CENTRE DE RECHERCHES EN ÉCONOMIE ET DROIT
Paris Center for Law and Economics

UNIVERSITÉ PARIS II
PANTHÉON-ASSAS

CRED WORKING PAPER N^o 2022-4

Bye-box: An Analysis of Non-Promotion on the Amazon Marketplace

03.06.2022

Matthias Hunold[*], Ulrich Laitenberger[†] and Guillaume Thébaudin[‡]

[*]University of Siegen, Unteres Schloß 3, 57068 Siegen, Germany; e-mail: matthias.hunold@uni-siegen.de

[†]i3, Telecom Paris, Institut Polytechnique de Paris, 91120 Palaiseau, France and CRED - Université Paris II; e-mail: ulrich.laitenberger@telecom-paris.fr.

[‡]i3, Telecom Paris, Institut Polytechnique de Paris, 91120 Palaiseau, France; e-mail: guillaume.thebaudin@telecom-paris.fr.

Bye-box: An Analysis of Non-Promotion on the Amazon Marketplace*

Matthias Hunold†, Ulrich Laitenberger‡, and Guillaume Thébaudin§

June 3, 2022
First version: January 31, 2021

Abstract

We study seller and product recommendations of the hybrid e-commerce platform Amazon. Using web-scraped data, we find that Amazon makes the visibility of offers of third-party suppliers in the "buybox" dependent on prices on competing marketplaces like Walmart and eBay. Amazon's own offers are visible regardless of their competitiveness. We find that the absence of seller recommendations makes recommendations to related products more effective and Amazon tends to steer consumers in these situations more often to products it sells itself. We discuss that this behavior is difficult to reconcile with the hypothesis of an independent marketplace operator.

Keywords: Amazon marketplace, buybox, self-preferencing, algorithm bias, recommendation algorithms.

JEL Class: D40, L42, L81

*We thank participants at the French-German E-Commerce Workshop 2021, the ICT Conference at the ZEW 2021, the EARIE annual meeting 2021, the AFREN doctoral meeting 2021, the MaCCI Annual Meeting 2022 as well as seminar participants in the Dyson School of Management Strategy Seminar, the University of Cologne - Supply Chain Research Seminar, the Centre of Competition Policy seminar in Norwich, economics seminar at University of Siegen, the Economics and business seminar in University of Luxembourg, the Economic Workshop seminar in University of Würzburg, the marketing seminar at Goethe University Frankfurt and the MSI seminar at KU Leuven, and specifically Özlem Bedre-Folie, Marc Bourreau, Francisco Clavorà Braulin, Tomaso Duso, Sean Ennis, Chris Forman, Daniel Garcia, Dominik Gutt, Paul Heidhues, Reinhold Kesler, Benjamin Leyden, Johannes Muthers, Carlo Reggiani, Jochen Reiner, Michelangelo Rossi, Armin Schmutzler, Bernd Skiera and Sebastian Wismer for helpful comments. This work is supported by the French Research Agency (ANR) under grant ANR-21-CE26-0018-01 ("HYBRID"). The authors alone are responsible for the content.

†University of Siegen, Unteres Schloß 3, 57068 Siegen, Germany; e-mail: matthias.hunold@uni-siegen.de

‡i3, Telecom Paris, Institut Polytechnique de Paris, 91120 Palaiseau, France; e-mail: ulrich.laitenberger@telecom-paris.fr.

§i3, Telecom Paris, Institut Polytechnique de Paris, 91120 Palaiseau, France; e-mail: guillaume.thebaudin@telecom-paris.fr.

1 Introduction

Online marketplaces enable consumers to find different product offers from various sellers. Marketplace operators use complex algorithms that determine which sellers and products the consumers see on the marketplace websites. These “recommendations” can benefit consumers by facilitating their search. It seems natural that these recommendations also take the platform’s interests into account and not only those of the consumers. Hybrid sales platforms, which operate a marketplace and also sell on it as retailers, are under scrutiny as there is the fear that the interests of consumers and the hybrid platforms might not be well aligned. Prominent cases in point concern Google Shopping[1] and Amazon’s hybrid role as both a marketplace and a seller.[2] New legislative proposals, such as the European Unions’s Digital Markets Act (DMA) and the American Innovation and Choice Online Act, aim at restricting self-preferencing more generally, at least for particularly important online platforms.[3] An economic experts’ report on the Digital Markets Act even proposes making self-preferencing per-se illegal for gatekeepers (Cabral et al., 2021).

Whereas policy can build on various published theories on steering behavior, the corresponding empirical evidence is still comparatively scarce (see Section 2). In this article, we contribute by studying key design features of the Amazon website to better understand the functioning of the marketplace with respect to customer steering and the interplay between the marketplace operations and the retail business. Amazon is an interesting case of a hybrid sales platform that provides both seller recommendations for a given product as well as product recommendations.

The Amazon website facilitates the purchasing process by recommending products and a particular seller for a given product. The promotion of one seller among those available on a product page in the so-called buybox is a key design feature which significantly influences sales (Hagiu et al., Forthcoming). Unsurprisingly, different studies investigate whether Amazon uses the buybox for self-preferencing (Chen et al., 2016; Gómez-Losada and Duch-Brown, 2019; Lee and Musolff, 2021).

[1] See European Commission –Case AT.39740 –Google Search (Shopping), 2017, https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1516198535804&uri=CELEX:52018XC0112(01), last accessed June 3, 2022.

[2] See European Commission –Antitrust: Commission sends Statement of Objections to Amazon for the use of non-public independent seller data and opens a second investigation into its e-commerce business practices, 2020, https://ec.europa.eu/commission/presscorner/detail/en/ip_20_2077, last accessed June 3, 2022.

[3] See "The Digital Markets Act: ensuring fair and open digital markets", https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/digital-markets-act-ensuring-fair-and-open-digital-markets_en and "S.2992 - American Innovation and Choice Online Act", https://www.congress.gov/bill/117th-congress/senate-bill/2992/text, last accessed June 3, 2022.

We investigate the fact that Amazon does not always display the buybox, even in cases where multiple sellers are available. Amazon claims it does so in the interest of consumers in certain situations, in particular when it observes harmful prices. This behaviour is interesting as, first, it is not clear whether Amazon applies the same standards in instances where it is one of a product's sellers, and, second, it tends to reduce the product's sales and thus potentially Amazon's commission income. However, sales of other products may increase if Amazon redirects consumers to them. Consumers who search for a given product but do not see its buybox might be more attentive to Amazon's recommendations of alternative products. A hybrid sales platform might use this channel as a means to favor products which it sells itself.

Based on a description of the practises in question (Section 3), we develop our hypotheses in Section 4 and describe the data used for our study in Section 5. Through webscraping between September to December 2020, we collected daily information for about 3500 products about the positioning of sellers and related product recommendations from the US version of Amazon. Our sample consists of the the 50 most popular products in 50 categories from Amazon as well as from the e-commerce platform of Walmart. We also retrieved prices for the competing sales platforms eBay and Walmart.

In Section 6, we conduct empirical tests to validate the hypotheses on the buybox suppression. Focusing on cases where the product is in stock, we observe that the buybox is always visible when Amazon is one of the sellers, while it is invisible in 39% of cases when only third-party sellers are available. Using regression analyses, we find that Amazon is more likely to show the buybox when sellers with a certain quality are present and when the minimum price is lower than competitive benchmarks. Using out of sample predictions, we find that if Amazon would apply the same algorithm to cases where it is a seller itself, it should not show the buybox in about 13% of cases. Moreover, we quantify the effects of buybox visibility on sales. Our regression analyses reveal that the sales rank of a product, and thus the sales of the product on the Amazon marketplace, are significantly lower when the product's buybox is invisible. This underlines the importance of the buybox and the power of suppressing the buybox as a tool for steering demand.

In Section 7, we assess the effects of product recommendations on sales. For this, we focus on the "Compare with similar items" design element which typically appears on a product page and provides organic recommendations to other products. We find that these recommendations are more effective in terms of sales in the absence of the buybox. Moreover, in the absence of the buybox, there are more frequently recommendations of alternative products which are sold by Amazon. We exclude various explanations whereby these patterns could be rationalized under the hypothesis of an independent marketplace operator.

We conclude in Section 8 by summarizing our results and their limitations. We also discuss managerial implications and relate our results to various policy debates, such as transparency, consumer protection, price parity clauses, as well as the self-preferencing of hybrid sales platforms.

2 Related Literature

We relate to the empirical literature on product rankings and pricing on sales platforms. One strand of this literature examines how the arrangement of search results influences consumer choice on sales platforms (Chen and Yao, 2016; De los Santos and Koulayev, 2017; Koulayev, 2014; Ghose et al., 2012, 2014; Ursu, 2018). Ursu (2018) exploits a random variation in the ranking of hotels at the online travel agent Expedia and finds that consumers are more likely to click on an offer that is higher ranked to obtain detailed information on it.

Albeit growing, the empirical literature on the ranking biases of online platforms is limited (Krämer and Schnurr, 2018). Reasons for this include the difficulty with identifying causality in the absence of controlled experiments to study these research questions. Hunold et al. (2020) investigate whether hotel booking portals (OTAs) assign lower positions to hotels in their search results if they charge lower hotel prices at other OTAs or in direct online sales. In the present article, we also find empirical relationships between price ratios to other distribution channels and the visibility of the buybox on the Amazon website. This is also interesting against the backdrop that Amazon has used price parity clauses for many years[4]. The paper by Chen and Tsai (2019) studies the behavior of Amazon regarding its "Frequently Bought Together" recommendation tool. They find that, at a given point in time, Amazon is less likely to recommend such a complementary product when the product is offered only by third-party but not Amazon. Our study differs by focusing on and relating buybox suppression and the "compare with similar items" table. We thus employ a qualitatively different identification strategy regarding self-preferencing and study different parts of the Amazon marketplace. Lee and Musolff (2021) study the Amazon buybox using a structural model. For a given product, they

[4] See "Amazon eases price restrictions on third-party vendors" in Financial Times, March 12 2019, https://www.ft.com/content/3beea4a6-445b-11e9-b168-96a37d002cd3, last accessed June 3, 2022.

find that self-preferencing Amazon as a seller in the buybox increases static consumer welfare, mainly through increased price competition and lower prices. They also find that this lowers entry incentives and, thereby, offsets nearly all of the algorithm's gains in consumer welfare relative to the random recommendation baseline. Our study differs by relating buybox suppression and the "compare with similar items" table to identify self-preferencing regarding product recommendations.

We also relate to other articles studying the Amazon marketplace (Zhu and Liu, 2018; Chen et al., 2016; Gómez-Losada and Duch-Brown, 2019; Reimers and Waldfogel, 2021). Chen et al. (2016) developed a methodology to detect algorithmic pricing in the Amazon marketplace. In this work from the field of information systems, there is, however, no economic analysis. Zhu and Liu (2018) show that Amazon is more likely to target successful product spaces when choosing which products to sell itself. Reimers and Waldfogel (2021) study the effects of Amazon star ratings on consumer welfare in book publishing on sales. We build on their approach to estimate demand effects.

By studying how Amazon steers consumers with the buybox and the related products' table, we relate to the theoretical literature on consumer steering by intermediaries (Raskovich, 2007; Inderst and Ottaviani, 2012; Hagiu and Jullien, 2011, 2014; De Cornière and Taylor, 2014, 2019; Hunold and Muthers, 2017; Shen and Wright, 2019). Hagiu and Jullien (2011) analyze distortions in search engine results lists. In an environment in which customers have heterogeneous search costs and the platform generates sales via per-click payments, Hagiu and Jullien (2011) predict distortions in the result lists in the sense that a less suitable product is displayed first in order to generate additional income from the product providers. Calvano and Jullien (2018) show that biases in the recommendations of such algorithms can also occur if the recommender has no financial incentives regarding which product the consumer consumes, as recommendation systems can be inefficiently risk-averse. We complement this literature by demonstrating empirical patterns of steering in terms of product and seller recommendations on the one hand and their effects on demand on the other hand.

We study Amazon's recommendations in view of its hybrid role and thereby add to the literature on steering incentives in the case of vertical integration (Bourreau and Gaudin, 2022; De Cornière and Taylor, 2014, 2019; Drugov and Jeon, 2019; Hagiu et al., Forthcoming). De Cornière and Taylor (2014) show that integrated search engines distort search results, but the welfare effect is unclear, as the integrated search engine can have a strong positive incentive to generate demand. Similarly, De Cornière and Taylor (2019) examine biased recommendations from merchants and show that bias can harm consumers when the payout functions of sellers and consumers conflict over the optimal recommendation. Hagiu et al. (Forthcoming) analyze the dual role of an intermediary that is both seller and marketplace. They contrast a prohibition of this practice with other policies that restrict the imitation of innovative third-party products through the platform or the redirection of demand for the intermediary's product. Using the example of a streaming platform, Bourreau and Gaudin (2022) and Drugov and Jeon (2019) study incentives to bias recommendations to consumers toward vertically integrated content. We complement this literature by providing empirical evidence that is consistent with self-preferencing.

3 Background knowledge on Amazon's marketplace

3.1 Business model

Amazon is the largest e-commerce platform in the United States and in Europe. In the United States, over 1.1 million sellers were active on the platform in 2019, offering a total of 12 million products.[5] Amazon accounts for a share of 38.7% of the US e-commerce market in 2020, followed by the marketplaces of Walmart (5.3%) and eBay (4.7%).[6]

Amazon generates revenue mainly through the own sale of products on its platform. In 2019, its online stores contributed US $141.25 billion to Amazon's global net revenues. Commission fees charged to third-party sellers were the second largest revenue stream with $53.76 billion in 2019.[7] The share of units sold by a third-party has been steadily growing, and since 2017 has accounted for over 50% of all products sold.[8] Amazon invites third-party sellers to participate in the *Fulfilled-by-Amazon* (FBA) program on its platform. For FBA sellers, Amazon handles the logistics, like inventory, shipping, and returns, in exchange for higher commission payments. Sellers can also choose to take responsability for the logistics of the product they sell on the platform. These are called *Fulfilled-by-Merchants* (FBM) sellers and they only pay commission payments to Amazon for the intermediation. Finally, Amazon offers consumers a

[5] See `https://www.statista.com/statistics/1086664/amazon-3p-seller-by-country/`, last accessed June 3, 2022.

[6] See `https://www.marketingcharts.com/industries/retail-and-e-commerce-112285`, last accessed June 3, 2022.

[7] See `https://www.statista.com/statistics/672747/amazons-consolidated-net-revenue-by-segment/`, last accessed June 3, 2022.

[8] See `https://www.statista.com/statistics/259782/third-party-seller-share-of-amazon-platform`, last accessed June 3, 2022.

loyalty program called *Prime*, which entitles consumers to fast shipping over a wide range of products –typically those offered by Amazon itself or FBA products –as well as unlimited access to digital streams, in exchange for a monthly fee.

3.2 Competition policy cases and legislation

Due to its strong market position, some business practices have raised concerns. For a long time, Amazon had a price parity policy which prohibited third-party sellers from selling elsewhere on the web at a lower price. Amazon ended this practice in 2013 in Europe following anti-trust investigations, and recently ended it in the US in 2019 under threat of regulation. The European Commission recently launched an investigation regarding the "dual" mode of Amazon, of being both a marketplace operator and seller on it. In particular, the European Commission is concerned that Amazon's ability and practice of collecting transaction data on its marketplace and its use of it as a retailer, as well as the way in which it steers consumers through its platform design, could be anti-competitive.[9]

The current proposal of the Digital Markets Act (DMA) for the EU contains various rules that so-called gatekeepers need to follow. These include rules against parity clauses and self-preferencing.[10] The DMA defines criteria for how a firm would be categorized as a gatekeeper and it seems that most commentators expect Amazon will be one of them.[11]

3.3 Seller recommendation: The buybox

Seller aggregation on Amazon. A key feature of e-commerce marketplaces is that multiple sellers can offer the same product to consumers. Sellers may differ in certain dimensions, such as the quality of service. A platform can influence how easily consumers can compare the sellers and their offers through the design of its website. The design can also influence the degree of competition between the sellers.

The design of the Amazon platform is such that when consumers search for a product through the Amazon search box, the search results link to a unique, product-level web-page with information on the product characteristics. The list of the different offers available for each product is accessible on a separate page with information on every offer, such as price, shipping details,

[9] See https://ec.europa.eu/commission/presscorner/detail/en/ip_20_2077, last accessed June 3, 2022.

[10] See Article 5b and Article 6(1)d of the DMA proposal https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52020PC0842, last accessed June 3, 2022.

[11] See "Which platforms will be caught by the Digital Markets Act? The 'gatekeeper' dilemma", available at https://www.bruegel.org/2021/12/which-platforms-will-be-caught-by-the-digital-markets-act-the-gatekeeper-dilemma/, last accessed June 3, 2022.

and seller metrics. Consumers can access this page through a link on the product page. Other platforms like Walmart and Cdiscount have similar designs.

Instead, platforms like eBay or Alibaba do not aggregate offers for the same product by different sellers on one page. They allow sellers to customize their offer on an own subpage, where they can decide on the details to provide. Consumers are thus likely to be faced with non-standardized websites when comparing sellers.

Figure 1: Product page with buybox at Amazon.com

Basic functionality of the buybox. On the Amazon marketplace, a product page does not (directly) show the list of sellers offering the product. Rather, for a majority of products Amazon usually selects one unique offer among the different ones available to be featured prominently in the top-right section of the product detail page –the so-called *buybox*. This can be seen on the left-hand side of Figure 1. In this web site element, consumers see the price as well as the different characteristics of the selected offer (shipping details, seller metrics...) and can click on the “add-to-cart” button, also present in this web site element, if they wish to purchase the product from the promoted seller. Importantly, the price of the offer featured in the buybox is also shown as the product’s attributes on the search result page for cross-product comparison purposes.

The link to the web site listing all available offers is located right under the buybox. It is thus not prominently displayed on the product page. Given this design, consumers may be more prone to directly click on the “add-to-cart” button on the buybox element rather than engaging in a comparison process of all the different offers available. It is reported that over 80% of sales are made through the buybox (Hagiu et al., Forthcoming), the remaining one being made on the offer comparison page. Thus, given the prominence and thereby the demand which it offers, being featured in the buybox appears to be crucial for sellers aiming at higher sales.

Buybox algorithm. Amazon rotates prominence among the different eligible sellers. The algorithm behind seller promotion in the buybox is proprietary and the exact formula behind the buybox allocation has not been released. Amazon, however, advises sellers on its platform with some conditions that they need to fulfill in order to be eligible for buybox prominence. These are: (i) competitive pricing, (ii) Prime-eligibility or free and fast shipping, (iii) provide consumers with great experience (measured through past metrics regarding one-time shipping, product defects...), and (iv) having the product in stock.[12]

Thus, the buybox allocation algorithm can be seen as a mapping from all the characteristics of each offer available to the share of buybox each offer gets. Chen et al. (2016) estimate the importance of several seller characteristics and show that competitive pricing is one of the most important determinants for buybox allocation, followed by the amount of positive feedback sellers receive from consumers who have purchased from them. The study also indicates that Amazon appears to have a high tendency to promote its own offer for products it sells.

Amazon's claim about its prominence allocation rule leads sellers to compete on the relevant dimensions in order to be featured in the buybox, as their sales are likely to be proportional to their buybox share. Amazon's aim thus appears to ensure a high degree of competitive intensity at the product level on its platform.

3.4 Recommendations of other products

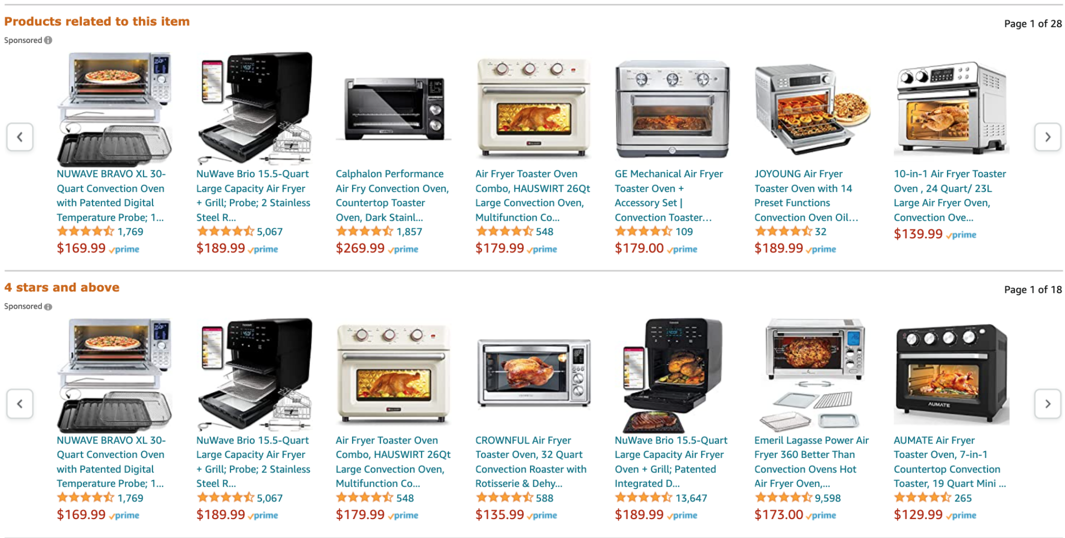
Amazon recommends other products when consumers have already arrived at a specific product page. Amazon makes these recommendation by matching past purchases and ratings to similar items, and combining these similar products to a list of recommended products. The technique is called item-based collaborative filtering.

Sponsored sections. Third-party sellers can pay to have their products advertised on the product pages of other similar products. These are displayed horizontally in sections on product pages under names such as "Products related to this item" or "4 stars and above" and labeled as "sponsored", as can be seen in Figure 2. Only basic information is provided about displayed

[12] See `https://sellercentral-europe.amazon.com/gp/help/external/201687550?language=en_GB&ref=efph_201687550_cont_GGLVN6Y9PZ6W4NSD`, last accessed June 3, 2022.

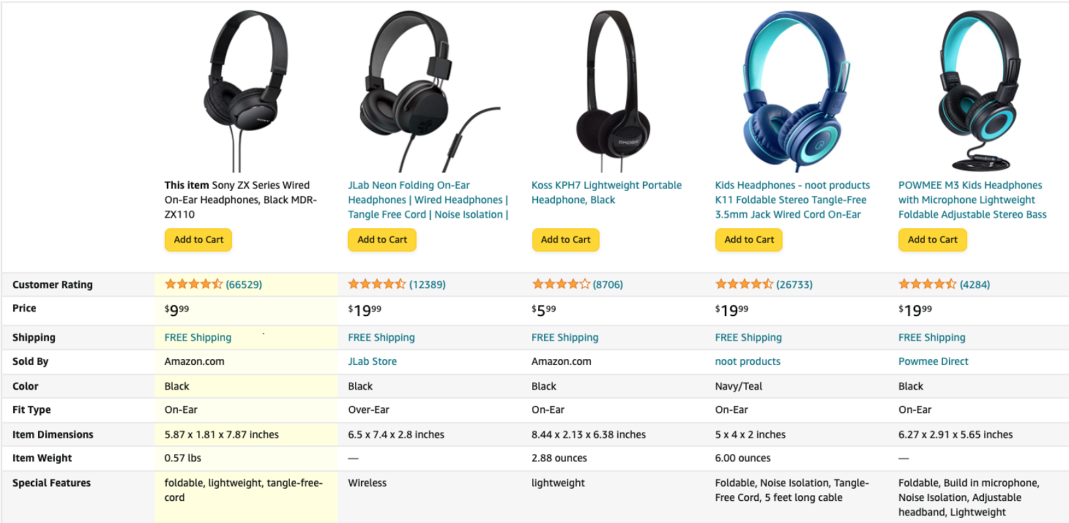
products: name, price, product ratings, and a picture. If consumers are interested in one of these products, they can click on the links to be referred to the product page of the product.

Figure 2: Sponsored product recommendations –section on product pages



Non-sponsored sections. Amazon provides product recommendations which are not influenced by third-party sellers' monetary payments in two elements of the product page. The most important one is the *Compare with similar items* section (which is available in our data for 80% of products). Consumers can easily compare the product they were considering in the first place with up to five other similar products in terms of technical characteristics as well as price, seller identity, and ratings. It is located under the product information on the web-page but directly accessible from the top of the page through a link next to the buybox section. Amazon facilitates the purchase of products displayed by showing an"add to cart" button directly under the product pictures, as can be seen in figure 3. One can guess that the functionality of this element on the web-page is to enable consumers to easily compare and navigate across available alternatives on the marketplace if consumers are not fully satisfied with the characteristics of the product they were initially considering. Another prominent section is called *Frequently bought together*. This section provides consumers with complementary products to the one they are considering buying. These can, for instance, be case covers if consumers are about to purchase a tablet.

Figure 3: "Compare with similar items" table on product pages



	This item Sony ZX Series Wired On-Ear Headphones, Black MDR-ZX110	JLab Neon Folding On-Ear Headphones | Wired Headphones | Tangle Free Cord | Noise Isolation |	Koss KPH7 Lightweight Portable Headphone, Black	Kids Headphones - noot products K11 Foldable Stereo Tangle-Free 3.5mm Jack Wired Cord On-Ear	POWMEE M3 Kids Headphones with Microphone Lightweight Foldable Adjustable Stereo Bass
	Add to Cart	Add to Cart	Add to Cart	Add to Cart	Add to Cart
Customer Rating	(66529)	(12389)	(8706)	(26733)	(4284)
Price	$9.99	$19.99	$5.99	$19.99	$19.99
Shipping	FREE Shipping	FREE Shipping	FREE Shipping	FREE Shipping	FREE Shipping
Sold By	Amazon.com	JLab Store	Amazon.com	noot products	Powmee Direct
Color	Black	Black	Black	Navy/Teal	Black
Fit Type	On-Ear	Over-Ear	On-Ear	On-Ear	On-Ear
Item Dimensions	5.87 x 1.81 x 7.87 inches	6.5 x 7.4 x 2.8 inches	8.44 x 2.13 x 6.38 inches	5 x 4 x 2 inches	6.27 x 2.91 x 5.65 inches
Item Weight	0.57 lbs	—	2.88 ounces	6.00 ounces	—
Special Features	foldable, lightweight, tangle-free-cord	Wireless	lightweight	Foldable, Noise Isolation, Tangle-Free Cord, 5 feet long cable	Foldable, Build in microphone, Noise Isolation, Adjustable headband, Lightweight

4 Hypotheses

4.1 Reasons for buybox suppression

We observe that Amazon does not always display the buybox even when the product is available (Figure 4). This means that no offers are present directly on the product page and the "add-to-cart" and "buy now" buttons are missing as well. We refer to this as the buybox being "invisible" or "suppressed."

Figure 4: Product page with suppressed buybox

Buybox suppression is puzzling at first glance as the non-promotion of any seller bears the risk that consumers might not purchase the good without recommendation. Thereby, Amazon potentially foregoes commission payments.

Amazon's explanation. Amazon claims that it acts in the best interest of consumers when suppressing the buybox, as reflected by these statements on its website:

> "Amazon regularly monitors the prices of items on our marketplace and compares them with other prices available to our customers. If we see pricing practices on a marketplace offer that harms customer trust, Amazon can remove the buy-box [...].
>
> Pricing practices that harm customer trust include [...] setting a price on a product or service that is significantly higher than recent prices offered on or off Amazon."[13]

Moreover, there are commentators who claim that Amazon's non-promotion is likely to be related to competing sales channels being more competitive than the sellers on its marketplace.[14] In line with these claims, a product's buybox is more likely to be invisible when the best price for the product on the Amazon marketplace is relatively high.

Hypothesis 1 (H1)**.** *The buybox is more likely to be invisible when the best price available on the Amazon marketplace for the product is high relative to competitive benchmarks.*

These competitive benchmarks are defined more precisely below.

If Amazon's claim is true, we would expect that, when Amazon itself is a seller of a product, it will offer the product at a competitive price that does not "harm customer trust". By this, making the buybox invisible is not necessary even if applying the same standards as in cases where Amazon is not one of the sellers. This leads to two hypotheses.

Hypothesis 2a (H2a)**.** *Amazon applies the same competitive criteria for making a product's buybox on its marketplace invisible in cases where it is itself a seller of the produce as in cases where only third parties sell the product.*

Hypothesis 2b (H2b)**.** *A product's buybox is always visible when Amazon is selling the product.*

---

13 See https://sellercentral.amazon.com/gp/help/external/G5TUVJKZHUVMN77V?language=en_US, last accessed June 3, 2022.

14 "By far the top reason that Amazon suppresses the buybox is that they think all of the offers from all of the sellers for the item are priced too high. Amazon likes for customers to have the perception of buying items at the best price available.", see https://www.fulltimefba.com/understanding-the-suppressed-buy-box-on-amazon, last accessed June 3, 2022; another quote is "Removing the Buy Box when a lower-priced product is found off Amazon makes it less likely a customer will purchase it on Amazon. A missing Buy Box in this case is an effort by Amazon to protect the customer (and Amazon's brand) from a negative experience where Amazon isn't home to the lowest price.", See https://www.buyboxexperts.com/12-reasons-amazon-listing-missing-buy-box-how-get-back, last accessed June 3, 2022.

4.2 Effects of buybox invisibility on demand

Typically, the buybox is present and consumers use it when buying the product on the Amazon marketplace. The buybox channels over 80% of the sales (Hagiu et al., Forthcoming). If consumers wish to purchase the product on the marketplace but the buybox is invisible, they can click on a "see all buying options" button located in the same top-right corner where the buybox is usually shown. They can then choose one of the available offers.

How do consumers react when there is no buybox on the product page? Some consumers might be confused and think that the product is currently not available for purchase, given missing price and shipping details in the typical position on the product page. Other consumers might search for alternative products on Amazon or turn toward other sales platforms that might also sell the product. We thus expect that a significant share of consumers will no longer buy the product on the Amazon marketplace when the buybox is missing.

Hypothesis 3 (H3)**.** *The sales of a product on the Amazon marketplace are lower when its buybox is invisible.*

4.3 Effects of a product's presence in a comparison table on its sales

The table titled "Compare with similar items" on the product page allows consumers to compare the product with substitutes (see Figure 3). We refer to this table as the "comparison table." The comparison table can stimulate the consumer to buy one of the listed substitutes instead of the product on the current page. The next hypothesis is thus straightforward.

Hypothesis 4 (H4)**.** *The sales of a product on the Amazon marketplace are higher when it is listed in the comparison table of another product's page.*

There might also be a link between the visibility of the buybox on a product page and the frequency with which consumers look at the comparison table on that page and end up buying a substitute.

Hypothesis 5 (H5)**.** *Product j being present in the comparison table on product i's page increases demand of product j more if the buybox of product i is invisible.*

4.4 Vertical independence, comparison table, and buybox invisibility

The comparison table on a particular product page typically contains three to five substitutes. In the comparison table on a particular product page, we observe that different products are present at different points in time. We are interested in the selection of products in the comparison table. Fundamentally, we want to study whether the listing decisions are consistent with

Hypothesis 6a (H6a)**.** *The Amazon marketplace acts like an operator that is independent of the Amazon sales department.*

We want to test this against the following hypothesis of self-preferencing:

Hypothesis 6b (H6b)**.** *The Amazon marketplace acts like an integrated operator that favors its own sales department.*

First of all, we investigate whether a product has a higher probability of appearing in the comparison table if Amazon is one of its sellers.

Hypothesis 7 (H7)**.** *Product j appears in the comparison table of a product i more often if Amazon is one of the sellers of product j.*

We acknowledge that results consistent with H7 may be consistent with the hypothesis 6a of an independent platform operator. It might be that consumers value the fact that Amazon sells a product, such that also an independent marketplace operator might prioritize products sold by Amazon.

To distinguish between H6a and H6b, an investigation of the relationship between buybox suppression on the page of product i and the likelihood that a substitute product j that is sold by Amazon appears in the comparison table appears to be promising. Let us explain why.

A platform operator that internalizes the profits of the integrated retail business generally has incentive to increase their profits if that does not decrease the operator profits by too much. However, we observe that Amazon does not always prioritize products it sells in the comparison table. Indeed, it is plausible that it is of value for an integrated platform operator to not appear too "biased" in the eyes of independent sellers, consumers, and possibly authorities.

Even if some degree of self-preferencing occurred, it would thus not be drastic, but rather arising from a trade-off between pushing own sales and appearing unbiased (independent). As a consequence of this trade-off, the self-preferencing would rather occur in instances where it is more profitable but not too obvious.

Suppose that H5 is true: product j being present in the comparison table on product i's page increases the demand of product j more if the buybox of product i is invisible. In this case, listing a product sold by Amazon in the comparison table is thus more profitable for Amazon if the buybox is invisible. A platform operator that maximizes joint profits should thus more often list its own products in the comparison table when the buybox is invisible. Instead, under the hypothesis of an independent marketplace operator, the fact that the buybox of product i is visible or not should not affect the selection of substitute offers in the comparison table. We will therefore test the following hypothesis:

Hypothesis 8 (H8)**.** *A product j sold by Amazon appears more often in the comparison table on a product page when the buybox on that product page is invisible.*

The behavior described in H8 is difficult to rationalize under the null hypothesis H6a of an independent marketplace operator. However, it might be consistent with gradual self-preferencing as described above and thus with H6b.

5 Data

5.1 Data collection

Our main source are daily web-scraped data from the US Amazon product pages. We complement the data with price information from the online marketplaces of Walmart and eBay. Walmart is the biggest retail company in the US and opened its e-commerce website to third-party sellers in 2009. It is currently the second largest e-retailer in the US, behind the Amazon marketplace.The eBay marketplace has been operating since 1995, and is currently the third largest e-retailer in the US. We obtained Walmart data through web-scraping Walmart product pages and obtained eBay data through a third party data supplier (Keepa).

We selected products in a symmetric manner across the Walmart and Amazon platforms. We selected 50 categories of products available on both platforms, ranging from headphones, dog food, and puzzles to stand-up paddle boards. We excluded digital products, books and fresh food. The list of product categories included in our data set can be found in Table A1 in appendix A. For each platform and each category, we used the list of the 50 best-selling products in September 2020. Across the two websites we then matched the two websites each of the 2,500 products of the list of best sellers. We were able to find about 60% of Walmart's best sellers on the Amazon marketplace and 50% of the Amazon best sellers on Walmart's.

Simultaneous web-scraping of each product page on both platforms took place concurrently, on a daily basis, for three months from September 12 to December 12, 2020. Everyday, we randomized the time at which each product data was collected on both platforms. Data gathered on the Amazon marketplace primarily included information about the buybox availability on the product page and, if available, information about the offer that was prominently displayed. We also collected the list of the various other offers available for the product and their details, in particular whether or not Amazon was present as a seller of the products. Additionally, we gathered data on each product's sales rank –a measure of current product popularity on the Amazon platform, as well as the manufacturer's suggested retail price (MSRP). Both are displayed on the product page. We complemented our data set with additional historical information on Amazon products that we got from the Keepa API. We collected similar information for the Walmart marketplace, including data about the featured offer on the product page as well as the list of all the different offers available for a product on this marketplace.

We focus our study on the Amazon marketplace and thus excluded all Walmart best-selling products which were not available for purchase on Amazon. Overall, our data set includes information about 3,851 products. Technical problems during data collection, due to server down-times or access changes, also led to the non-availability of some information. When this happened, we excluded the corresponding observation from our dataset.

5.2 Descriptive statistics

One observation in our main data set includes Amazon data at the product-time level, complemented with eBay and Walmart price data, and contains aggregates on the different sellers offering the product at each time. We focus on observations for which at least one seller is selling the product as new. Table 1 shows descriptive statistics which we will discuss in the following paragraphs.

Seller types (Panel A). Our data set includes product observations where Amazon itself is active as a seller, competing alongside third-party sellers on the marketplace, and observations where only third-party sellers offer a product. Amazon is acting as a seller for about half of the observations. For some products it acts as a seller for only certain time periods, possibly due to stock-outs or changes in assortment decisions.[15] At least one third-party seller is selling the product in 92% of the observations. The presence of Amazon as a seller of a specific product may vary over time. Amazon may temporarily leave the market for some products, possibly due to stock-outs. Amazon's exit as a product's seller may also be permanent. Among third-party

[15] We do not find indications that the presumable reasons for Amazon's temporal absence as a product's seller play a role for our investigation. The corresponding results can be found in Appendix O2.1.

Table 1: Descriptive statistics

	Mean	S.D.	Min	Max	# Obs.
Panel A: Availability of sellers by type					
Amazon	0.53	0.50	0.00	1.00	274568
3rd party (not Amazon)	0.92	0.27	0.00	1.00	274568
FBA seller	0.47	0.50	0.00	1.00	274568
FBM seller	0.76	0.43	0.00	1.00	274568
Panel B: Prices by seller type					
Amazon price	87.00	187.82	0.75	4001.26	144651
3rd parties' minimum price	99.30	179.75	0.66	3000.00	253140
FBA sellers' minimum price	72.55	136.07	2.69	1997.99	129065
FBM sellers' minimum price	110.83	193.73	0.66	3000.00	208631
All sellers' minimum price	93.53	176.77	0.66	4001.26	274424
Panel C: Buybox characteristics					
Buybox visible	0.83	0.38	0.00	1.00	274568
Buybox price	89.34	176.27	0.75	4001.26	225674
Min price when buybox is visible	88.77	175.33	0.66	4001.26	227510
Amazon in bb if bb visible + Amz seller	0.93	0.26	0.00	1.00	146280
Panel D: Product market characteristics					
- Log Sales Rank	- 8.75	2.37	0.00	- 15.31	247391
Product age (in #months)	71.48	53.64	0.00	267.00	269988
# Offers	9.24	13.03	1.00	280.00	274568
Minimum rec. sales price (MRSP)	102.76	195.92	0.96	2999.99	269132
Walmart (WM) price	94.10	178.65	0,01	5073.49	183042
eBay price	94.28	172.11	1.54	2779.00	157607
Product available at WM or eBay	0.79	0.40	0.00	1.00	274568

NOTE: This table shows the summary statistics for daily observations of our main variables of interest for 3,851 products. The variables are organized in four groups: Panel A shows the availability of specific seller types by day. Panel B summarizes the corresponding prices and Panel C denotes the buybox choices of Amazon in our data set. Panel D reports other product market characteristics.

sellers, we observe that products are more often sold by *Fulfilled-by-merchant* (FBM) sellers than *Fulfilled-by-Amazon* (FBA) sellers.

Prices (Panel B). Products sold by Amazon have, on average, a lower price than products sold by third-party sellers. This may partly stem from a different assortment (a composition effect) but may also result from these sellers indeed having lower prices for the same product offer. We also observe heterogeneity in the prices between third-party sellers, with prices of FBA sellers being, on average, lower than prices from FBM sellers and Amazon.

Buybox characteristics (Panel C). Amazon displays the buybox on the product page in 83% of our observations. Noticeably, the average price set by the seller featured in the buybox is of the same order of magnitude but slightly lower than the lowest price of the marketplace prices (including Amazon's price). For observations where Amazon is present as a seller of the product, we observe that Amazon gives prominence to its own offer in the buybox 93% of the time.[16] As these statistics may include observations when Amazon was not present, we further analyze the relationship between buybox presence and prices in the next subsection.

Product market characteristics (Panel D). Information about product demand is not directly available. Amazon provides information on the sales rank of each product, based on the number of purchases among products belonging to the same category. This variable enables us to measure the popularity of each product with respect to similar products. An increase in this measure indicates a decrease of the relative demand for this product within its category on the marketplace. Our data also includes information on the current number of sellers for a product, the current product rating plus the number of reviews for each product. In our data, products are available for purchase on either one of the two rival marketplaces Walmart and eBay 79% of the time. We observe that the price of products on Walmart and eBay are, on average, of the same order of magnitude, although slightly higher, than the buybox price on the Amazon marketplace. The data also suggests that, on average, Amazon's price when it acts as a seller is lower than the average prices on competing sales channels whereas the average third-party price on Amazon are, on average, higher.

6 Buybox availability and product demand

6.1 Descriptive evidence

Seller participation and buybox visibility. We differentiate observations with respect to the presence of the buybox on the product page in Table 2. Most strikingly, the buybox is never absent when Amazon is a seller of the product. We also observe that observations with no buybox exhibit, on average, a higher number of sellers than observations with a buybox (row 4). For observations with no buybox, there is almost always an FBM seller present (row 3) while at least one FBA seller is present in almost half of the cases (row 4). Products where no buybox is shown seem to have a lower sales rank than products where the buybox features a seller. This could indicate a relationship between the non-promotion of sellers and the popularity of the product *vis-à-vis* other products in the same category.

Result 1. *The buybox on a product page is always visible when Amazon sells the product. This*

[16] Other studies, such as Lee and Musolff (2021), employ data from repricing companies. Their data may include more products that third-party sellers typically sell.

Table 2: Descriptive statistics according to buybox visibility

	Buybox visible		Buybox invisible		All obs.	
	Mean	S.d.	Mean	S.d.	Mean	S.d.
	Product market characteristics					
Amazon is a seller	0.64	0.48	0.00	0.00	0.53	0.50
FBA seller present	0.48	0.50	0.44	0.50	0.47	0.50
FBM seller present	0.72	0.45	0.97	0.17	0.76	0.43
# Offers	8.29	12.15	13.84	15.85	9.24	13.03
-Log sales rank	-8.41	2.25	- 10.40	2.28	-8.75	2.37
	Minimum price on Amazon marketplace (p_{MinAmz})					
Freq. $p_{MinAmz} > p_{MRSP}$	0.12	0.33	0.58	0.49	0.20	0.40
$(p_{MinAmz} - p_{MRSP})/p_{MinAmz}$	-0.12	0.85	0.05	0.43	-0.09	0.79
Freq. $p_{MinAmz} > p_{AvgSubst}$	0.41	0.49	0.51	0.50	0.43	0.49
$(p_{MinAmz} - p_{AvgSubst})/p_{MinAmz}$	-0.05	0.73	0.01	0.57	-0.04	0.71
Freq. $p_{MinAmz} > min(p_{WM}, p_{eBay})$	0.40	0.49	0.85	0.36	0.48	0.50
$(p_{MinAmz}$-$min(p_{WM}, p_{eBay}))/p_{MinAmz}$	0.01	0.33	0.25	0.26	0.05	0.33
N	227511		47057		274568	

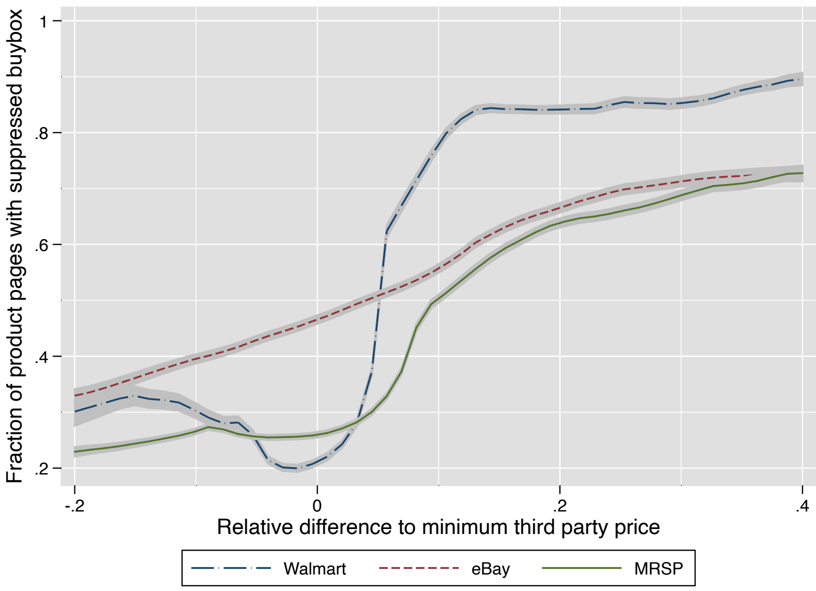
NOTE: Summary of daily observations of variables, differentiated by the presence of the buybox on the product page. The variables are organized in two groups: Product market characteristics and differences of minimum price available on the Amazon marketplace with some reference price.

is consistent with H2b.

Minimum price on Amazon and buybox presence. In line with hypothessis H1, there appear to be systematic differences in various price relations depending on whether the buybox is present or not. When the minimum price on Amazon is relatively high, the buybox is shown less often. High and low refers to relations of the Amazon price with either (i) the list price at Amazon (minimum recommended sales price –MRSP), the average category price at Amazon (ii), and the minimum price occurring on competing sales channels (iii). Observations with a buybox have, on average, a minimum price higher than the MRSP 12% of the time, whereas 58% of products without a buybox have a price above the MRSP. Similarly, products with an Amazon buybox are generally priced lower on either Walmart or eBay 40% of the observations, whereas products with no buybox exhibit a minimum price higher than the Walmart price in 85% of the observations.[17] These products show a positive average relative price difference with respect to the minimum price on competing platforms and a negative one for products with a buybox. Figure 5 illustrates these observations. We study these relationships systematically in the next subsection.

[17] In online appendix O1.2, we provide further descriptive statistics indicating that only third-party sellers matching the Walmart price can win a relevant share of the buybox impressions in our data.

Figure 5: Buybox invisibility and price difference



NOTE: The figure shows the share of product-date observations for which the buybox is invisible as a function of the relative difference between the minimum price available on the marketplace for a product and the a) Walmart price, b) the eBay price and c) the manufacturer recommended selling price, normalized by the minimum price, and is positive when X is lower than the minimum price.

6.2 Determinants of buybox visibility

We study the determinants behind Amazon's choice of whether or not to present the buybox, that is, to show the sales button for a particular seller in the buybox on the product page. To do so, we use product-level data available on the Amazon website. We complement this data with information on products in a competing marketplace (Walmart and eBay) in order to investigate whether Amazon's behavior is related to inter-channel competition.

Regression model. The task is to explain how Amazon decides whether or not to promote a seller in the buybox. Our approach relies on using variation over time in the availability and prices of sellers on the platform and elsewhere. We estimate a linear model with observations at the level of the product (i) and date (t):

$$BuyboxVisible_{i,t} = \beta' X_{i,t} + \gamma' Z_{i,t} + \xi_i + \xi_{c,t} + \varepsilon_{i,t}, \quad (1)$$

where $BuyboxVisible_{i,t}$ is a variable taking value 1 if Amazon is showing the buybox and else 0. Equation (1) can be conceptualized as a reduced-form equation that comes out of a mechanism that determines whether to show the buybox or not as a result of the prices and qualities of the sellers relative to potential substitutes in the Amazon marketplace and elsewhere. $X_{i,t}$ are time-varying product market characteristics like the availability of sellers and distinct types, a demand measure, and the minimum price. $Z_{i,t}$ contains measures of substitute products on and off the platform. ξ_i are product and $\xi_{c,t}$ are date-category fixed effects.

Identification. The key challenge is to identify whether there is a causal link between our competition variables and Amazon's decision over whether to show the buybox. There are several issues to address. To account for unobserved time-constant heterogeneity between products, we employ product fixed effects. Hereby, we essentially compare how the buybox presence varies over time with the changing competitive landscape surrounding a product. When deciding whether to display the buybox, Amazon might take factors affecting the whole product category into account, e.g., a higher interest in the category on a specific day. Higher demand for a category could lead to less buybox suppression for all categories, and omitting this information might bias our results. We control for this with date-category fixed effects. This rules out that a parallel change in the same category is driving the result. Finally, by including fixed effects for the hour of the day, we capture that, at times, the overall demand on Amazon may be higher or lower.

Regression outcomes. Table 3 contains the regression results of the presence of the buybox on a product page at a particular point in time as a function of market characteristics and minimum price levels. A positive coefficient indicates that the explanatory variable positively affects the likelihood that the buybox is present. As mentioned in the previous section, the presence of Amazon as a seller is a perfect predictor of buybox presence. We thus restrict the sample to product observations where Amazon does not appear as a seller.[18] The baseline specification (1) includes product market characteristics. These include the lag of the inverse sales rank in order to control for the popularity of the product on the marketplace, the presence of different types of third-party sellers and the number of offers available, the minimum price available and the availability of the product in competing marketplaces. In the additional specifications of Table 3, we added sequentially how the product's minimum price on Amazon relates to the MRSP, the average price of similar products, and the lowest price on competing sales channels. These variables take the value 1 for products where the minimum available price is strictly higher than the price to which it is compared to.

[18] Regressions on the whole sample controlling for Amazon's presence as a seller are available and discussed in Table O6 in online appendix O3.

Table 3: Determinants of buybox visibility – OLS regression results

	(1)	(2)	(3)
- Lag log sales rank	0.03***	0.03***	0.03***
	(0.00)	(0.00)	(0.00)
FBA seller present	0.09***	0.07***	0.07***
	(0.02)	(0.02)	(0.02)
# Sellers	-0.02***	-0.03***	-0.03***
	(0.01)	(0.01)	(0.01)
Minimum price	-0.02***	-0.02***	-0.02***
	(0.00)	(0.00)	(0.00)
Product available at Walmart or eBay	0.02	0.00	0.01
	(0.03)	(0.02)	(0.02)
I[Min price > MRSP]		-0.08***	-0.08***
		(0.01)	(0.01)
I[Min price > Avg. subst. price]		-0.04***	-0.04***
		(0.01)	(0.01)
I[Min price > min(price WM/price eBay)]		-0.22***	-0.22***
		(0.02)	(0.02)
Share of AMZ products among substitutes			0.02
			(0.02)
Constant	1.13***	1.19***	1.18***
	(0.05)	(0.05)	(0.05)
Observations	112,491	112,491	112,491
R-squared	0.81	0.82	0.82

Notes: Dependent variable: Buybox availability. Unit of observation: Product-Date. Linear regressions include fixed effects at the product, date, and hour of the day level as well as controlling for product age. Column (3) includes only products which show product-comparison tables (see Section 7). Standard errors are robust to heteroscedasticity and adjusted for serial correlation inside clusters. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

We interpret the regression as follows: Amazon is more likely to show the buybox for products with high popularity in the marketplace (a higher value of the negative sales rank reflects more sales). The presence of FBA sellers appears to positively affect the buybox presence. This pattern might arise because Amazon gets higher commission payments from transactions with FBA sellers compared to FBM sellers. As Amazon handles the logistics of FBA sellers, the relationship might also be driven by quality considerations and customer satisfaction. The level of the minimum price also significantly affects the buybox presence. The additional columns reveal that when the minimum price is higher compared to the MRSP there is a negative impact on buybox presence. This indicates that Amazon discourages sales at prices above those recommended by the manufacturer. Similarly, products with a higher minimum price relative to the average price of similar products in the same category are less likely to have a buybox as well. This variable measures the price competitiveness of the product relative to (potential) substitute products within the Amazon marketplace. Column (2) shows that, after controlling for MRSP and a comparison of the average prices of similar products, the buybox is less present when the product is expensive in comparison to the Walmart price. The coefficient associated with this latter variable is of a higher level (20%) than the ones associated with MRSP and the average price of similar products. In Column (3) we further control for the share of products among potential substitutes that Amazon is currently selling (the set of substitutes is defined by products appearing in a comparison table). We do not find that Amazon suppresses the buybox significantly more often when it is currently selling a high share of substitutes. We will come back to this point in Section 7.

Factors moderating the platform competition effect and robustness. Our regression analysis indicates that Amazon tends to remove the buybox when competing websites have a lower price. To investigate whether Amazon's buybox decision is affected by other factors, we run the specifications in Table O4 in the online appendix O2.1 using different interactions. Summarizing, we find that buybox absence due to lower prices on competing sales platforms is (1) independent of whether Amazon has ever, never or recently also offered the product, (2) negatively correlates with the number of available sellers, (3) independent with the number of similar products available on the marketplace. Online appendix O2.1 contains a detailed discussion. We also run different robustness checks in online appendix O3. For instance, we employ different specifications for the price differences across platforms (Table O5) and we investigate potential selection biases by including product observations on dates where Amazon was also selling the product (Table O6). The results do not change qualitatively. Online appendix O3 contains a detailed discussion.

Result 2. *On the Amazon marketplace, the buybox of a product is less frequently visible if the minimum price of the product is high relative to different other prices, that is, the prices of substitutes on the marketplace, the recommended retail price, and the product's price on other platforms. This is consistent with hypothesis H1.*

Recall that we focus on the best-selling products of different categories, such that the result is not driven by rarely sold niche products.

6.3 Prediction of buybox visibility when Amazon is a seller

Figure 5 and the regression results in subsection 6.2 reveal that Amazon makes the visibility of the buybox dependent on prices of other products within the Amazon marketplace and the price level of the product in comparison to competing marketplaces. The descriptive statistics in subsection 6.1 show that Amazon always makes a product's buybox visible when it is one of the sellers of the product.

We simulate how often Amazon would not show the buybox when it sells the product if it were to apply the same algorithm as for third-party only markets. To do so, we use product-day observations where Amazon is not one of the sellers to train a model predicting whether the buybox is visible. We then apply the model to product-day observations where Amazon was indeed a seller (out of sample).

To predict whether the buybox should be removed according to the estimated model (Amazon's algorithm) and the observables, we treat the presence of Amazon like the presence of an FBA seller. This seems appropriate in the context of Amazon's announced policy of removing the buybox if it sees harmful pricing practises (subsection 4.1). The assumption effectively means that if a buybox should be invisible as the price that an FBA seller charges is considered harmful, the fact that Amazon charges the same price should not change the assessment of the price being harmful, other things equal.

We estimate a model similar to column (3) of Table 3 using OLS and a logit estimator. We include for the specification "OLS1" fixed effects for the category-date and the hour of the data collection. The specifications "OLS2" and Logit include fixed effects for the category, date, and hour of the data collection. The omission of product fixed effects does not change results qualitatively.

Table 4: Prediction of buybox visibility using different models

	(1)	(2)	(3)	(4)	(5)
Product-day-observations where...	Amazon not a seller			Amazon a seller	
Buybox	invisible	visible		invisible	visible
% (full sample)	36.5	63.5		0.0	100.0
% (competing price available)	47.6	52.4		0.0	100.0
Prediction %	... in sample		%correct	... out of sample	
model: OLS1	39.1	60.9	86.9	13.4	86.6
model: OLS2	39.5	60.5	86.4	13.2	86.8
model: Logit	37.0	63.0	84.8	14.1	85.9
Sample means for different cases (based on OLS1) if competitor price is available					
% Competitor cheaper	89.1	30.6		84.6	54.7
Price difference	25.2	-15.5		20.6	5.7

Notes: Calculations based on OLS and Logit regression models with dependent variable "Buybox visible" using data when Amazon is not a seller, like in column (3) in Table 3. "Buybox visible" is predicted if the (latent) model indicates that $\hat{y} > 0.5$.

Table 4 summarizes the results. For the predictions, we set *buybox visible* to one if the predicted value ($\hat{y}$) exceeds 0.5 and to zero otherwise. With this, the models predict shares of the buybox visibility the buybox visibility correctly in 85-87 percent of the in sample observations where Amazon was not observed as a seller (columns 1 to 3). When predicting the buybox visibility out-of-sample for the product-day observations where Amazon was indeed one of the sellers (columns 4 and 5), the buybox should be invisible in about 13 to 14 percent of the cases. This relies on treating Amazon equal to other FBA-sellers. If Amazon is one of the sellers, the buybox is thus visible in cases where it would not be visible given other covariates (such as prices) if there were "just" other FBA sellers but not Amazon itself.

One can grasp the relevance of the prices for the visibility of the buybox by studying the differences of the lowest sales price on the Amazon marketplace and the minimum of the eBay and Walmart price. The variable 'price difference' is this price difference in relation to the price level and the variable 'Competitor cheaper' is an indicator that is one if the difference is positive, so that the price on the Amazon marketplace is higher than the price on the competing website. Let us first compare the sample means of these variables in the first two data columns where Amazon is not a seller. If the buybox is invisible, the competitor's price is cheaper in more than 89 percent of the cases with an average price difference of about 25 percent. If the buybox is visible, instead, the competitor is only cheaper in less than 30 percent of the observations and, on average, the offer on the Amazon marketplace is about 16 percent cheaper.[19]

Columns 4 and 5 contain the sample averages for the observations where Amazon is a seller; this sample is split based on our model predictions of whether the buybox should be visible or not. For the cases where the buybox is predicted invisible (column 4), the sample means are close to those where the buybox is indeed invisible when Amazon is not a seller (column 5). This indicates that in these cases the buybox should indeed be invisible if one were to apply Amazon's policy of relating buybox visibility to price competitiveness (see Section 4) also to instances where Amazon sells the product itself. This yields

Result 3. *When Amazon sells a product itself, for the buybox to be visible it does not require as competitive prices and conditions as it requires in instances where it is not a seller.*

Moreover, in cases where Amazon is a seller and the buybox is predicted to be visible, the prices are considerably less competitive compared to when Amazon is not a seller and the buybox is

[19] When conditioning on observations where an FBA seller was present, the share of observations where the competitor is cheaper is lower (24.9 percent) and the average relative price difference is also even lower (-22.2).

visible (the competitor is cheaper in 55% versus 31% of the cases and the relative price difference is 6% versus -16%). This indicates that consumers face, on average, higher prices relative to the competitive benchmark when the buybox is visible and Amazon is one of the sellers.

In isolation, the fact that the buybox is never invisible when Amazon is a seller of a product but partly invisible when Amazon is not a seller (Table 2) could be consistent with Amazon's announced policy of removing the buybox if it sees pricing practices that harm customers. It could be that the prices are always competitive enough when Amazon is a seller to make the buybox visible. However, the latter finding is difficult to reconcile with this policy. It suggests that Amazon's explanation of the buybox removal policy could benefit from being explicit about whether this applies only to third party sellers or also to own offers.

6.4 The impact of buybox visibility on sales

We now investigate the effect of buybox visibility on the realized sales of that product on Amazon. This also sheds light on the sequence of events. We use a similar approach as to Reimers and Waldfogel (2021) and estimate the following equation:

$$-\,log(SalesRank_{i,t}) = -log(SalesRank_{i,t-1})\alpha + BBInvisible_{i,t}\beta + \gamma' X_{i,t} + \xi_i + \xi_t + \varepsilon_{i,t}, \quad (2)$$

where $-log(SalesRank_{i,t})$ is the log of the current sales rank of the product in the category multiplied by -1, and $-log(SalesRank_{i,t})$ is its lag. A motivation for including the lag is to capture all the past marketplace choices and factors which might also influence the sales rank today.[20] For ease of interpretation, we multiply times -1 such that a higher value on the left-hand side corresponds to more demand. The right-hand side variable $BBInvisible_{i,t}$ is an indicator which equals 1 if the buybox is not visible; $X_{i,t}$ are product and market characteristics; ξ_i are product and ξ_t are date fixed effects.

We estimate three different specifications and report the results in Table 5. In all specifications demand is higher if Amazon is a seller, if there are FBA sellers, and if there is a higher number of sellers. The higher the lowest price offered for this product, the lower the demand. Furthermore, demand also decreases when the price of other products in the same category on Amazon are lower and if the minimum price for the product is above the MRSP. Hence, we

[20] The inclusion of the lag of the dependent variable could suggest the use of a dynamic panel data model, as the estimator could be inconsistent. However, as noted by Greene (2012) on p. 400, the bias converges to zero as the number of periods (T) increases. We observe products for a period of three months, such that $T = 87$.

Table 5: Impact of buybox invisibility on demand

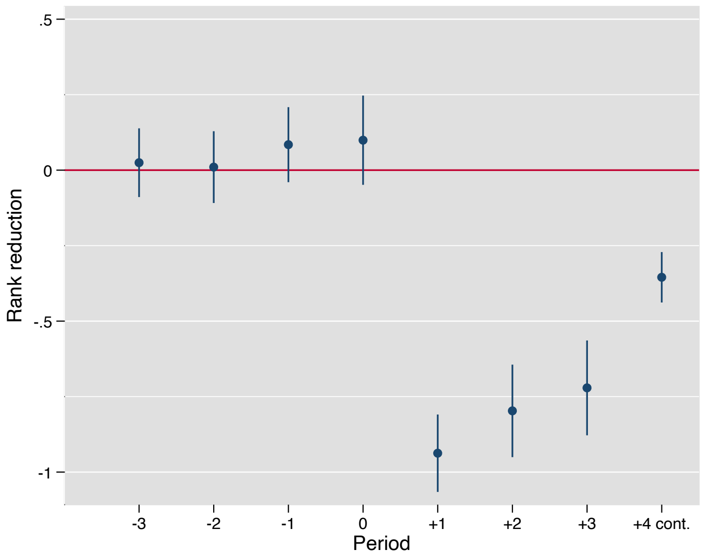
	(1)	(2)	(3)
- Lag log sales rank	0.83***	0.83***	0.83***
	(0.00)	(0.00)	(0.00)
Buybox invisible	-0.10***	-0.09***	-0.11***
	(0.01)	(0.01)	(0.01)
Amazon sells product	0.10***	0.08***	0.07***
	(0.01)	(0.01)	(0.01)
FBA seller present	0.02***	0.01***	0.01**
	(0.00)	(0.00)	(0.00)
# Sellers	0.02***	0.02***	0.01***
	(0.00)	(0.00)	(0.00)
Min price	-0.01***	-0.01***	-0.01***
	(0.00)	(0.00)	(0.00)
I[Min price > Avg. subst. price]		-0.05***	-0.04***
		(0.00)	(0.00)
Product available at Walmart or eBay		-0.01	-0.02
		(0.01)	(0.01)
I[Min price > min(price WM/ price eBay)]		-0.04***	-0.04***
		(0.00)	(0.00)
Min price * Buybox invisible			0.00***
			(0.00)
Constant	-1.47***	-1.44***	-1.42***
	(0.04)	(0.04)	(0.04)
Observations	243,358	243,358	243,358
R-squared	0.98	0.98	0.98

Notes: Dependent variable: (minus) log sales rank. Unit of observation: product-date. Linear regressions include fixed effects at the product, category-date, and hour of the day level. Controls include the age of the product. Standard errors are robust to heteroscedasticity and adjusted for serial correlation inside clusters.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

can confirm significant substitution effects of offers on the platform. Furthermore, a product with an obscured buybox is associated with a lower demand of this product within its category. Specifically, the sales rank of a product without a buybox worsens by about 50 to 58%. In the partial adjustment model of Equation (2), we set $ln(y_t) = ln(y_{t-1})$ to calculate the full effect of a right hand-side variable. To do so, we divide the coefficient of interest (that of "Buybox invisible" in this case) by *one minus the coefficient of the lagged dependent variable.* For the results in column 1 of Table 5, this yields $\frac{-0.10}{1-0.83} = -0.59$. Assuming the same distributional pattern as in other studies, such as Reimers and Waldfogel (2021), we multiply this figure by one half. This means that an invisible buybox, other things equal, is on average associated with a decrease in sales of about 30%.

In column 2 of Table 5, the indicators of the minimum price *in* Amazon being larger than prices *outside* Amazon (notably Walmart and eBay) have significantly negative coefficients, indicating

Figure 6: Event study for demand reduction due to buybox suppression



NOTE: This figure shows the decrease of a products demand following a buybox suppression. The effects are calculated by a regression as in Table 5, column (2) where the indicator "No buybox" is interacted with indicators for the different lags and leads. The fourth lag ("+4 cont.") also includes all subsequent periods during which the buybox is still suppressed. These coefficients are further divided by the 1 minus the coefficient of the log lag sales rank.

substitution effects between Amazon and Walmart/eBay as well as among different substitutes in the Amazon marketplace. In column 3, we add the interaction between the buybox being invisible and the minimum price for the product. The associated coefficient is significantly positive, meaning that a product's price has less effect on demand if the buybox is invisible. This is consistent with the conjecture that those who buy the product although the buybox is absent are less price sensitive.

Timing and causality. In Figure 6, we investigate the time structure between buybox invisibility and demand to better understand the causality between the two variables. We include indicator variables taking on value one when the buybox has been suppressed for the first time for at least four days in a row, as well as indicator variables for the leads and lags of this buybox deletion event. We find that the buybox suppression decreases demand significantly starting the day after the suppression ("+1"), but not before.

Result 4. *The regression results are consistent with the hypothesis H3 that buybox invisibility significantly decreases demand for the product.*

7 Buybox invisibility and substitute products' selection

7.1 Comparison table and descriptive statistics

We study another important element of the product pages: a table with the title "Compare with similar items" that contains information on substitute products available on the Amazon marketplace (see Figure 3). In this comparison table, which we note is ever displayed for 80% of products in our dataset, consumers can easily compare the product they were considering in the first place with up to five similar products in terms of technical characteristics, as well as price, seller identity, and ratings. It is noteworthy that on the page of the same product, the comparison table contains different products at different points in time. The choice of products to display in this element is not influenced by monetary payments of third-party sellers, contrary to other recommendation sections on the web-page. Using a link located next to the buybox section directing to this element, consumers can easily compare and navigate across available alternatives on the marketplace.

Table 6: Descriptive statistics of table "Compare with similar items"

	Mean	S.D.	Min	Max	N
Panel A: Statistics on comparison table on product i's page					
Comparison table visible	0.80	0.40	0.00	1.00	221772
# Products ever recommended	9.06	3.35	1.00	24.00	221772
# Products displayed at a time	3.84	0.89	3.00	5.00	177647
Panel B: statistics of all products j ever in comparison table					
Amazon is a seller	0.64	0.48	0.00	1.00	2009154
3rd party sellers present	0.83	0.37	0.00	1.00	2009154
Min price	93.36	160.65	0.01	5256.24	2008100
Log sales rank	8.09	2.28	0.00	15.06	1896277
# Sellers	8.49	13.55	1.00	163	2009154
# Recommendations received	2.99	5.25	0.00	44.00	2009154
Panel C: statistics of products j when present in comparison table					
Amazon is a seller	0.69	0.46	0.00	1.00	681679
3rd party sellers present	0.85	0.36	0.00	1.00	681679
Minimum price	88.00	154.44	0.01	4617.99	681574
Log sales rank	7.89	2.30	0.00	14.55	644182
# Sellers	8.60	12.66	1.00	159.00	681679
# Recommendations received	5.60	6.57	1.00	44.00	681679

NOTE: Panel A: statistics regarding the "compare with similar items" table on page of product i calculated at the product i-date level. Panel B: summary statistics at product j-date level of all products j which ever appeared in the recommendation table of the set of products i. Panel C: statistics of products j from panel B restricted to the dates they were actually displayed in comparison table.

We first investigate whether a placement of a product j in the comparison table of product

i affects the demand of product j. Moreover, we analyze whether a product j in this table receives more attention and realizes more sales when Amazon refrains from showing a buybox on the page of product i. Recall from subsection 6.4 that buybox invisibility on the page of product i appears to reduce the demand for product i. Consumers who do not see product i's buybox might continue to search on Amazon for similar products. Amazon facilitates this by displaying alternative products in the comparison table on the product pages. We denote these products by j. Consumers might look at these alternative *products* more often if the buybox is invisible. Building on these insights, we study whether Amazon is potentially making use of these recommendations to divert consumers more toward products which it is selling itself, especially when the buybox is absent.

Table 6 summarizes the data we use for these analyses. Observe in Panel A that in 80% of cases, Amazon includes a comparison table on the product page of product i. On the page of a given product i, there are on average nine products which appear in the comparison table at different points in time during our observation period. Out of these nine alternatives, about four alternatives appear in the comparison table at the same time. We amend this data set with information about all these alternative products. We get a data set at the level of pairs of the currently shown product i and each of its potential substitutes j (Panel B). With this we study the characteristics of the products recommended at time t (Panel C). A comparison of the descriptive statistics in panels B and C reveals that out of the potential substitutes, products for which Amazon is present as a seller are recommended with a higher frequency (0.69 vs. 0.64) and appear to be cheaper (US \$88.56 vs. \$93.60). On average, each product receives about five recommendations each time it is recommended.

7.2 Effects of product recommendations on demand

We investigate the effect of a product j being displayed in the "Compare with similar items" table of a product i page on the realized demand of a product j on the Amazon marketplace. We also investigate whether recommending a product j leads to more sales when the recommendation occurs on the page of a product i where the buybox is absent. Using the data set of all products j that were *ever* recommended, we estimate the following equation, similar to the analysis in subsection 6.4:

$$-log(SalesRank_{j,t}) = \alpha \cdot [-log(SalesRank_{j,t-1})] + \beta \cdot log(\#InTables_{j,t}) + \gamma' X_{j,t} + \xi_j + \xi_t + \varepsilon_{j,t}. \tag{3}$$

The dependent variable $-log(SalesRank_{j,t})$ contains product j's logged sales rank in the category multiplied by -1 and $-log(SalesRank_{j,t})$ its lag. The right-hand side variable $log(\#InTables_{j,t})$ contains the logged number of cases we see in which product j at time t is in a comparison table in the set of products i that we are observing. The vector $X_{j,t}$ contains product and market characteristics, whereas ξ_i are product and ξ_t are date fixed effects.

Table 7: Impact of presence of product j in comparison tables on its sales rank

	(1)	(2)
- Lag log sales rank	0.82***	0.82***
	(0.00)	(0.00)
Amazon sells j	0.06***	0.06***
	(0.00)	(0.00)
FBA sellers present for j	0.04***	0.04***
	(0.00)	(0.00)
# Sellers j	0.02***	0.02***
	(0.00)	(0.00)
Minimum price j	-0.01***	-0.01***
	(0.00)	(0.00)
Buybox j invisible	-0.14***	-0.14***
	(0.00)	(0.00)
Product j at eBay	0.02**	0.02**
	(0.01)	(0.01)
I[Min price j > eBay price j]	-0.04***	-0.04***
	(0.00)	(0.00)
log(1+#j in comparison tables)	0.01***	
	(0.00)	
log(1+#j in tables where Buybox i visible)		0.01***
		(0.00)
log(1+#j in tables where Buybox i invisible)		0.02***
		(0.00)
Constant	1.53***	1.54***
	(0.03)	(0.03)
Observations	638,836	638,836
R-squared	0.98	0.98

Notes: unit of observation: product-date; observations: all products j observed in recommendation tables of products i; dependent variable: (minus) log sales rank; linear regression with fixed effects at product, category-date, and hour of the day levels; controls include product age; standard errors robust to heteroscedasticity and adjusted for serial correlation inside clusters.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 7 exhibits the regressions results of two different specifications. Column 1 contains an explanatory variable of the number of cases where product j appears in a comparison table of a product i, irrespective of the visibility of the buybox on the page of product i. Being recommended more increases the demand for product j significantly, suggesting that a portion of consumers on the page of product i redirect to the alternatives displayed in the comparison table.

In column 2, we differentiate the cases where product j is in a comparison table on the product page i with respect to the visibility of the buybox on that page. We observe that the coefficient on the cases in the comparison table coming from product pages i where the buybox is invisible is more than 50% higher than the coefficient on cases from product pages where the buybox is visible.[21] This suggests that the absence of the buybox makes consumers more attentive to alternative product recommendations and thus more likely to re-direct toward other alternatives displayed in the recommendation section.

Regarding the size of the effect, we apply the same reasoning as in subsection 6.4. We find that the effect of a 1% increase in the number of appearances of a product j in comparison tables on pages of products i with an invisible buybox improve the sales rank by 9.6%, which implies an increase in sales of about 5%. The corresponding sales increase is about 2.5% for appearances of product j in comparison tables on pages where the buybox is visible. Here, a caveat applies as we count only the appearance in comparison tables of products i which occur in our sample. As products i in our sample are the best-selling ones, it is not surprising that we find significant effects even though we do not measure what happens in other comparison tables. Our findings are representative of the set of all comparison tables on the marketplace if, on average for each product j, appearing in comparison tables in our sample is perfectly correlated with appearing in comparison tables not in our sample. Otherwise, the true effects of appearing in comparison tables might be even larger.

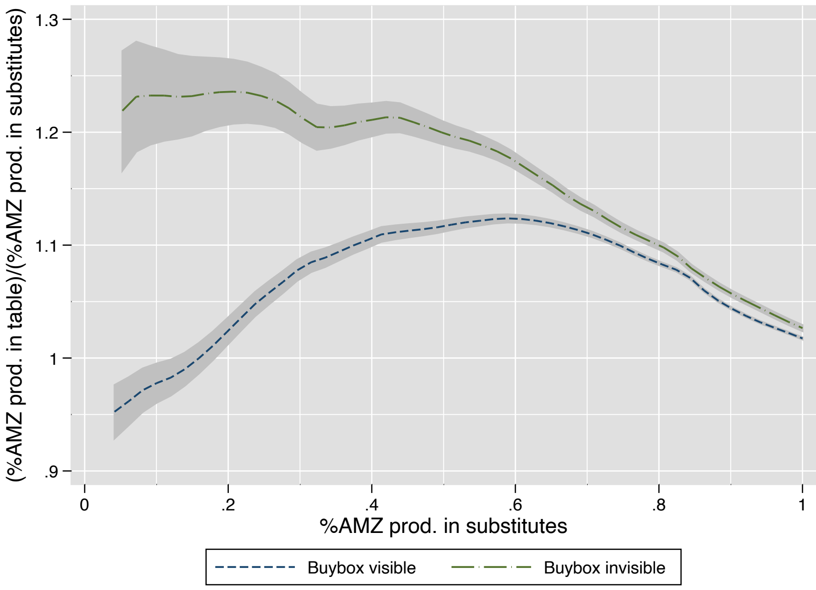
Result 5. *The regression results indicate that a product's demand is significantly higher when it is present in a comparison table on a product page. This is consistent with H4. The effect is about one and a half times as strong when the buybox on that product page is invisible. This is consistent with H5.*

[21] The coefficients are 0.01698 and 0.01084, such that the relative difference is 56.6%. The difference of the coefficients is significantly different from zero with *p=0.0647*.

7.3 Buybox (in)visibility and composition of comparison table

The analyses of the previous subsection indicate that the demand of product j increases more if it is presented in the comparison table on the page of a product i if the buybox of that product i is invisible. We now study the composition of the comparison table. In particular, we investigate whether the composition of the table on the page of product i depends on whether the buybox is visible on the page of product i.

Figure 7: Amazon's presence in comparison table relative to its overall presence



NOTE: The figure compares Amazon's presence for products j in the comparison table relative to its overall presence, given that the buybox for product i is shown or not.

A simple mean comparison between panels B and C in Table 6 suggests that Amazon might promote products it sells itself more often. Before we turn to the multivariate regression analyses, let us look at Figure 7. Essentially, we construct it using two fractions at the product i level: (a) the number of products in the comparison table that are sold by Amazon divided by the number of products in the comparison table; (b) among the products which ever appear in the comparison table of product i as substitutes, the number of products that are sold by Amazon. The figure depicts the ratio of fraction (a) over fraction (b) on the vertical axis for different levels of fraction (b) on the horizontal axis. As a reference, the ratios would be equal to 1 if the products appearing in the comparison tables were randomly drawn from the potential substitutes. We observe in the figure that the lines are mostly above 1, indicating that products sold by Amazon appear more often than under random drawing. Note that the figure contains two lines, differentiated by whether the buybox on the page of product i was visible at the time the comparison table was shown. The line consisting of observations where the buybox was invisible is significantly above the line where the buybox was visible. This could indicate that the non-random selection of products sold by Amazon in the comparison table is more pronounced if the buybox is invisible.

For the regression analyses, we employ an approach similar to Chen and Tsai (2019) on a data set of potential substitutes j of each product i that could potentially appear in the comparison table. To learn which factors make an appearance in the table more likely, we estimate the model

$$InTable_{j,i,t} = \beta' X_{j,t} + \gamma' Z_{j,i,t} + \xi_{j,i} + \xi_{i,t} + \varepsilon_{j,i,t}, \tag{4}$$

where $InTable_{j,i,t}$ is a binary variable taking value 1 if product j is recommended in the comparison table of product i at time t. $X_{j,t}$ are time-varying product j market characteristics, such as the availability of seller types (e.g., FBA), demand measures, the lowest price, and buybox presence. $Z_{j,i,t}$ denotes an interaction between Amazon's presence as a seller in product j and the presence of the buybox on product i's page, whereas $\xi_{j,i}$ are product pair j and i fixed effects and $\xi_{i,t}$ are product i-date fixed effects. Note that we do not control for the buybox presence of product i here as it is captured by the product i, t fixed effect.

Table 8 contains the regression results. These suggest that a product j appears more frequently in the comparison table if it is more popular and FBA. A low minimum price also increases the frequency – both in absolute terms and when the minimum price is lower than the average minimum price among products in the substitutes choice set. When Amazon sells product j itself, the likelihood of the product being in the comparison table is 19% higher. This is consistent with H7.

In column 2, we add an indicator that is one if product j has no buybox. We find the frequency of appearing in the table is lower if product j has no buybox. In column 3, we add as a variable the interaction between Amazon selling the product j with an indicator of the buybox of product i being absent. We observe that when the buybox is invisible for product i, product j has a 3% increased likelihood of appearing in the comparison table on the page of product i if Amazon is a seller of product j.

In column 4, we further investigate the newly identified effect of an invisible buybox increasing the likelihood of an Amazon-sold product appearing in the comparison table. We study how

Table 8: Determinants of product j appearing in comparison table on page of product i

	(1)	(2)	(3)	(4)
- Lag log sales j	0.03***	0.03***	0.03***	0.03***
	(0.00)	(0.00)	(0.00)	(0.00)
FBA sellers present for j	0.02***	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)
# Sellers j	-0.01***	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)
Minimum price j	-0.00***	-0.00***	-0.00***	-0.00***
	(0.00)	(0.00)	(0.00)	(0.00)
I[Min price j > eBay price j]	-0.01**	0.01	0.01	0.01
	(0.00)	(0.00)	(0.00)	(0.00)
I[Min price j > Avg. subst. price i]	-0.01**	-0.01	-0.01	-0.01
	(0.00)	(0.00)	(0.00)	(0.00)
No Bbox on page of product j		-0.34***	-0.34***	-0.34***
		(0.01)	(0.01)	(0.01)
AMZ sells j	0.19***	-0.00	-0.01	-0.00
	(0.00)	(0.01)	(0.01)	(0.01)
AMZ sells j * AMZ sells i			-0.00	-0.00
			(0.01)	(0.01)
AMZ sells j * No Bbox i			0.03***	0.07***
			(0.01)	(0.02)
AMZ sells j * share AMZ subst. i				-0.02
				(0.02)
AMZ sells j * No Bbox i * share AMZ subst. i				-0.08**
				(0.04)
Constant	0.57***	0.64***	0.64***	0.64***
	(0.02)	(0.01)	(0.01)	(0.01)
Observations	1,647,259	1,647,259	1,647,259	1,647,259
R-squared	0.51	0.52	0.52	0.52

Notes: dependent variable is indicator taking value 1 if product j is present in the comparison table on product i's page; unit of observation: product pair $j-i$-date; linear regressions include fixed effects at the product pair, product-i-date and hour of the day level; standard errors: robust to heteroscedasticity and adjusted for serial correlation inside clusters.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

this effect depends on the fraction of products sold by Amazon of all potential substitutes for the comparison table on the page of a given product i. This corresponds to the differentiation on the horizontal axis of Figure 7. We find that the effect of buybox invisibility on the selection of an Amazon product decreases if the fraction of Amazon-sold products in the set of substitutes increases. This pattern is broadly consistent with the pattern observed in Figure 7 where we do not control for confounding factors.[22]

Result 6. *Products appear in the comparison table more frequently if Amazon sells them, con-*

[22] In Table O7 in online appendix O3, we also decompose this interaction of intervals of the absolute numbers of Amazon products among substitutes.

sistent with H7. In particular, products sold by Amazon appear in the comparison table of a product page more frequently if the buybox on that page is invisible. This is consistent with H8.

7.4 Alternative explanations

Reverse causality. The previous result relies on the interpretation of the partial correlation in Table 8 between buybox non-availability and the number of products sold by Amazon in the comparison table. In particular, we interpret this as a causal effect of the suppression of the buybox of product i on the selection of products j into the comparison table.

Reverse causality is possible and potentially interesting as well: Amazon might suppress the buybox on a product page more often if the page's comparison table contains many products that Amazon sells. Nevertheless, we will now explain why our econometric results contradict this reverse causality.

Specifically, we want to distinguish between the following two cases.

- Case 1: Amazon's decision of whether to show the buybox does NOT depend on the content of the comparison table. Instead, the share of Amazon products in the comparison table of product page i depends on whether the buybox on the product i page is available.
- Case 2: Amazon's decision of whether to show the buybox depends on the composition of the comparison table. In particular, Amazon might suppress the buybox and thus increase demand for products in the comparison table when Amazon sells many of these.

To distinguish between the two cases, we exploit variation in the presence of products sold by Amazon among the products that potentially appear in the comparison table. Recall that on the page of a particular product i, different products appear in the comparison table at different points in time. We define the set of products ever appearing in the table of product i in our data set as the potential substitute products. For each product i, we compute the share of products sold by Amazon in this set and call this variable $PotentialShareAmz_i$.

We know from Table 3 that the $PotentialShareAmz_i$ is positively correlated with the number of Amazon products in the comparison table of product i. Let us reconsider the regression model from Section 6 specified in equation 1 where we now explicitly include the $PotentialShareAmz_i$ as an explanatory variable:

$$BuyboxVisible_{i,t} = \beta_0 \cdot PotentialShareAmz_i + \beta' X_{i,t} + \gamma' Z_{i,t} + \xi_i + \xi_t + \varepsilon_{i,t}. \qquad (1b)$$

In the regression model, the coefficient β_0 should be positive if Amazon's decision of whether to show the buybox of product i depends on the composition of the comparison table on product i's page (Case 2). Instead, if there is no such effect, β_0 should be zero (Case 1).

Column (3) of Table 3 contains the results of this regression. Specifically, β_0 is not statistically different from zero. This means that we can reject Case 2 in favor of Case 1, which confirms our previous interpretation of the results in Table 8.

Consumer preference for Amazon products in comparison table. We document in Section 7.3 that products sold by Amazon appear in the comparison table of a product page more frequently if the buybox on that page is invisible. We argue that this is difficult to reconcile with an integrated marketplace operator which, at least partially, internalizes the profits of the affiliated retail business. Let us explore an alternative explanation whereby an independent marketplace operator would also present products sold by Amazon more often in the comparison table on a page where the buybox is invisible. On a given page of product i, an independent marketplace operator finds it optimal to display a certain number of product offers where Amazon is visible as a seller. The rationale could be that some consumers are only, or mostly, interested in products which are obviously sold by Amazon.

Hence, if the buybox on the page of product i is invisible or if the buybox is visible but Amazon is not in the buybox, the marketplace operator might want to add a product sold by Amazon in the comparison table. According to the regression results reported in Table 8, this does not appear to be the case as the coefficient of the interaction "AMZ sells j * AMZ sells i" is not significantly different from zero. This speaks against the alternative explanation and reinforces Result 5.

8 Conclusion

Contribution. We have studied key design features of the Amazon website to better understand how the marketplace functions with respect to customer steering and the interplay between the marketplace operations and the retail business. We analyze Amazon's marketplace and focus on the product pages. By presenting detailed descriptive statistics as well as extensive regression analyses, we have illustrated that both the buybox and the comparison table influence consumer behavior in the Amazon marketplace. Moreover, we provide results indicating that these tools are combined in a way that is difficult to reconcile with the hypothesis of an independent marketplace operator. We hereby illustrate how subtle and difficult it is to detect forms of self-preferencing that might take place on hybrid sales platforms.

Limitations and scope for future research. We have carefully collected and analyzed the data by means of extensive regression analyses where we control for confounding factors with a rich set of control variables as well as extensive sets of fixed effects. We also conduct additional analyses for robustness and to exclude reverse causality. However, we acknowledge at least three limitations to our study which provide avenues for future research. First, we analyze real-world data from several digital platforms where many decisions can take place in very short time periods. This can limit the scope for definite causal interpretations of the empirical findings. Running controlled field experiments would be desirable. Second, the regression results are consistent with the idea that Amazon's behavior of suppressing the buybox leads to a decrease in the popularity of the products on the Amazon marketplace. The immediate channel of how we expect buybox suppression to reduce demand is by increasing obfuscation in the purchasing process, starting on the product page due to the less visible "add-to-cart" button. We cannot generally exclude the existence of other confounding factors which relate to buybox invisibility. For instance, buybox invisibility might be correlated with missing price information on the search result page for consumers looking for a suitable product or the position of the product in the search results. The buybox coefficient may capture to some extent such confounding effects which might be part of a more general attempt by the marketplace to make the product less prominent. Future work could separate these effects in a controlled experiment. Third, in view of the complex market structure and design features of the Amazon marketplace, we do not provide a formal model that would capture all relevant effects and lead to a definite welfare statement.

Managerial considerations. Removing the buybox makes purchase for consumers more difficult and leads to a reduction in sales. However, Amazon suppresses the buybox in a significant fraction of cases where Amazon is not a seller of the product and the prices for the product are high relative to competitive benchmarks. One rationalization of this strategy is that removing the buybox may help Amazon to maintain sales, as too high prices can be harmful because consumers may turn away from the marketplace and buy somewhere else in the future. It may thus help to build trust among customers in the marketplace and achieve a more profitable balance between Amazon's profit margin and sales volume.

If third-party suppliers learn that a buybox suppression drastically reduces sales, they might be more cautious in their future price setting. One could imagine that, in an ideal world, sellers would not charge too high prices when anticipating that it makes the buybox disappear and thereby drastically reduces demand. This is not the case, however. We observe that Amazon suppresses the buybox in nearly one third of cases for third-party-only products and that sales for products without buybox drop. This leads to a reduction in sales, which would be most problematic for Amazon in the extreme case that all non-buying consumers would turn away from its marketplace. Then, removing the buybox would lead to a drastic reduction of fee revenues from third parties and hence harm Amazon's profits. Although it may seem that removing the buybox can only be profitable in the long run, Amazon may even benefit in the short run.

In the short run, the marketplace's loss in sales for a product without a buybox can be compensated if Amazon can successfully divert consumers toward other products. It seems plausible that a large fraction of consumers might do this once they do not see a buybox for a particular product. Although we cannot measure with the available data whether this is the case, we provide evidence which suggests that showing a product j in a "compare with similar items table" on the site of a product i without a buybox curbs sales of product j. In the best case for Amazon, these more effective recommendations may allow it to keep all consumers buying from the Amazon marketplace. We find that, when the buybox of product i is invisible, Amazon recommends a product j more frequently if it is one of the sellers. Amazon may benefit more from own product sales compared to third-party sales which only yield commission payments but not the full retail margin. Taken together, Amazon may even increase its profits in the short run by removing the buybox of certain products which it does not sell itself. As argued before, in the long run Amazon might benefit from a better reputation of its marketplace on the consumer side and thus more overall visits.

Policy considerations. Our study also provides insights for economic policy. First, Amazon states that it can remove the buybox if it observes pricing practises on the marketplace offer that harms customer trust. However, we find that when Amazon sells a product itself, for the buybox to be visible it does not require as competitive prices and conditions as it requires in instances where it is not a seller. An obvious policy question is whether a hybrid sales platform may *treat* comparable offers of its own and of independent sellers *differently.* In addition, for *transparency*, it would be desirable if the sales platform informed market participants about whether it applies its policies, including those that allegedly aim at protecting customers, only to independent sellers or also to offers of its own sales department.

Second, buybox suppression can create a *search friction for consumers* – we document that it is associated with significantly lower sales. Consumers may not know how to buy the product when the buybox is invisible or – possibly wrongly – take on the belief that the product is not worth buying. Informing consumers that prices of a product are currently relatively high while leaving the convenient way of nevertheless purchasing the product by means of the buybox might be a more efficient way of protecting consumers from buying at possibly too high prices.

Third, significantly reducing the sales potential for sellers that charge higher prices on Amazon compared to other competitive platforms, such as eBay and Walmart, could create incentives similar to *price parity clauses* (PPCs). PPCs have been forbidden as practices that might lead to excessively high commission rates by competition authorities and legislators in various European countries, whereas they are legal in, for instance, the US. Note that we do not claim that the observed behavior leads to effects equivalent to PPC, but only argue that such a risk exists.

Fourth, the Amazon marketplace recommending products that are sold by Amazon particularly often in the comparison table may be justifiable if consumers particularly like to buy products sold by Amazon. An independent marketplace operator might find it optimal to behave in the same manner. However, it is unclear why, in a situation of high prices for product i, consumers should be steered more towards a product j which is sold by Amazon. Yet this is what we find in the data. A rationalization is that steering is more profitable for the hybrid sales platform in this case – which is *consistent with self-preferencing* but difficult to rationalize for an independent marketplace operator. These insights can provide food for thought in view of potential new policies, such as the digital markets act (DMA) in Europe. Among other things, this proposed regulation aims at limiting or prohibiting self-preferencing of hybrid sales platforms that qualify as gatekeepers. We contribute by illustrating a potential subtle channel of behaviour which might constitute self-preferencing. This suggests that for identifying self-preferencing, one should go a step further than simply analyzing whether offers of the integrated entity appear more frequently than comparable offers of independent sellers. In addition, one should check whether offers appear particularly frequently in situations where displaying these offers is particularly profitable for the hybrid sales platform.

References

Bourreau, Marc and Germain Gaudin, "Streaming platform and strategic recommendation bias," *Journal of Economics & Management Strategy*, 2022, *31* (1), 25–47.

Cabral, Luis, Justus Haucap, G Parker G, G Petropoulos, Tommasso Valletti, and Marshall Van Alstyne, "The EU Digital Markets Act," Scientific analysis or review, Policy assessment, KJ-02-21-116-EN-N, Luxembourg 2021.

Calvano, Emilio and Bruno Jullien, "Can we trust the algorithms that recommend products online? A theory of biased advice with no pecuniary incentives and lab evidence," Technical Report, Mimeo 2018.

Chen, Le, Alan Mislove, and Christo Wilson, "An Empirical Analysis of Algorithmic Pricing on the Amazon Marketplace," in "Proceedings of the 25th international conference on World Wide Web" 2016, pp. 1339–1349.

Chen, Nan and Hsin-Tien Tsai, "Steering via Algorithmic Recommendations," *Available at SSRN 3500407*, 2019.

Chen, Yuxin and Song Yao, "Sequential Search With Refinement: Model and Application With Click-Stream Data," *Management Science*, 2016, *63* (12), 4345–4365.

Cornière, Alexandre De and Greg Taylor, "Integration and search engine bias," *The RAND Journal of Economics*, 2014, *45* (3), 576–597.

_ **and** _ , "A model of biased intermediation," *The RAND Journal of Economics*, 2019, *50* (4), 854–882.

De los Santos, Babur and Sergei Koulayev, "Optimizing Click-Through in Online Rankings with Endogenous Search Refinement," *Marketing Science*, 2017, *36* (4), 542–564.

Drugov, Mikhail and Doh-Shin Jeon, "Vertical Integration and Algorithm Bias," *mimeo*, 2019.

Ghose, Anindya, Panagiotis G. Ipeirotis, and Beibei Li, "Designing Ranking Systems for Hotels on Travel Search Engines by Mining User-Generated and Crowdsourced Content," *Marketing Science*, 2012, *31* (3), 493–520.

_ , _ , and _ , "Examining the Impact of Ranking on Consumer Behavior and Search Engine Revenue," *Management Science*, 2014, *60* (7), 1632–1654.

Gómez-Losada, Álvaro and Néstor Duch-Brown, "Competing for Amazon's Buy Box: A Machine-Learning Approach," in Witold Abramowicz and Rafael Corchuelo, eds., *Business Information Systems Workshops*, Springer International Publishing Cham 2019, pp. 445–456.

Greene, William H., *Econometric Analysis*, Pearson Education, 2012.

Hagiu, Andrei and Bruno Jullien, "Why Do Intermediaries Divert Search?," *The RAND Journal of Economics*, 2011, *42* (2), 337–362.

_ and _ , "Search Diversion and Platform Competition," *International Journal of Industrial Organization*, 2014, *33*, 48–60.

_ , Tat-How Teh, and Julian Wright, "Should Amazon be allowed to sell on its own Marketplace?," *The RAND Journal of Economics*, Forthcoming.

Hunold, Matthias and Johannes Muthers, "Resale Price Maintenance and Manufacturer Competition for Retail Services," *The RAND Journal of Economics*, 2017, *48* (1), 3–23.

_ , Reinhold Kesler, and Ulrich Laitenberger, "Rankings of Online Travel Agents, Channel Pricing, and Consumer Protection," *Marketing Science*, 2020, *39* (1), 92–116.

Inderst, Roman and Marco Ottaviani, "Competition Through Commissions and Kickbacks," *The American Economic Review*, 2012, *102* (2), 780–809.

Koulayev, Sergei, "Search for Differentiated Products: Identification and Estimation," *The RAND Journal of Economics*, 2014, *45* (3), 553–575.

Krämer, Jan and Daniel Schnurr, "Is there a need for platform neutrality regulation in the EU?," *Telecommunications Policy*, 2018, *42* (7), 514–529.

Lee, Kwok Hao and Leon Musolff, "Entry Into Two-Sided Markets Shaped By Platform-Guided Search," Technical Report, Mimeo, Princeton University 2021.

Raskovich, Alexander, "Retail Buyer Power Through Steering," *Economics Letters*, 2007, *96* (2), 221–225.

Reimers, Imke and Joel Waldfogel, "Digitization and pre-purchase information: the causal and welfare impacts of reviews and crowd ratings," *American Economic Review*, 2021, *111* (6), 1944–71.

Shen, Bo and Julian Wright, "Why (don't) firms free ride on an intermediary's advice?," *International Journal of Industrial Organization*, 2019, *64*, 27–54.

Ursu, R. M, "The Power of Rankings: Quantifying the Effects of Rankings on Online Consumer Search and Purchase Decision," *Marketing Science*, 2018, *37* (4), 530–552.

Zhu, Feng and Qihong Liu, "Competing with complementors: An empirical look at Amazon.com," *Strategic management journal*, 2018, *39* (10), 2618–2642.

Appendices

A Sample information

Table A1: List of product categories included in data collection

Broad category	Narrow category
Baby	Baby toys
	Strollers
	Baby bottles
	Disposable diapers
Beauty & Personal care	Shampoos
	Men razors
	Hair brushes
	Hair dryers
	Lips makeup
Electronics	Monitors
	Laptops
	Printers
	Home Theater Systems
	Headphones
Garden & Outdoors	Gas Grills
	Shears & Scissors
	Pressure Washers
	Lawn Mowers & Tractors
Health & Households	Thermometers
	Manual Toothbrushes
	First Aid Kits
Home	Irons
	Storage Drawer Carts
	Home Office Chairs
	Vacuums
	Brooms
Kitchen & Dining	Blenders
	Compact Refrigerators
	Microwave Ovens
	Pressure cookers
	Slow cookers
	Coffee Makers
	Toasters
Office products	Pens
	Scissors
Pet supplies	Cats flea & tick control
	Aquariums
	Dogs food
	Dogs toys
Sports & Outdoors	Stand up Paddle Boards
	Tents
	Exercise Bikes
Tools & Home improvements	Drills
	Air Conditioners
	Step Ladders
Toys & Games	Puzzles
	Electronic learning
	Board games
	Kitchen toys
	Dolls

Online Appendix

O1 Additional descriptive statistics

O1.1 Buybox presence across categories

Table O1: Buybox presence across categories

	# Prod.	# Sellers		Price		Buybox presence		% never buybox
Category		Mean	S.d.	Mean	S.d.	Mean	S.d.	
Baby	305	8.20	15.26	56.77	100.10	0.87	0.27	2.62%
Beauty & Pers. care	389	6.45	5.56	16.33	31.87	0.86	0.31	4.63%
Electronics	387	11.24	14.21	305.64	403.51	0.79	0.29	2.07%
Garden & Outdoors	313	6.93	7.90	177.13	191.33	0.84	0.29	2.88%
Health & Households	201	3.93	4.94	16.87	14.86	0.91	0.23	1.00%
Home	372	6.85	8.12	73.95	98.71	0.86	0.27	1.88%
Kitchen & Dining	545	8.12	10.66	82.84	70.97	0.77	0.32	3.30%
Office products	162	12.44	12.23	10.63	7.08	0.86	0.27	0.00%
Pet supplies	322	7.43	7.67	23.51	32.52	0.78	0.35	7.14%
Sports & Outdoors	210	3.73	5.10	209.49	197.37	0.89	0.25	2.86%
Tools & Home improv.	231	7.41	9.09	155.41	205.43	0.83	0.29	3.03%
Toys & Games	414	19.01	20.18	27.95	34.34	0.78	0.32	3.87%
All category	3851	8.85	12.18	97.17	187.01	0.83	0.30	2.94%

NOTE: This table shows summary statistics calculated at the product level, differentiated by broader sales categories. For the calculation of the presence of the buybox, only product-day observations where Amazon was not present (and therefore the buybox could have been suppressed) were taken into account.

O1.2 Buybox choices and inter-platform competition

Our data enables us to study when and how Amazon decides which of the available sellers' product to feature in the buybox. We study how this decision depends on prices on Walmart. We distinguish between situations where Amazon is present as a seller and those where it is not.

When Amazon sells. When Amazon is selling a product itself, we observe in Table O2 that the buybox is always present on the product page. In row 2 from left to right we see that the more third-party sellers price below Walmart, the lower the percentage of time Amazon features itself in the buybox. Row 3 indicates that Amazon's price competitiveness with respect to other sellers on its platform has an analogous decreasing pattern. When no Amazon marketplace seller besides Amazon has a price less than the Walmart price (first column), Amazon almost always chooses to promote its own offer in the buybox, and appears to be the most competitive offer on

Table O2: Amazon Buybox and Walmart price when Amazon sells

	Walmart price is lower than third party which is			
	cheapest	2^{nd} cheapest	3^{rd} cheapest	4^{th} cheapest
% Buybox presence	100%	100%	100%	100%
% AMZ in the buybox	97.59%	88.58%	85.17%	85.38%
% AMZ price < cheapest 3P	93.57%	47.76%	49.60%	57.91%
% 3P in the buybox				
cheapest 3P	77.52%	95.14%	62.80%	52.12%
2^{nd} cheapest 3P	12.48%	4.22%	31.30%	32.03%
3^{rd} cheapest 3P	5.04%	0.39%	3.64%	11.60%
4^{th} cheapest 3P	2.39%	0.10%	1.48%	3.92%
5^{th} cheapest 3P	0.71%	0.15%	0.30%	0.33%

NOTE: This table shows buybox choice statistics when Amazon is selling, depending on where the Walmart price stands among the third-party sellers which have been ordered for each product according to their price (from lowest to highest). The first row shows the % of obs. for which the buybox is present on the product page. The second and third rows show the % of obs. where Amazon's own offer is featured in the buybox and the % of obs. where its offer is the cheapest on the marketplace. The remaining rows show the % of obs. for which each seller in the price rank appears as being featured in the buybox.

its marketplace about 97% of the time. When there is at least one third party-seller with a lower price than Walmart, we see in the bottom part of Table O2 in columns (3) to (5) that sellers with prices lower than Walmart win the buybox in at least 95% of cases (whenever Amazon is not winning it), while the remaining sellers rarely win.

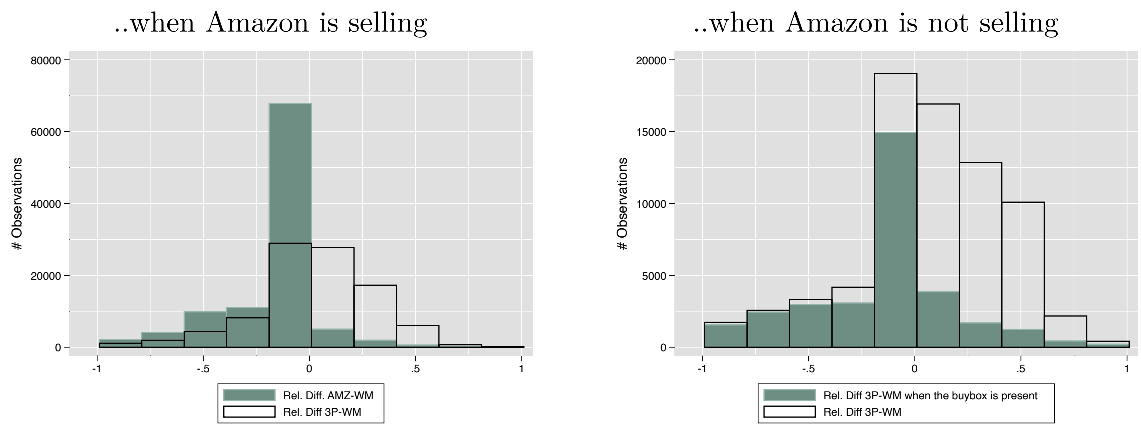
Table O3: Buybox choices with respect to Walmart pricing when Amazon does not sell

	Walmart price is lower than third-party seller who is the			
	cheapest	2^{nd} cheapest	3^{rd} cheapest	4^{th} cheapest
% Buybox presence	14.17%	79.56%	69.73%	69.49%
% 3P in the buybox				
cheapest 3P	78.26%	95.98%	68.61%	60.42%
2^{nd} cheapest 3P	12.88%	3.20%	25.64%	24.71%
3^{rd} cheapest 3P	4.92%	0.53%	2.55%	11.35%
4^{th} cheapest 3P	1.95%	0.18%	1.29%	1.29%
5^{th} cheapest 3P	0.93%	0.04%	1.35%	1.44%

NOTE: This table shows buybox choice statistics when Amazon is not acting as a seller. The columns condition on the relation of the Walmart price to the prices of third-party sellers on Amazon. The first row shows the % of obs. for which the buybox is present on the product page. The remaining rows show the % of obs. for which each seller on Amazon, ranked according to price from low to high, is featured in the buybox.

When Amazon does not sell. Table O3 describes the cases where Amazon does sell a product itself. We observe that if no third-party seller on the Amazon marketplace offers the product cheaper than Walmart, Amazon does not show the buybox on the product page in about 85% of cases (column 1, row 1). If at least one third-party seller offers a lower price than Walmart, the buybox is present in 80 to 90% of cases (columns 2 to 4). Furthermore, sellers featured in the buybox tend to be those that set a lower price than Walmart.

Figure O1: Distribution of prices on the Amazon marketplace sellers relative to the Walmart price



NOTE: The left figure shows in green the distribution among all observations where Amazon is present as a seller of relative price differences between Amazon itself and Walmart, computed as $\frac{\text{AMZ price-WM price}}{\text{AMZ price}}$. The transparent outline is the distribution for these observations of the relative price differences between the cheapest third-party seller available and Walmart, computed as $\frac{\text{Min 3P price-WM price}}{\text{Min 3P price}}$. The right figure shows in green the distribution among all products where Amazon is not present and which have a buybox of relative price differences between the cheapest third-party seller available and Walmart. The distribution in transparent also contains observations for which the buybox is not displayed.

Thus, Amazon appears to only choose not to promote any sellers in cases where it does not sell the product, while the absence of any third-party sellers who are more competitive than the Walmart price seems to be correlated with non-promotion on the product page.

Price dispersion between Amazon sellers and Walmart. On the left in Figure O1, we observe that for products where Amazon sells, the relative difference between Amazon's own price and the Walmart price is almost always negative. Apparently, Amazon tends to set prices equal to or lower than Walmart. For these observations, the relative price difference between the least expensive third-party seller on Amazon and Walmart appears to be positive more often. Moreover, for these observations, we see from Table O3 that Amazon almost always self-selects itself in the buybox, with a price very likely to be lower than or equal to the Walmart price given the high mass located right above zero on the left histogram of Figure O1.

The right histogram in Figure O1 depicts observations where Amazon is not a seller. The third-party sellers' relative price distribution with respect to Walmart (transparent bars) show a high mass right of zero. If we restrict ourselves to cases where Amazon shows no one in the buybox (dark bars), the positive mass decreases, thus confirming Amazon's tendency to choose not to promote any seller whenever they are not price competitive enough with respect to Walmart.

O2 Additional regression results

O2.1 Moderating factors for buybox suppression

Our regression analysis in Subsection 6.2 indicates that Amazon tends to remove the buybox when lower prices are to be found off-Amazon. To investigate whether Amazon's buybox decision is affected by other factors, we run the specifications in Table O4. In column (1), we differentiate based on the length of Amazon's absence. We distinguish short time periods of absence (less than one month), which might reflect periods of stock-outs for Amazon, from periods of absence longer than one month. The latter may indicate that Amazon exited the product market because there is no longer any profit in sales. The remaining cases cover products which Amazon has never sold according to our data, either by choice (due to lack of profitability) or by inability to do so. We find that the coefficient associated with the indicator that Walmart is more price competitive to be of the same order of magnitude for these three distinctions, suggesting that Amazon's buybox policy in relation to the Walmart price is not affected by the profitability of the markets for Amazon as a seller. In column (2), we interact the Walmart competitiveness variable with the number of sellers available for the product, which may indicate the product's popularity for third-party sellers. We do not see a differential effect in Amazon's buybox decision with respect to the number of sellers. In column (3), we see that Amazon is more likely to suppress the buybox for products of categories where Amazon has more products available for purchase in its marketplace. The number of items available within a product category broadly measures the popularity of the category on Amazon, which is also likely correlated with the number of highly substitutable products in the marketplace. Finally, we observe that the overall price level of products does not affect Amazon's buybox choices driven by Walmart price competitiveness.

Table O4: Moderating factors

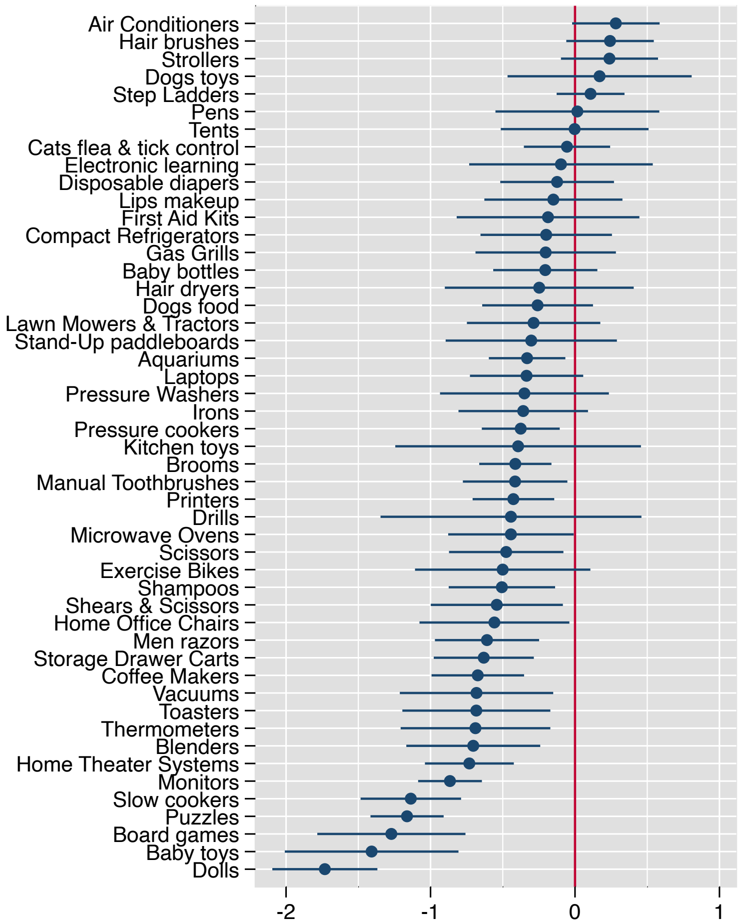
	(1)	(2)	(3)	(4)
- Lag log sales rank	0.03***	0.03***	0.03***	0.03***
	(0.00)	(0.00)	(0.00)	(0.00)
FBA seller present	0.07***	0.07***	0.07***	0.07***
	(0.02)	(0.02)	(0.02)	(0.02)
# Sellers	-0.03***	-0.01	-0.03***	-0.03***
	(0.01)	(0.01)	(0.01)	(0.01)
Min price	-0.02***	-0.02***	-0.02***	-0.02***
	(0.00)	(0.00)	(0.00)	(0.00)
Available at WM/eBay	0.01	0.00	0.00	0.00
	(0.02)	(0.02)	(0.02)	(0.02)
I[Min price > MRSP]	-0.08***	-0.09***	-0.08***	-0.08***
	(0.01)	(0.01)	(0.01)	(0.01)
I[Min price > avg. subst. price]	-0.04***	-0.04***	-0.04***	-0.04***
	(0.01)	(0.01)	(0.01)	(0.01)
I[Min price > min(WM,eBay)]		-0.18***	-0.21***	-0.22***
		(0.02)	(0.03)	(0.02)
I[Min price > min(WM,eBay)] * AMZ absent < 1 month	-0.22***			
	(0.02)			
I[Min price > min(WM,eBay)] * AMZ absent > 1 month	-0.23***			
	(0.02)			
I[Min price > min(WM,eBay)] * AMZ never sold	-0.20***			
	(0.03)			
I[Min price > min(WM,eBay)] * # Sellers		-0.03***		
		(0.01)		
I[Min price > min(WM,eBay)] * # subst.			-0.00	
			(0.00)	
I[Min price > min(WM,eBay)] * Min price				0.00
				(0.00)
Constant	1.19***	1.18***	1.19***	1.19***
	(0.05)	(0.05)	(0.05)	(0.05)
Observations	112,491	112,491	112,491	112,491
R-squared	0.82	0.82	0.82	0.82

Notes: Dependent variable: buybox availability. Unit of observation: product-date. Linear regressions include fixed effects at the product, date, and hour of the day level. Standard errors are robust to heteroscedasticity and adjusted for serial correlation inside clusters.

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

O2.2 Buybox suppression and demand reduction by product category

Figure O2: Buybox suppression and demand reduction by category



NOTE: The figure shows the worsening of the sales rank by product category following a buybox suppression. For instance, products that are in the "Blender" category worsen in their sales ranks by 50% if the buybox is suppressed. The effects are calculated by a regression as in Table 5, column (2) where the indicator "No buybox" is interacted with indicators for the product category; these coefficients are further divided by the 1 minus the coefficient of the log lag sales rank.

O3 Robustness checks

In subsection 6.2 we showed that controlling for many other factors, Amazon makes showing the buybox dependent from prices at Walmart and eBay. Following up the main results, we first employ different specifications for the price differences across platforms. Second, we investigate potential selection biases by also including products on dates where Amazon was also selling the product. Finally, we study whether our results are robust to unobserved shocks on Amazon in the same product category.

O3.1 Alternative definition of price differences across platforms

In the main regression analyses we measure price deviations on Walmart by an indicator variable taking value one whenever the minimum third-party price on Amazon is higher than the minimum price between Walmart and eBay.[23] Taking the indicator variable allows us to compare products with different absolute average price levels and decreases the potential spurious correlation between price and the deviation, while being parsimonious with our specification and having a straightforward interpretation. In the following, we discuss the robustness of this approach. In column (1), we show that we get qualitatively the same result when taking the relative price differences: The larger the difference between the minimum third-party price on Amazon to the minimum off-Amazon price, the less likely the buybox is shown. In column (2), we distinguish between positive and negative relative differences. Consistent with our main specification, we observe that the buybox is less likely shown the larger the positive relative price differences is, i.e., if prices on competing sales channels are cheaper. There is no effect for negative deviations, consistent with the view that Amazon takes the competing price levels as an important factor for determining when not to show the buybox. Finally, we study whether Amazon takes all levels of price differences into account in the same way. To test this, we partitioned the positive deviations into whether they were below 5%, between 5 and 10%, 10 and 20%, or beyond, and built corresponding indicator indicator variables which we included in column (3). We find that there is an increasing pattern of the coefficients until deviations above 20%. This suggests that Amazon has some tolerance for smaller deviations below 10%, while for larger deviations the effect essentially becomes constant.

O3.2 Including observations when Amazon is present

In our regression analyses in Table 3, we focused only on products which were currently not sold by Amazon itself, but only by third parties. This was motivated by the observation in Table 2 that the buybox was always present in our sample when Amazon sells. For this purpose we re-estimate the specifications employed in Table 3 using the full sample of product-date combinations, taking into account observations when Amazon was present as a seller, and adding an indicator "Amazon presence" when this was the case. In Table O6 we show that

[23] For Walmart we focused on the Walmart buybox price led by the assumption that consumers would pay most attention to this price. We also considered the minimum price available for each product on Walmart as a benchmark and find that our findings qualitatively unaffected. The respective results are available upon request.

this does not alter our results qualitatively. First we note that the coefficient of the indicator variable "Amazon presence" is highly significant and positive. We also see that the off-Amazon price negatively affects the presence of the buybox. However, in column (4) we see that this only applies when Amazon is not present as a seller.

Table O5: Different specifications of price differences

	(1)	(2)	(3)
- Lag log sales rank	0.03***	0.03***	0.03***
	(0.00)	(0.00)	(0.00)
FBA seller present	0.08***	0.08***	0.07***
	(0.02)	(0.02)	(0.02)
# Sellers	-0.03***	-0.03***	-0.03***
	(0.01)	(0.01)	(0.01)
Min price	-0.02***	-0.02***	-0.02***
	(0.00)	(0.00)	(0.00)
Available at WM/eBay	-0.02	-0.02	0.00
	(0.03)	(0.03)	(0.03)
Rel.diff Min-MRSP	-0.05***	-0.04***	-0.04***
	(0.01)	(0.01)	(0.01)
Rel.diff Min-Avg. subst. price	-0.09***	-0.07***	-0.07***
	(0.03)	(0.03)	(0.02)
Rel.diff Min-min(WM, eBay)	-0.14***		
	(0.03)		
Rel.diff Min-min(WM, eBay) ≤ 0		-0.03	
		(0.02)	
Rel.diff Min-min(WM, eBay) > 0		-0.42***	
		(0.04)	
I[$0 <$ Rel.diff Min-min(WM, eBay) ≤ 0.05]			-0.14***
			(0.02)
I[$0.05 <$ Rel.diff Min-min(WM, eBay) ≤ 0.1]			-0.22***
			(0.02)
I[$0.1 <$ Rel.diff Min-min(WM, eBay) ≤ 0.2]			-0.26***
			(0.02)
I[Rel.diff Min-min(WM, eBay) > 0.2]			-0.26***
			(0.02)
Constant	1.09***	1.11***	1.13***
	(0.05)	(0.05)	(0.05)
Observations	112,491	112,491	112,491
R-squared	0.81	0.81	0.82

Notes: Dependent variable: Buybox availability. Unit of observation: Product-Date. Linear regressions include fixed effects at the product, date, and hour of the day level. Standard errors are robust to heteroscedasticity and adjusted for serial correlation inside clusters.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table O6: Determinants of buybox availability (including observations when Amazon sells)

	(1)	(2)	(3)	(4)
- Lag log sales rank	0.02***	0.02***	0.02***	0.02***
	(0.00)	(0.00)	(0.00)	(0.00)
Amazon sells product	0.63***	0.63***	0.59***	0.45***
	(0.01)	(0.01)	(0.01)	(0.02)
FBA seller present	0.03***	0.03***	0.02***	0.02***
	(0.01)	(0.01)	(0.01)	(0.01)
# Sellers	-0.01***	-0.01***	-0.01***	-0.01***
	(0.00)	(0.00)	(0.00)	(0.00)
Min price	-0.01***	-0.01***	-0.01***	-0.01***
	(0.00)	(0.00)	(0.00)	(0.00)
Available at WM/eBay	0.01	0.01	0.01	0.00
	(0.02)	(0.02)	(0.02)	(0.02)
I[Min price > MRSP]	-0.10***	-0.10***	-0.08***	-0.07***
	(0.01)	(0.01)	(0.01)	(0.01)
I[Min price > Avg. subst. price]		-0.03***	-0.02***	-0.02***
		(0.01)	(0.01)	(0.01)
I[Min price > min(WM, eBay)]			-0.12***	
			(0.01)	
I[Min price > min(WM, eBay)] * AMZ				0.00
				(0.00)
I[Min price > min(WM, eBay)] * No AMZ				-0.28***
				(0.01)
Constant	0.77***	0.78***	0.80***	0.89***
	(0.03)	(0.03)	(0.03)	(0.03)
Observations	245,324	245,324	245,324	245,324
R-squared	0.81	0.81	0.82	0.83

Notes: Dependent variable: Buybox availability. Unit of observation: Product-Date. Linear regressions include fixed effects at the product, date, and hour of the day level. Standard errors are robust to heteroscedasticity and adjusted for serial correlation inside clusters.

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

O3.3 Moderation factors for recommendations

Table O7: Moderating factors for recommendations

	(1)	(2)
- Lag log sales rank j	0.03***	0.03***
	(0.00)	(0.00)
FBA seller present j	0.00	0.00
	(0.00)	(0.00)
# Sellers j	-0.00	-0.00
	(0.00)	(0.00)
Min price j	-0.00***	-0.00***
	(0.00)	(0.00)
I[Min price j > eBay price j]	0.01	0.01
	(0.00)	(0.00)
I[Min price j > Avg. subst. price i]	-0.01	-0.01
	(0.00)	(0.00)
Buybox invisible j	-0.34***	-0.34***
	(0.01)	(0.01)
AMZ sells product j	0.05***	-0.01
	(0.01)	(0.01)
AMZ sells product j * Buybox invisible i	0.08***	
	(0.02)	
AMZ sells product j * Buybox invisible i * # AMZ subst.	-0.01***	
	(0.00)	
AMZ sells product j * # AMZ subst.	-0.01***	
	(0.00)	
AMZ sells product j * Buybox invisible i * I[# AMZ subst. $\in [1,3]$]		0.07***
		(0.02)
AMZ sells product j * Buybox invisible i * I[# AMZ subst. $\in [4,6]$]		0.05***
		(0.01)
AMZ sells product j * Buybox invisible i * I[# AMZ subst. > 7]		-0.00
		(0.01)
Constant	0.65***	0.64***
	(0.01)	(0.01)
Observations	1,647,259	1,647,259
R-squared	0.52	0.52

Notes: Dependent variable: indicator taking value 1 if product p is recommended on product i's page. Unit of observation: Product pair $p - i$-Date. Linear regressions include fixed effects at the product pair, product-i-date, and hour of the day level. Standard errors are robust to heteroscedasticity and adjusted for serial correlation inside clusters.

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