7}$ and $q_{0.02} < \Delta^{-}\pi < q_{0.98}$
- FBA: $q_{0.2} < \Delta^{+}\pi < q_{0.8}$ and $q_{0.2} < \Delta^{-}\pi < q_{0.8}$
- Amz: $q_{0.2} < \Delta^{+}\pi < q_{0.8}$ and $q_{0.4} < \Delta^{-}\pi < q_{1}$

	N^{+}	Average $\Delta^{+}\pi$	SD $\Delta^{+}\pi$	N^{-}	Average $\Delta^{-}\pi$	SD $\Delta^{-}\pi$
FBM	140	30	13	118	-31	27
FBA	210	35	23	227	-52	30
Amazon	186	140	106	111	-142	148

Table 17: Statistics Truncated distributions

Mills Ratio Mills ratio for V^+. Assume $V^+ \sim N(0, \sigma_V$. I assume $\sigma_V = sd(\Delta^+\hat{\pi})$ from the truncated distribution.

$$E_{jt}[\Delta\pi^+_{jt} - \theta - V_{jt}|D_{jt} = 1] = \tag{37}$$

$$E_{jt}[\Delta\pi^+_{jt}|D_{jt} = 1] - \theta - E_{jt}[V_{jt}|D_{jt} = 1] = \tag{38}$$

$$E_{jt}[\Delta\pi^+_{jt}|D_{jt} = 1] - \theta - E_{jt}[V_{jt}|V_{jt} > \Delta\pi^+_{jt} - \theta] = \quad \text{(since } \Delta\pi^+_{jt} - \theta - V_{jt} < 0 \text{ by rev. pref.)} \tag{39}$$

$$E_{jt}[\Delta\pi^+_{jt}|D_{jt} = 1] - \theta - \sigma_V\lambda(z) \quad \text{(where } \lambda(z) = \frac{\phi(\frac{\Delta\pi^+_{jt} - \theta}{\sigma_V})}{1 - F(\frac{\Delta\pi^+_{jt} - \theta}{\sigma_V})}) = \frac{\phi(\frac{\Delta\pi^+_{jt} - \theta}{\sigma_V})}{F(\frac{-\Delta\pi^+_{jt} + \theta}{\sigma_V})}) \tag{40}$$

To compute $\lambda(z)$

1. Pick a value of θ
2. Compute $\lambda(z_{jt})$
3. Compute the average across j and t
4. Compute average across $\theta \in [\underline{\theta}, \bar{\theta}]$

Mills ratio for V^-. Assume $V^- \sim N(0, \sigma_V$. I assume $\sigma_V = sd(\Delta^-\hat{\pi})$ from the truncated distribution.

$$E_{jt}[\Delta\pi^-_{jt} + \theta + V_{jt}|D_{jt} = 1] = \tag{41}$$

$$E_{jt}[\Delta\pi^-_{jt}|D_{jt} = 1] + \theta + E_{jt}[V_{jt}|D_{jt} = 1] = \tag{42}$$

$$E_{jt}[\Delta\pi^+_{jt}|D_{jt} = 1] + \theta + E_{jt}[V_{jt}|V_{jt} < -\Delta\pi^-_{jt} - \theta] = \quad \text{(since } \Delta\pi^-_{jt} + \theta + V_{jt} < 0 \text{ by rev. pref.)} \tag{43}$$

$$E_{jt}[\Delta\pi^+_{jt}|D_{jt} = 1] - \theta - \sigma_V\lambda(z) \quad \text{(where } \lambda(z) = \frac{\phi(\frac{-\Delta\pi^-_{jt} - \theta}{\sigma_V})}{F(-\frac{\Delta\pi^-_{jt} - \theta}{\sigma_V})}) \tag{44}$$

	V^+	V^-	$(\lvert V^+\rvert + \lvert V^-\rvert)/2$
FBM	12.2	-25.3	-18.7
FBA	17.3	-24.5	-20.9
Amz	97.11	-141.4	119.3

Table 18: Caption

Table 19: Estimated Bounds for 3P-sellers Fixed Costs.

$\bar{V}$=($\lvert V^+\rvert + \lvert V^-\rvert$)/2	$\hat{\theta}$	Estimated Bounds
19 US$	$\hat{\theta}_{3P,FBM}$	[11 US$, 50 US$]
21 US$	$\hat{\theta}_{3P,FBA}$	[14 US$, 73 US$]
119 US$	$\hat{\theta}_{Amz}$	[20 US$, 261 US$]

I now compute the confidence region of $[\theta^0_L, \theta^0_U$ and the confidence region of θ^0. Following Imbens and Manski (2004), the confidence region of $[\theta^0_L, \theta^0_U]$ is

$$C_{1-\alpha}(\theta^0_L, \theta^0_U) = \left[\hat{\theta}^L - q_{1-\alpha/2}\frac{\hat{\sigma}_L}{\sqrt{N^+}}; \hat{\theta}^U + q_{1-\alpha/2}\frac{\hat{\sigma}_U}{\sqrt{N^-}}\right] \tag{45}$$

Instead, the confidence region of θ^0 is

$$C_{1-\alpha}(\theta^0) = \left[\hat{\theta}^L - q_{1-\alpha}\frac{\hat{\sigma}_L}{\sqrt{N^+}}; \hat{\theta}^U + q_{1-\alpha}\frac{\hat{\sigma}_U}{\sqrt{N^-}}\right] \tag{46}$$

Table 20: 95% Confidence Region $[\theta_L, \theta_U]$

$\hat{\theta}$	CR Bounds
$\hat{\theta}_{3P,FBM}$	[8 US$, 55 US$]
$\hat{\theta}_{3P,FBA}$	[11 US$, 77 US$]
$\hat{\theta}_{Amz}$	[5 US$, 288 US$]

Table 21: Bound for the fixed cost of offering one product during a week when using Fulfilled by Amazon (FBA) or Fulfilled by Merchants (FBM). $\alpha = 0.05$

Table 22: 95% Confidence Region θ

$\hat{\theta}$	CR Bounds
$\hat{\theta}_{3P,FBM}$	[9 US$, 54 US$]
$\hat{\theta}_{3P,FBA}$	[11 US$, 77 US$]
$\hat{\theta}_{Amz}$	[8 US$, 284 US$]

Table 23: Bound for the fixed cost of offering one product during a week when using Fulfilled by Amazon (FBA) or Fulfilled by Merchants (FBM). $\alpha = 0.05$

A.6 Comparison with Eizenberg (2014)

Assuming a bound on the error component of the fixed cost is methodologically similar to Eizenberg (2014), that assumes an upper and a lower bound on the fixed cost, as both ways allow to avoid the selection issue generated by the error component of the fixed cost. I focus on the estimation of θ, but the same calculations can be carried out for the computation of $\widetilde{\theta}$ and θ_a. I also drop the time index for simplification.

Consider the bounds on the fixed cost of product j sold by seller r (I drop the time subscript to simplify the notation).

$$F = \theta + V_{jd} \leq E_{\xi,\omega}\big[\pi(\mathcal{J}_r) - \pi(\mathcal{J}_r - \{j\})\big] = \bar{F}$$
$$F = \theta + V_{jd} \geq E_{\xi,\omega}\big[\pi(\mathcal{J}_r + \{j\}) - \pi(\mathcal{J}_r)\big] = \underline{F}$$

Eizenberg (2014) then takes two assumptions:

- Ass. Eizenberg 1: $\sup_j\{F\} = F^U < \infty, \ \inf_j\{F\} = F^L > -\infty$
- Ass. Eizenberg 2: $[F^L, F^U] \subset$ *supp(expected change in variable profit due to the elimination or addition of a single product by firm f)*. Let denote the support of the fixed costs by $[S^L, S^U]$

The identified lower and upper bound of F are then

$$L = \begin{cases} S^L & \text{if } j \in \mathcal{J}_d \\ \underline{F} & \text{if } j \notin \mathcal{J}_d \end{cases}$$

$$U = \begin{cases} \bar{F} \text{ if } j \in \mathcal{J}_d \\ S^U \text{ if } j \notin \mathcal{J}_d \end{cases}$$

These bounds apply to any potential product offered by firm f

$$L \leq F \leq U \ \forall j \tag{47}$$

Finally, the selection issue rising from V_{jd} is dealt with by taking the unconditional expectation of Eq (47)

$$E[L_{jd}] \leq E[F] \leq E[U]$$
$$\iff E[L] \leq E[+V_{jd}] \leq E[U]$$
$$\iff E[L] \leq \theta + \underbrace{E[V_{jd}]}_{=0} \leq E[U]$$
$$\iff E[L] \leq \theta \leq E[U]$$

To estimate V^L and V^U, I use $\min_j\{S\}$ and $\max_j\{S\}$ where

$$S = \begin{cases} \bar{F} \text{ if } j \in \mathcal{J}_d \\ \underline{F} \text{ if } j \notin \mathcal{J}_d \end{cases} \tag{48}$$

The estimated set is then given by

$$\bar{l} = \frac{1}{J}\sum L \tag{49}$$

$$\bar{u} = \frac{1}{J}\sum U \tag{50}$$

Finally, the confidence region is

$$[\bar{l} - \frac{sd(L)}{\sqrt{J}}q_{1-\alpha}, \bar{u} + \frac{sd(U)}{\sqrt{J}}q_{1-b}] \tag{51}$$

Since I am aggregating across firms, except for Amazon, I will compute the previous equations for FBM and FBA sellers together. In Table 25, I report the comparison with my results.

Table 24: 95% Confidence Region for 3P-sellers Fixed Costs.

$\hat{\theta}_{3P,FBM,Eiz}$	[18 US\$, 50 US\$]
$\hat{\theta}_{3P,FBA,Eiz}$	[14 US\$, 80 US\$]
$\hat{\theta}_{Amz,Eiz}$	[57 US\$, 367 US\$]

Table 25: Bound for the fixed cost of offering one product during a week when using Fulfilled by Amazon (FBA) or Fulfilled by Merchants (FBM). $\bar{V} = 5$ US\$. $\alpha = 0.05$

A.7 Combinations of Clusters of Products

During each week, I split the products sold by Amazon in four clusters:

- cluster 1: $p_{jat} <= q_{0.25}$
- cluster 2: $q_{0.25} < p_{jat} <= q_{0.5}$
- cluster 3: $q_{0.5} < p_{jat} <= q_{0.75}$
- cluster 4: $q_{0.75} <= p_{jat}$

where q_x is the x percentile in the Amazon prices distribution. In the counterfactuals, I evaluate which combinations of these four clusters the entrant 3P-seller is going to offer. In Table 26, I report the number of times a combination is chosen in equilibrium by the FBA entrant and FBM entrant.

Combinations	FBA Count	FBM Count
(cluster 1, cluster 2, cluster 4)	22	21
(cluster 1, cluster 2, cluster 3, cluster 4)	11	11
(cluster 2, cluster 3, cluster 4)	7	5
(cluster 1, cluster 2, cluster 3)	6	5
(cluster 2, cluster 4)	6	2
(cluster 1, cluster 3, cluster 4)	4	3
(cluster 1, cluster 2)	3	1
(cluster 4)	1	1
(cluster 1, cluster 4)	1	1

Table 26: **Frequency of Combinations of Clusters Offered in Equilibrium.** Comparison between FBA Entrant and FBM Entrant across weeks

B Appendix

Figures

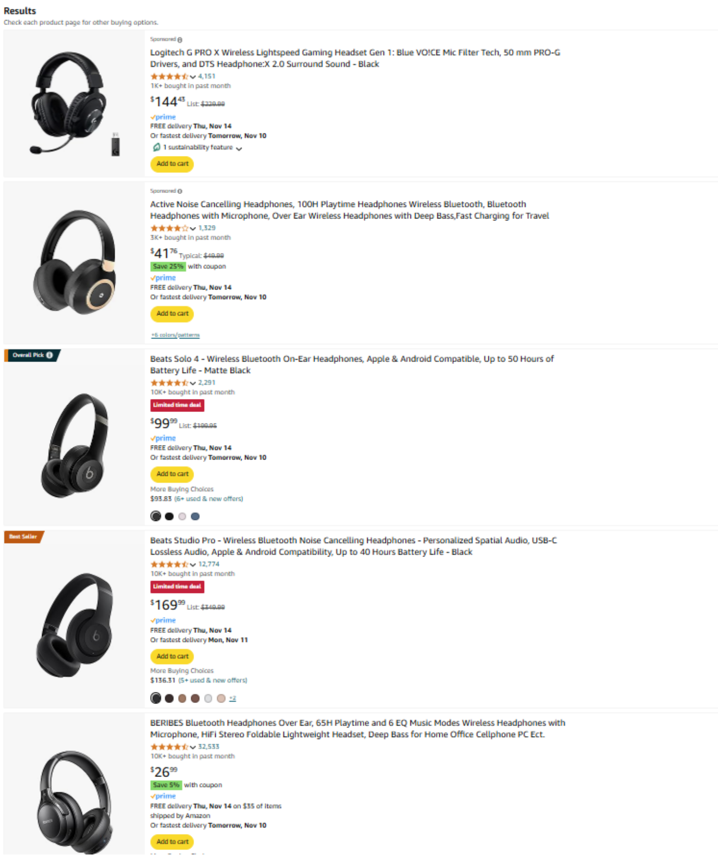


Figure 9: **List of headphones.** List displayed searching from the keyword "headphones" (9th November 2024).

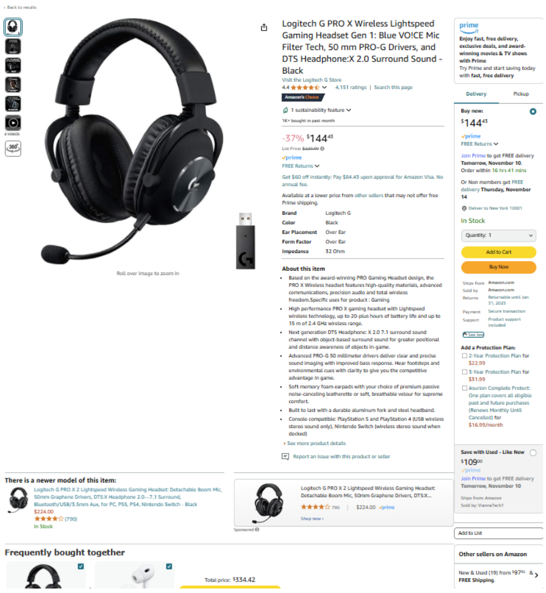


Figure 10: **Example Product** (distinguished by a distinct barcode from the other products). The Buy-Box seller is displayed in the left-hand side window. On the bottom-right, there is a smaller window to access the list of all offers available for this product.

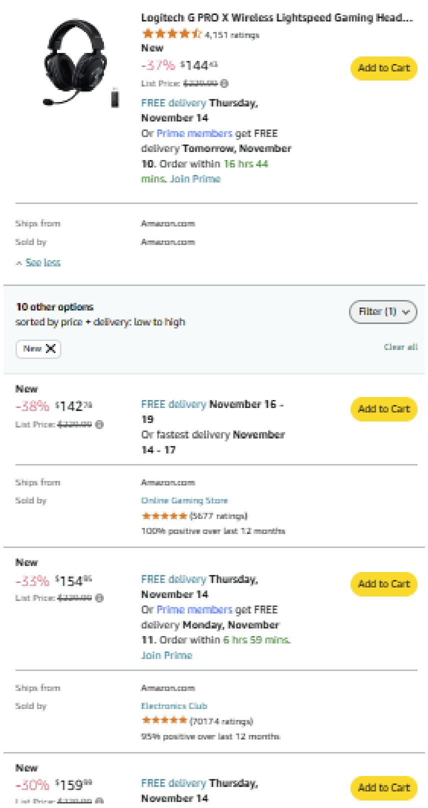


Figure 11: **List of Sellers for the Product.**

C Appendix

C.1 Incorporating the Buy-Box in the Demand Model

To understand how the Buy-Box algorithm affects sellers' pricing, I extend the demand model to include a search cost proportional to the probability of getting the Buy-Box. In other words, if seller f has a probability equal to 1 of getting the Buy-Box for product j in period t, then the search cost is null; instead, if seller f has a probability equal to 0 of getting the Buy-Box for product j in period t, then the search cost is infinite. The search cost is then decreasing in the probability of getting the Buy-Box

C.1.1 Search Model

To design the model, I use as reference the simultaneous search model for differentiated products demand in Moraga-Gonzalez et al. (2023).

Before inspecting product the offer of product j sold by seller f, consumers are not aware of the product characteristics, prices [33], ξ_{jft} and ε_{jft}.

After having inspected the product, consumer i indirect utility for product j sold by seller f is

$$u_{ijft} = \delta_{jft} + \sigma_{\varepsilon}\varepsilon_{ijft} \tag{52}$$

where ε_{ijft} follows the EV Type I distribution. The indirect utility for the outside option is

$$u_{i0t} = \varepsilon_{i0t} \tag{53}$$

Let S be the subset of products inspected by consumers. The cost of inspecting these products is

$$c_{iSt} = \sum_{jf \in S} c_{jft} + \sigma_{\lambda}\lambda_{iSt} \tag{54}$$

$$= \bar{c}_{St} + \sigma_{\lambda}\lambda_{iSt} \tag{55}$$

where λ_{iSt} is EV Type I distributed. c_{jft} can be modeled to depend on the Buy-Box, so that the higher the probability of being in the Buy-Box, the lower the search cost.

The expected utility of consumer i from inspecting all products sold by sellers in a subset S is

$$E[\max_{jf \in S} u_{ijft}] - c_{iSt} \tag{56}$$

where the expectation is taken with respect to ε_{ijft}.

Letting F_{ε} denote the CDF ε_{ijft}, the random variable $\max_{jf \in S} u_{ijft}$ has a CDF given by $\Pi_{jf \in S} F_{\varepsilon}((u - \delta_{jft})/\sigma_{\varepsilon})$. Using this, we obtain

$$m_{St} = \sigma_{\varepsilon} \log(1 + \sum_{jf \in S} \exp[\delta_{jft}/\sigma_{\varepsilon}]) - \bar{c}_{St} \tag{57}$$

Consumer i picks the subset of sellers to visit that maximizes the expected gain $m_{iSt} - \sigma_{\lambda}\lambda_{iSt}$.

The optimal search set of consumer i is then

$$S^*_{it} = \arg\max_{S \in \boldsymbol{S}} E[m_{St} - \sigma_{\lambda}\lambda_{iSt}]$$

where $\boldsymbol{S}$ is the set of all possible subsets S.

Since λ_{iSt} is iid TIEV, we can compute the probability that S^*_{it} takes a value S, which we denote P_S

$$P_S = \frac{\exp[m_{St}/\sigma_{\lambda}]}{\sum_{S' \in \bar{S}} \exp[m'_{St}/\sigma_{\lambda}]} \tag{58}$$

Once the products have been inspected, the probability she buys alternative j from seller f is

$$P_{jf|S} = \frac{\exp[\delta_{jft}/\sigma_{\varepsilon}]}{1 + \sum_{r \in S} \exp[\delta_r/\sigma_{\varepsilon}]} \tag{59}$$

[33] I assume consumers hold correct conjectures about the prices in equilibrium

The probability that consumer i buys product j sold by seller f is then

$$s_{jft} = \sum_{S \in \boldsymbol{S}_{jf}} P_S P_{jf|S} \tag{60}$$

P_S can then be rewritten as

$$P_S = \frac{\exp[m_{St}/\sigma_\lambda]}{\sum_{S' \in \boldsymbol{S}} \exp[m_{S't}/\sigma_\lambda]} \tag{61}$$

$$= \frac{\exp[\frac{\sigma_\varepsilon}{\sigma_\lambda} \log(1 + \sum_{jf \in S} \exp[\delta_{jft}/\sigma_\varepsilon]) - \bar{c}_{St}]}{1 + \sum_{S' \in \boldsymbol{S}/\varnothing} \exp[\frac{\sigma_\varepsilon}{\sigma_\lambda} \log(1 + \sum_{jf \in S} \exp[\delta_{jft}/\sigma_\varepsilon]) - \bar{c}_{St}]} \tag{62}$$

$$= \frac{(1 + \sum_{jf \in S} \exp[\delta_{jft}/\sigma_\varepsilon])^{\frac{\sigma_\varepsilon}{\sigma_\lambda}} \exp[\bar{c}_{St}]}{1 + \sum_{S' \in \boldsymbol{S}} (1 + \sum_{jf \in S} \exp[\delta_{jft}/\sigma_\varepsilon])^{\frac{\sigma_\varepsilon}{\sigma_\lambda}} \exp[\bar{c}_{St}]} \tag{63}$$

The derived market share of product j sold by seller f is then equal to

$$s_{jft} = \sum_{S \in \boldsymbol{S}_{jft}} P_S P_{jf|S} \tag{64}$$

$$= \sum_{S \in \boldsymbol{S}_{jf}} \frac{(1 + \sum_{jf \in S} \exp[\delta_{jft}/\sigma_\varepsilon])^{\frac{\sigma_\varepsilon}{\sigma_\lambda}} \exp[-\bar{c}_{St}]}{1 + \sum_{S' \in \boldsymbol{S}} (1 + \sum_{jf \in S} \exp[\delta_{jft}/\sigma_\varepsilon])^{\frac{\sigma_\varepsilon}{\sigma_\lambda}} \exp[-\bar{c}_{St}]} \cdot \frac{\exp[\delta_{jft}/\sigma_\varepsilon]}{1 + \sum_{r \in S} \exp[\delta_r/\sigma_\varepsilon]} \tag{65}$$

$$= \exp[\delta_{jft}/\sigma_\varepsilon] \sum_{S \in \boldsymbol{S}_{jf}} \frac{(1 + \sum_{jf \in S} \exp[\delta_{jft}/\sigma_\varepsilon])^{\frac{\sigma_\varepsilon}{\sigma_\lambda} - 1} \exp[-\bar{c}_{St}]}{1 + \sum_{S' \in \boldsymbol{S}} (1 + \sum_{jf \in S} \exp[\delta_{jft}/\sigma_\varepsilon])^{\frac{\sigma_\varepsilon}{\sigma_\lambda}} \exp[-\bar{c}_{St}]} \tag{66}$$

Furthermore, when $\sigma_\varepsilon = \sigma_\lambda = 1$, the market share is

$$s_{jft} = \exp[\delta_{jft}] \sum_{S \in \boldsymbol{S}_f} \frac{\exp[-\bar{c}_{St}]}{1 + \sum_{S' \in \boldsymbol{S}/\varnothing} (1 + \sum_{jf \in S'} \exp[\delta_{jft}]) \exp[-\bar{c}_{S't}]} \tag{67}$$

$$= \exp[\delta_{jft}] \frac{\sum_{S \in \boldsymbol{S}-jf} \exp[-\bar{c}_{jft} - \bar{c}_{St}]}{\sum_{S' \in \boldsymbol{S'}} (1 + \bar{\bar{\delta}}_{S't})} \exp[-\bar{c}_{S't}] \tag{68}$$

$$= \exp[\delta_{jft}] \frac{\exp[-\bar{c}_{jft}] \sum_{S \in \boldsymbol{S}-jf} \exp[-\bar{c}_{St}]}{\sum_{S' \in \boldsymbol{S}} (1 + \bar{\bar{\delta}}_{S't}) \exp[-\bar{c}_{S't}]} \tag{69}$$

$$= \exp[\delta_{jft}] \frac{\exp[-\bar{c}_{jft}] \sum_{S \in \boldsymbol{S}-jf} \Pi_{jf \in S} \exp[-\bar{c}_{jft}]}{\sum_{S' \in \boldsymbol{S}} (1 + \bar{\bar{\delta}}_{S't}) \exp[-\bar{c}_{S't}]} \tag{70}$$

Since

$$\sum_{S \in \boldsymbol{S}-jf} \Pi_{\in S} \exp[-\bar{c}] = \Pi_{g \in JFjf} (1 + \exp[-\bar{c}_g]) = \frac{\Pi_{g \in F} (1 + \exp[-\bar{c}_g])}{1 + \exp[-\bar{c}_{jft}]} \tag{71}$$

we have

$$s_{jft} = \exp[\delta_{jft}] \frac{\exp[-\bar{c}_{jft}] \frac{\Pi_{g \in F} 1 + \exp[-\bar{c}_g]}{1 + \exp[-\bar{c}_{jft}]}}{\sum_{S' \in \boldsymbol{S}} (1 + \bar{\bar{\delta}}_{S'}) \exp[-\bar{c}_{S't}]} \tag{72}$$

Then, since

$$\frac{1}{1+\exp[-\bar{c}_{jft}]} = (1+\exp[-\bar{c}_{jft}])^{-1} = \exp(\ln((1+\exp[-\bar{c}_{jft}])^{-1})) = \exp(-\ln(1+\exp[-\bar{c}_{jft}])) \tag{73}$$

we have

$$s_{jft} = \frac{\exp[\delta_{jft} - \ln(1+\exp(c_{jft}))]\Pi}{\sum_{S'\in \boldsymbol{S}}(1+\bar{\delta}_{S't})\exp[-\bar{c}_{S't}]} \tag{74}$$

Note that s_{i0} is given by

$$s_{i0t} = \exp[0]\frac{\sum_{S\in \boldsymbol{S}_0}\exp[-\bar{c}_{St}]}{\sum_{S\in \boldsymbol{S}}(1+\bar{\delta}_{S't})\exp[-\bar{c}_{St}]} \tag{75}$$

$$= \frac{1+\sum_{S\in \boldsymbol{S}/\varnothing}\exp[-\bar{c}_{St}]}{\sum_{S'\in \boldsymbol{S}}(1+\bar{\delta}_{S't})\exp[-\bar{c}_{St}]} \tag{76}$$

$$= \frac{\sum_{S\in \boldsymbol{S}}\Pi_{jf\in S}\exp[-\bar{c}_{St}]}{\sum_{S'\in \boldsymbol{S}}(1+\bar{\delta}_{S't})\exp[-\bar{c}_{St}]} \tag{77}$$

$$= \frac{\Pi_{f\in F}(1+\exp[-\bar{c}_{jft}])}{\sum_{S'\in \boldsymbol{S}}(1+\bar{\delta}_{S't})\exp[-\bar{c}_{St}]} \tag{78}$$

$$= \frac{\Pi}{\sum_{S'\in \boldsymbol{S}}(1+\bar{\delta}_{S't})\exp[-\bar{c}_{St}]} \tag{79}$$

Since $\sum_{jf=0}^{JF} s_{jft} = 1$, we have

$$\frac{\Pi + \sum_{jf=1}^{JF}\exp[\delta_{jft} - \ln(1+\exp[c_{jft}])]\Pi}{\sum_{S'\in \boldsymbol{S}}(1+\bar{\delta}_{S't}\exp[-\bar{c}_{S't}]} = 1 \tag{80}$$

$$\sum_{S'\in \bar{S}}(1+\bar{\delta}_{S't})\exp[-\bar{c}_{S't}] = \Pi\cdot(1+\sum_{jf=1}^{JF}\exp[\delta_{jft} - \ln(1+\exp[c_{jft}])]) \tag{81}$$

Therefore, we get

$$s_{jft} = \frac{\exp[\delta_{jft} - \ln(1+\exp[\bar{c}_{jft}])]}{1+\sum_{r=1}^{JF}\exp[\delta_r - \ln(1+\exp[\bar{c}_r])]} \tag{82}$$

$$s_{0t} = \frac{1}{1+\sum_{r=1}^{JF}\exp[\delta_r - \ln(1+\exp[\bar{c}_r])]} \tag{83}$$

C.1.2 Buy-Box Probability

Following a methodology similar to Lee and Musolff (2023), I approximate the Buy-Box algorithm into a discrete choice model.

Consider a fraction of time ψ between $t-1$ and t. The Buy-Box valuation of product j for seller f in period ψ is

$$v_{jf,\psi} = \alpha^{BB}p_{jft} + \boldsymbol{\beta}^{BB}\boldsymbol{X}_{jft} + \xi_{jBBf} + \varepsilon_{jBBf,\psi}$$

The valuation when no Buy-Box is provided is

$$v_{j0} = \varepsilon_{j0}$$

Assuming that ε_{jBBs} is distributed EV Type I, the **probability** of seller f being chosen in the **Buy-Box** is

$$Pr_{BB,jft} = \frac{\exp(\delta_{jft}^{BB} + \xi_{jBBs})}{1 + \sum_{sj=1}^{Sj} \exp(\delta_{jft}^{BB} + \xi_{jBBs})} \tag{84}$$

I provide here some very preliminary estimates using the standard market share inversion. I estimate the parameters of the model by 2SLS, using as instrument the number of available sellers for a product.

Variables	**Estimates (Standard Error)**
Prices	-0.0013 (0.001)
Shipping Cost	0.0095 (0.014)
Amazon	0.5078^{***} (0.075)
Fulfilled by Amazon	0.0986^{*} (0.053)
Product Fixed Effects	Yes

Table 27: Weekly Buy-Box Estimates. *$p < 0.05$, **$p < 0.025$, ***$p < 0.01$

C.1.3 Market Share Inversion

c_{jft} could be modeled as

$$c_{jft} = \ln(\exp[\omega \cdot (1 - \hat{Pr}_{BB,jft})]) \tag{85}$$

where $\hat{Pr}_{BB,jft}$ is the estimated probability of getting the Buy-Box (from Eq. (84)) and ω is a parameter to be estimated.

Using the market share inversion, we get the following linear equation to estimate

$$\ln(s_{jft}) - \ln(s_{0t}) = \delta_{jft} - \ln(1 + \exp[\hat{\tilde{c}}_{jft}]) \tag{86}$$

$$\iff \ln(s_{jft}) - \ln(s_{0t}) = \boldsymbol{\beta X_{jft}} - \alpha p_{jft} + \gamma_{month} + \gamma_j + \xi_{jft} - \omega \cdot (1 - \hat{Pr}_{BB,jft}) \tag{87}$$

where $(1 - \hat{Pr}_{BB,jft})$ can be interpreted as a search cost shifter.