What's In the Box? The Effect Of Self-Preferencing On Amazon Marketplace

Florian Dendorfer & Regina Seibel*

January 28, 2026

Abstract

We examine self-preferencing in Amazon's "Similar item to consider" box. We document a stark difference across domains: while Amazon frequently recommends alternatives to Amazon Basics (AB) products on the US website, it systematically suppresses these recommendations on the Canadian website. Leveraging this discontinuity, we estimate the causal effect of algorithmic exclusion due to self-preferencing. We find that non-AB products that merit a recommendation in the US but are denied one in Canada suffer a 11% decrease in the volume of sales. This finding quantifies the significant market distortion caused when gatekeepers favor their own inventory over rival goods.

*Dendorfer: Competition Bureau Canada. Seibel: Rotman School of Management, University of Toronto. Email: florian.dendorfer@cb-bc.gc.ca, and regina.seibel@rotman.utoronto.ca. We are grateful to Nicholas Janota and Heski Bar-Isaac for their valuable comments. The views expressed are from a personal perspective and do not represent the views or the position of the Commissioner of Competition of the Government of Canada.

1 Introduction

A central policy concern in the digital economy is "self-preferencing" – the practice by which a dual-role platform favors its own products over alternatives from competing sellers. While theoretical models suggest such behavior can be anti-competitive (De Corniere and Taylor, 2019; Dendorfer, 2024; Hagiu, Teh and Wright, 2022; Padilla, Perkins and Piccolo, 2022; Tirole and Bisceglia, 2023), empirically identifying self-preferencing is notoriously difficult. As Peitz (2023) notes, "it can be challenging to detect self-preferencing bias as opposed to legitimate differential treatment" (p.315). In other words, a platform might recommend its own product because it is favored by the algorithm, or simply because consumers prefer them over alternatives.

We overcome this identification challenge by exploiting a stark discontinuity in Amazon's algorithmic assignment of the "Similar item to consider" recommendation widget, or "box" for brevity, across domains. By scraping data for over 13,000 products across Amazon's US and Canadian domains, we document a striking pattern: while Amazon frequently recommends both private-label (Amazon Basics, or AB) and competing (non-AB) products in the US, it never recommends non-AB products in Canada. Furthermore, while AB product pages in the US sometimes feature competitor recommendations, AB pages in Canada are effectively "shielded" from such competition, as no recommendation is shown on their product page. This cross-border variation – where the algorithm effectively "turns off" recommendations for non-AB goods in one domain – provides a quasi-experimental setting to isolate the causal effect of recommendations, and, by extension, of the loss of recommendations due to self-preferencing.

When constructing our sample, we deliberately selected product groups that include AB items to be able to observe differential treatment of AB versus non-AB products. We confirm that many observable product characteristics, such as ratings and prices, are comparable across domains. Products have fewer reviews and lower sales ranks in Canada which is consistent with the smaller size of the Canadian market. In both domains, a product's sales rank, i.e., the relative rank of its volume of sales compared to other products in the same category, is correlated with whether the product is an AB product, labeled Amazon's Choice, sold by Amazon, and with its price.

We provide detailed evidence of a higher level of self-preferencing in Canada by describing how box assignment in Canada differs from box assignment in the US. We find that the single most important predictor of a product being recommended or not featuring a competitor box in Canada is simply whether a product is AB branded. The relatively lower share of AB products among top sellers in Canada is suggestive for AB products being less popular there relative to non-AB products. This is at odds with consumer preferences being the main driver of the differences in box assignment across domains. In line with this argument, we also find that the number of units sold plays no role in the box assignment in Canada, while in the US, more popular products are more likely to be recommended and less likely to feature a competitor box on their product page.

Next, we assess whether this form of self-preferencing effectively steers consumer behavior. We assume that, absent differences in the box assignment across domains and conditional on US sales, the product's volume of sales in Canada is orthogonal to whether a product is recommended in the US. We show that, under this assumption, we can recover the causal effect of a non-AB product being recommended through the box on its sales, as well as the causal effect of a competitor product being recommended on an AB product page on the sales of AB products. We find that non-AB products that merit a recommendation in the US but are denied one in Canada experience a sales rank that is approximately 40% worse (higher) and a sales volume that is 11% lower than that of their non-recommended peers. Conversely, we find that Amazon's own products benefit from the removal of the box from their product pages – though these estimates are noisier due to sample size constraints.

We then explore heterogeneity in the effect of self-preferencing on non-AB products along various dimensions. As expected, the sales loss due to algorithmic exclusion from recommendations is larger the more often a product is recommended in the US. It is also most pronounced for less popular "long tail" products, for which algorithmic discovery is a critical driver of demand. Finally, we show that products whose US product pages feature a box with another non-AB product benefit from self-preferencing in the Canadian domain. This is because, in Canada, the box may be absent altogether if the non-AB product is not replaced by an AB product.

We posit that the difference in self-preferencing across domains is driven by two factors: economic incentives and regulation. First, AB products enjoy less organic popularity in Canada than in the US, potentially incentivizing the platform to use algorithmic tools to protect its private-label products. Second, the regulatory environment in the US has historically been harsher towards anti-competitive practices of large tech companies.

The self-preferential behavior documented in this paper has substantial aggregate economic implications. A back-of-the-envelope calculation suggests that if the US market were subject to the same degree of self-preferencing observed in Canada, total non-AB sales would decline by approximately 2.2%. Given the scale of the Amazon marketplace in the US, this distortion corresponds to a loss of more than 160 million unit sales annually for non-AB products. By quantifying this cost of self-preferencing our study highlights the significant market distortions that can arise when gatekeepers favor their own inventory.

The remainder of this paper is organized as follows. Section 2 reviews the related literature. Section 3 describes the data. Section 4 outlines the estimation strategy and presents the results. Section 5 explores heterogeneity of the main result. A brief discussion of the results in Section 6 follows. Section 7 concludes.

2 Literature

The incentives for platforms to engage in self-preferencing, as well as its impact on the platform-run marketplace, have been extensively explored in the theoretical literature (De Corniere and Taylor, 2019; Dendorfer, 2024; Hagiu et al., 2022; Padilla et al., 2022; Tirole and Bisceglia, 2023; Zennyo, 2022; Zou and Zhou, 2024). This literature largely concludes that self-preferencing is anti-competitive and harms consumers. De Corniere and Taylor (2019) show that a vertically integrated intermediary biases its recommendations in favor of its own product, leading to a more mismatches and harming consumers, particularly when price is the predominant dimension of competition. Hagiu et al. (2022) find that self-preferencing can grant a platform a virtual monopoly by enabling it to foreclose rivals. In a similar vein, Dendorfer (2024) shows that self-preferencing increases a platform's pricing power over its own product, exacerbates double marginalization, and therefore harms consumers through higher prices. Tirole and Bisceglia (2023) argue that self-preferencing can be eliminated, and market efficiency restored, through a Pigouvian-style policy. Zou and Zhou (2024) and Zennyo (2022) are exceptions within this literature: Zou and Zhou (2024) find that self-preferencing induces third-party sellers to lower prices, while Zennyo (2022) show that self-preferencing leads to a reduction in the platform's commission fee, which is passed on to consumers. Note that whether self-preferencing is ultimately profitable for the platform may depend on a range of factors, including the demand response. For instance, if self-preferencing serves to exploit captive consumers, overall platform usage may decline, thereby reducing platform profits (Dendorfer, 2024; Padilla et al., 2022).

A fast-growing empirical literature seeks to identify evidence of self-preferencing and document its effects on Amazon Marketplace (Chen and Tsai, 2024; Farronato, Fradkin and MacKay, 2023, 2025; Lee and Musolff, 2021; Raval, 2022; Waldfogel, 2024). Much of this work is surveyed in Etro (2024). Farronato et al. (2023) use consumer search data from Amazon Marketplace and find that Amazon-branded products rank higher in search results, i.e., have a lower search rank, than comparable non-Amazon-branded products, though it remains unclear whether this reflects self-preferencing or consumers' inherent preferences for Amazon's own brands. In a related subsequent study, Farronato et al. (2025) conduct a field experiment that manipulates the presence of Amazon-branded products and, using a structural model, calculate that assigning the on average same search rank to Amazon-branded and non-branded products leads to a small decrease in consumer surplus, which may suggest that rank differences are not driven by self-preferencing.

Waldfogel (2024) finds that Amazon-branded products have lower search ranks and argues that self-preferencing must account for at least part of this effect. He also notes that the average sales rank of Amazon's own products worsened significantly after the company was designated a "gatekeeper" under the Digital Markets Act, possibly due to reduced self-preferencing in an effort to comply with regulatory requirements.

In contrast to these studies, we document differential treatment of the *same* product across geographic domains, with respect to product recommendations. We show that products of Amazon's Amazon Basics (AB) brand are treated differently from products of other brands in the Canadian domain compared to the US domain of Amazon Marketplace, a difference that cannot be explained by observable factors such as the price or the average rating. Similar to Waldfogel (2024), we suspect that this difference is due, at least in part, to varying levels of antitrust scrutiny between countries.

Lee and Musolff (2021) and Raval (2022) study products sold both by Amazon and by third-party vendors, finding that Amazon's offers are disproportionally selected into the "Buy Box", i.e., it is the default seller. Lee and Musolff (2021) estimate a structural model of Amazon's marketplace and, through one of their main counterfactuals, determine that the bias toward Amazon largely reflects consumers' genuine preferences for Amazon as a seller, rather than unfair steering. In contrast, we document that AB products receive more favorable treatment in product recommendations in the Canadian domain than in the US domain, a difference that we argue is not attributable to consumer preferences.

Lastly, Chen and Tsai (2024) examine "Frequently Bought Together" product recommendations on Amazon Marketplace. They find that a product is, on average, eight percentage points less likely to be recommended when Amazon stocks out of a complementary product, thus losing the incentive to promote the product. Similarly, we find that Amazon favors its own, AB products over non-AB products in the "Similar item to consider" recommendations. In addition, we are able to estimate the change in the volume of sales products experience as a result of self-preferencing.

3 Data

This section is structured as follows. Section 3.1 describes our data collection process and the contents of our data sample. In Section 3.2, we showcase the data by replicating findings from the existing literature. Section 3.3 documents evidence of self-preferencing in the "Similar item to consider" box.

3.1 Data Collection

Between January 24 and March 11, 2025, we scraped product pages for 13,069 Amazon Standard Identification Numbers (ASINs) from both the US (Amazon.com) and Canadian (Amazon.ca) domains of the Amazon Marketplace website. We obtained the ASINs from the first search result page on Amazon Marketplace for each of 200 search terms (e.g., "Clothing Steamers") that we suspect includes at least one product sold under Amazon's largest private label, Amazon Basics (AB). Although Amazon sells under other brands as well, virtually all private label products in our data sample are AB products.[1]

We collected the Buy Box price (in USD and CAD, respectively) excluding delivery fees, the number of reviews, and the average rating on a 1-to-5 star scale. Additionally, we recorded whether a product was designated as "Amazon's Choice", whether it was an AB product, and its sales rank.

A single product can have multiple sales ranks, one for each category it belongs to. Typically, we observe a broad category (e.g., "Electronics") along with at least one more specific category (e.g., "Home Cinema Cables"). In total, we identified 495 categories, many of which include AB products. To give the reader a sense of the categories included in our dataset, Table A.1 in the Appendix lists the twenty largest categories, along with the total number of products and the number of AB products in each.

Apart from sales ranks, we collected sales data. Product pages occasionally provide a lower bound on the volume of sales, such as "700+ bought in the past month" or "2k+ bought in the past month." There are 36 different such sales indicators: x00+, xK+, x0K+, and x00K+ for $x \in \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$. We observe sales information of this kind for 55% of observations in our data sample. We acknowledge that, due to censoring, the reported sales data are systematically understated, and we view them

[1] A small number of products are sold under Amazon's Rivet and Amazon Commercial brands. We treat them as AB products in our analysis.

as conservative estimates of actual sales volumes. Nonetheless, we believe the sales data are of sufficiently high quality to be used in our analysis in Section 4.

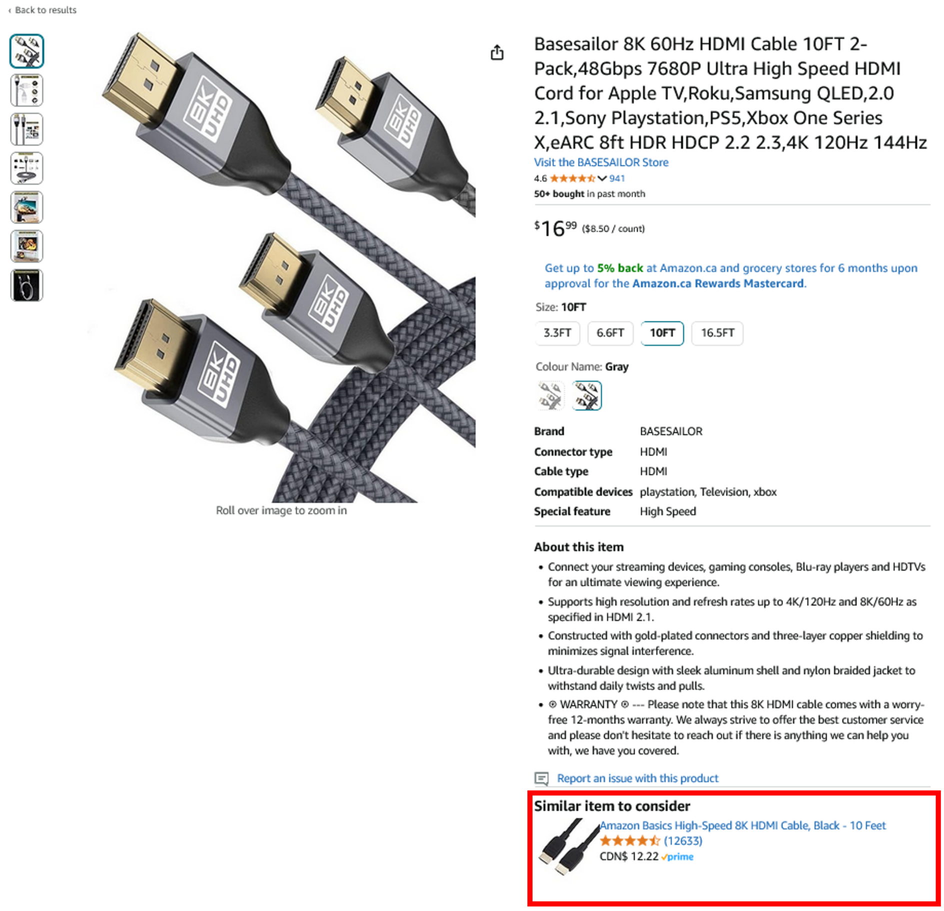


Figure 1: “Similar item to consider” box (red frame added).

Lastly, we record whether a product page includes a “Similar item to consider” box, sometimes labeled “Consider a similar item” or “Competitively priced item”, along with the ASIN of the product featured in the box. Figure 1 illustrates how this box appears on a product page. The feature is widespread: 32% of Canadian product pages in our sample contain the box, compared to 53% of US pages. Among products in our sample, 3% appear in the box in Canada, while this share is 8% for the US. While this specific widget is the focus of our study, it represents only one channel of potential self-preferencing. Amazon controls numerous other discovery mechanisms – including search ranking algorithms, the “Amazon’s Choice” badge, and the Buy Box, to name just a few – which may also be subject to similar preferential treatment.

For many products, we can identify both the (Buy Box-winning) seller and the shipper. All products for which we register the shipper are shipped or "fulfilled" by Amazon (FBA).[2] Similarly, Gutierrez (2021) reports that, in his dataset, the vast majority of products (81%) are shipped by Amazon. In our dataset, the seller with the largest number of products is by far Amazon itself. Third-party sellers tend to be small; half of them sell two or fewer products. In our sample, 4% of products are AB-branded since we deliberately sampled categories with AB presence. The share of AB products among all products sold on Amazon in the US overall is most likely below 1%.[3] With 28%, the share of first-party products in our sample is consistent with public numbers on revenue shares.[4]

3.2 Data and Descriptives

Table 1 presents sample averages for AB and non-AB products in the Canadian and US domains for various variables. AB products differ from non-AB products. On average, they are rated slightly higher and are about 30% cheaper than non-AB products in both domains. This aligns with findings by Farronato et al. (2023), who report similar statistics for their data sample. While Farronato et al. (2023) find that Amazon-branded products have about four times more reviews than other products, our sample shows that AB products have, on average, around eight times more reviews compared to non-AB products.

Compared to non-AB products, AB products have on average 79% and 67% lower (better) sales ranks in Canada and the US, respectively. This gap is substantially wider than the 24% difference reported in Farronato et al. (2023). Note that the raw sales ranks in our data are generally large, likely because our sample includes broad categories such as "Clothing" or "Tools & Home Improvement." However, even when we focus on the category in which each product attains its lowest rank, i.e. the "smallest" category, average ranks decrease in both domains, while the sales-rank gap between AB and non-AB products remains roughly unchanged. These descriptive

[2] For the Canadian domain, we successfully scraped the shipper information for 78% of product pages. Although we were unable to scrape the shipper information for the US domain, we believe that most, if not all of the products are FBA too.

[3] See https://www.momentumcommerce.com/amazons-private-label-market-share-shrinks-by-6-year-over-year-in-q1-2024/. Note however, that numbers reported refer to revenue shares. Amazon private labels are likely to be higher value products and thus we expect the share of products sold to be lower.

[4] See https://www.digitalcommerce360.com/amazon-ecommerce-facts-and-statistics.

	CA		US	
	non-AB	AB	non-AB	AB
Avg. rating	4.45 stars	4.48 stars	4.49 stars	4.55 stars
Review count	5,626	46,792	6,468	56,468
Avg price	62.05 CAD	42.18 CAD	41.12 USD	27.21 USD
Rank	26,259	5,401	35,600	11,654
Rank ("smallest" category)	14,504	3,586	20,644	7,143
No. products	12,546	523	12,546	523

Table 1: Summary statistics by domain and product type (AB vs non-AB)

statistics suggest that consumers might plausibly prefer AB products over other alternatives. Since AB products may enjoy greater prominence due to legitimate popularity, establishing self-preferencing requires distinguishing between algorithmic bias and organic consumer demand.

Table 1 also highlights that the Canadian and US domains are broadly similar based on observables, though the Canadian market is unsurprisingly smaller. The average sales rank and review count in Canada are 26% and 14% lower, respectively, than in the US. Differences in average nominal prices largely reflect the exchange rate.

Next, we investigate whether the average sales rank of AB products differs from that of non-AB products conditional on observables. For this and all subsequent regressions, we reweigh the data to account for the fact that a single product may appear in multiple categories. We define observations at the ASIN-category-domain level and weight each observation by the inverse of the number of categories associated with that ASIN in a given domain.[5] Accordingly, standard errors are clustered at the category-domain level. Our final dataset consists of 34,172 observations covering 13,069 unique products in each domain.

Table 2 presents the estimates from a regression of the log sales rank on product characteristics, including indicators for AB branding, "Amazon's Choice" status, and first-party (1P) selling, as well as log price, average rating, log review count, and category-domain fixed effects. We find that, all else equal, AB products have a 30% lower (better) sales rank.[6] Other coefficients align with economic intuition. Ceteris paribus, a 1% increase in price is associated with a 0.31% higher (worse) sales rank. The "Amazon's

[5]Let n_{am} be the number of observations for ASIN a in domain m. We weight each observation by n_{am}^{-1} in all regressions.

[6]This finding is consistent with Waldfogel (2024), who also documents that Amazon's private label products enjoy higher search ranks, conditional on similar observables, compared to competing goods.

Choice" label is linked to a substantial 70% reduction in sales rank. Higher ratings and higher review counts both correlate with better sales ranks, likely because these metrics reflect past sales performance. Surprisingly, first-party selling is associated with a slightly higher sales rank. We emphasize that these coefficients describe correlations rather than causal relationships; the AB coefficient, in particular, may reflect unobserved quality or consumer preferences rather than self-preferencing.

Dependent Variable:	log(rank)
Variables	
Amazon Basics	-0.35***
	(0.09)
Amazon's Choice	-1.2***
	(0.06)
1P	0.35***
	(0.05)
log(price)	0.31***
	(0.05)
rating	0.34**
	(0.15)
log(review count)	-0.14
	(0.15)
rating × log(review count)	-0.09**
	(0.04)
Fixed-effects	
category-domain	Yes
Fit statistics	
Observations	34,172
R^2	0.780
Within R^2	0.440

Clustered (category-domain) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 2: Determines of the sales rank.

3.3 Self-Preferencing in Canadian Product Recommendations

We now turn to describing how the assignment of the "Similar item to consider" box varies across domains. In our sample of 13,069 products, we observe a striking discontinuity in the Canadian domain: none of the non-AB products are promoted in the recommendation box. In contrast, 5.2% of non-AB products receive such recommendations in the US.[7] Among AB products, recommendation rates are high in both domains: 81.1% in Canada and 73.8% in the US.

A similar pattern emerges regarding which products host the recommendation box. In the Canadian domain, none of the AB product pages feature a "Similar item to consider" box (referred to as "having the box"). By contrast, in the US, 29.1% of AB product pages display the box. For non-AB products, the share of pages having the box is 33.7% in Canada and 54.1% in the US. These descriptive statistics are summarized in Table 3.

	Full Sample		AB products		non-AB products	
	US	CA	US	CA	US	CA
Recommended	7.9%	3.2%	73.8%	81.1%	5.2%	0.0%
Has Box	53.1%	32.3%	29.1%	0.0%	54.1%	33.7%
N	13,069	13,069	523	523	12,546	12,546

Table 3: Share of recommended products / products having the box by domain and product type (AB vs non-AB)

To provide more detail on how box assignment differs across borders, we regress the binary outcome of being recommended (or having a box) in Canada on the corresponding status in the US and an AB indicator. We also add category-level fixed effects to allow for the possibility that Amazon Canada's use of the box differs from Amazon US's use depending on the category. Reportedly, Amazon determines recommendations based on sales rank, price, the number of reviews and the rating.[8] To ensure our estimates do not simply capture correlations between AB status and observable quality, we add a vector of controls in a second specification: log sales rank, Amazon's Choice status, 1P status, log price, rating, and log review counts (for both domains).

[7] We note that these figures likely understate the total prevalence of recommendations, as we only observe boxes appearing on the product pages of items within our scraped sample.

[8] See `https://sites.duke.edu/theamazonproject/algorithms-and-amazon/`.

Column (1) of Table 4 shows that being an AB product increases the probability of being recommended in Canada by 76 percentage points. This probability increases by a further 10 percentage points if the product is also recommended in the US. In contrast, the probability for any non-AB product is zero, regardless of its US status. Importantly, AB status remains the single dominant predictor of recommendation in Canada even after controlling for sales rank and other characteristics in Column (2).

Column (3) reports the results for having the box. All else equal, non-AB product pages in Canada are 24 percentage points more likely to display a box if they also display one in the US. However, AB products have a 42 percentage point lower probability of having the box compared to non-AB products without the box in the US, and a 65 percentage point lower probability compared to non-AB products that do have a US box. Again, the inclusion of controls in Column (4) confirms that AB status is the primary determinant of whether a product page is shielded from competitor recommendations in Canada.

The primary challenge in identifying self-preferencing is the possibility that AB products are simply preferred by consumers, warranting their algorithmic promotion. However, the magnitude of the disparity – an increase in recommendation probability of 76 to 86 percentage points for AB products – is difficult to explain by consumer preferences alone.

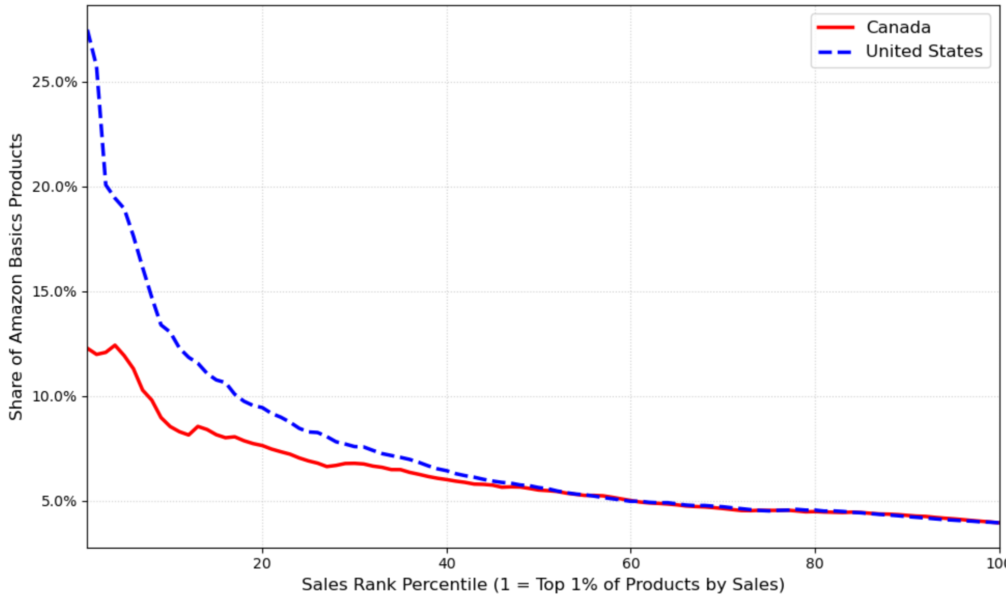
If popularity were the main driver, we would expect the product's sales rank to be a strong predictor of recommendation status, attenuating the coefficient on the AB dummy. However, Column (2) of Table 4 shows no significant correlation between a product's log sales rank and its recommendation status in Canada, conditional on AB status. While we cannot interpret the sales rank coefficient causally due to endogeneity, the direction of any bias reinforces our argument. If recommendations causally boost sales (lowering the sales rank), we would expect reverse causality to bias the coefficient on log sales rank downward (i.e., turn it negative). In this way, a coefficient that is statistically indistinguishable from zero contradicts the notion that Amazon recommends products primarily based on high sales performance and, by extension, consumer tastes. A similar argument applies to the regressions in columns (4) and (5), where the dependent variable is whether a recommendation box is shown (i.e., whether a product "has a box"). If having a box diverts sales, thereby increasing the sales rank, reverse causality would bias the coefficient upward. The absence of a strong positive relationship therefore suggests that popular products are not systematically less likely to host the box in Canada. In short, Amazon appears to assign the box in a way that increases competitive pressure for comparatively "strong" products, as measured by sales, rather than "weak" ones.

Dependent Variables:	Recommended (CA)		Has Box (CA)	
Model:	(1)	(2)	(3)	(4)
Variables				
Amazon Basics	0.76***	0.76***	-0.42***	-0.46***
	(0.03)	(0.03)	(0.05)	(0.05)
Recommended (US)	-0.00	0.00	-0.01	-0.01
	(0.00)	(0.00)	(0.02)	(0.02)
Amazon Basics × Recommended (US)	0.10***	0.10***	0.04	0.04
	(0.03)	(0.03)	(0.03)	(0.03)
Has Box(US)	0.00*	0.00*	0.24***	0.24***
	(0.00)	(0.00)	(0.03)	(0.03)
Amazon Basic × Has Box (US)	-0.02	-0.02	-0.23***	-0.23***
	(0.04)	(0.04)	(0.06)	(0.06)
log(sales rank) (CA)		0.00		-0.01*
		(0.00)		(0.00)
Fixed-effects				
Category	Yes	Yes	Yes	Yes
Controls				
X_{CA}	No	Yes	No	Yes
X_{US}	No	Yes	No	Yes
Fit statistics				
Observations	17,086	17,086	17,086	17,086
R^2	0.828	0.828	0.370	0.376
Within R^2	0.811	0.811	0.133	0.142

Clustered (category) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 4: Regression of being recommended / having the box in CA on being recommended / having the box in US and product type (AB vs non AB)

Figure 2: Share of Amazon Basics Products by Sales Rank: Canada vs. US



Furthermore, if the preferential treatment of AB products in Canada were driven by a stronger local taste for Amazon's brand, this should be reflected in the sales distribution. Figure 2 plots the share of AB products among the top X percent of best-selling products in both domains. While the aggregate share of AB products is similar across countries (converging at the 100th percentile), AB products are under-represented among top-selling items in Canada compared to the US. In the top 1% of sales ranks, AB products account for 27% of items in the US but only 12% in Canada. Despite this, none of the remaining 88% of top-selling non-AB products in Canada receive a recommendation. In contrast, roughly 8% of comparable top-selling non-AB products in the US are recommended. The systematic exclusion of highly popular non-AB products from recommendations in Canada—while less popular AB products are promoted—contradicts the view that the algorithm merely reflects Canadian consumers' idiosyncratic tastes.

Importantly, we do not claim that the "Similar item to consider" box is the only form of self-preferencing, nor that self-preferencing is absent in the US. Rather, the practice appears significantly more pronounced in the Canadian domain with respect to the box specifically. When we mirror the specification in Table 4 to predict box assignment in the US (see Table A.2), we find that, while the AB status is significant, its predictive power is substantially smaller.[9] Furthermore, in the US, the coefficient on log sales rank is negative and significant for recommendations (indicating popular products are recommended) and positive for having the box (indicating less popular products host the box). This suggests that while Amazon US may engage in some self-preferencing, the Canadian algorithm deviates far more sharply from meritocratic, sales-based assignment.

4 The Sales Effects of Box Assignment

The analysis in the previous section implies that some non-AB products would have been recommended in Canada if it were not for the higher degree of self-preferencing. In turn, some AB products might not have been recommended if their recommendation was based only on consumer preferences. To assess the consequences of self-preferencing, we next describe our identification strategy and estimate the effects of being recommended and of having the box on product sales.

4.1 Identification

Our identification strategy exploits two empirical findings: (1) identical products appear in both the US and Canadian domains, and (2) the algorithm in the Canadian domain completely suppresses recommendations for non-AB products and boxes for AB products, as we have shown in Section 3. Although our primary interest lies in Canadian market sales, the former finding allows us to leverage US sales to control for shared consumer preferences and unobserved product quality. The latter finding allows us to eliminate concerns of reverse causality, provided that the box assignment in the US is reflecting US preferences only. When there are no recommendations or boxes, those cannot depend on idiosyncratic tastes of Canadian consumers. Intuitively, since US box assignments don't cater to Canadian consumers either, the difference in US box assignment must be responsible for the the difference in sales between products with a US recommendation/box and those without, not the other way around.

Before moving on to the analysis, we provide a simple formalization of the above argument using a potential outcomes approach. Denote $Y_i^m(d)$ as the potential outcome of product i in domain m under treatment status

[9] Note that Table A.2 does not contain interactions effects of AB and Recommended/Has Box (CA). This is because such an interaction would be multicollinear, only AB products are recommended in CA and only non-AB products have the box in CA. This should be taken into account when interpreting the coefficients.

$d \in \{0, 1\}$, where treatment is either being recommended, or having the box on the product page. Our outcome of interest is typically either the log sales rank or the log sales. Moreover, we denote the actual treatment status of product i in domain m as D_i^m, where $D_i^m = 1$ if the product is recommended/ has the box and $D_i^m = 0$ otherwise.

Assumption 1 *Suppose outcomes vary by domain and treatment status as follows:*

i) $Y_i^{CA}(d) = t_i^{CA} + Y_i^{US}(d)$

ii) $\mathbb{E}[t_i^{CA} | Y_i^{US}, D_i^{US} = 1] = \mathbb{E}[t_i^{CA} | Y_i^{US}, D_i^{US} = 0]$

where t_i^{CA} represents the idiosyncratic tastes of Canadians.

Assumption 1 i) implies that outcomes in Canada may be shifted relative to US outcomes under the same treatment, i.e., whether the box was assigned. That shifter is allowed to be product-specific, so the assumption represents a fairly general and intuitive relationship between outcomes in the US and Canada. Assumption 1 ii) is our core identifying assumption. It requires that, conditional on achieving the same observed sales performance in the US, treated and untreated products share the same expected Canadian market shifter. It rules out the specific concern that Amazon's US algorithm selects products for recommendation based on unobservable traits that specifically appeal to Canadian consumers more (or less) than they appeal to US consumers. Given that the US algorithm is optimizing for US demand, it is plausible that any idiosyncratic Canadian preference is orthogonal to the US selection mechanism conditional on observable success.

To estimate the treatment effect, we employ ANCOVA regression specifications of the following form:

$$Y_i^{CA} = \alpha + \gamma D_i^{US} + \delta Y_i^{US} + u_i \tag{1}$$

This specification estimates the coefficient γ as the difference in Canadian outcomes between products that are treated in the US and those that are not treated, conditional on the realized US outcome. Under the condition of Lemma 1 – respectively, that no non-AB products are recommended and no AB products have the box in the Canadian domain – γ therefore recovers the negative of the Average Treatment Effect on the Treated (ATT) in the US domain.

Lemma 1 *If no product is treated in the Canadian domain, i.e.* $D_i^{CA} = 0$ $\forall i$*, then:*

$$\begin{aligned} &\mathbb{E}[Y_i^{CA} | Y_i^{US} = y, D_i^{US} = 1] - \mathbb{E}[Y_i^{CA} | Y_i^{US} = y, D_i^{US} = 0] \\ &= -\mathbb{E}[Y_i^{US}(1) - Y_i^{US}(0) | Y_i^{US}(1) = y, D_i^{US} = 1] \end{aligned} \tag{2}$$

The proof of Lemma 1 can be found in Appendix A.1.

The intuition behind this result rests on the fact that the Canadian domain acts as a counterfactual environment where the treatment is universally absent ($D_i^{CA} = 0$). Consequently, observed outcomes in Canada reveal the untreated potential outcome, $Y_i^{CA}(0)$, for all products. By Assumption 1(i), this Canadian outcome serves as a proxy for the product's untreated potential outcome in the US, $Y_i^{US}(0)$, shifted only by idiosyncratic Canadian tastes t_i^{CA}. When we compare a product treated in the US ($D_i^{US} = 1$) to an untreated product ($D_i^{US} = 0$) that achieved the same observed outcome y in the US, we are effectively comparing a product that needed the treatment to reach level y against a product that reached level y organically. In the US, these products appear observationally equivalent. However, in Canada, the treatment boost is removed. The product that was treated in the US reverts back to its organic level, while the originally untreated product remains at its organic level. The difference in their Canadian outcomes therefore isolates the magnitude of the treatment effect that products recommended in the US lost, effectively recovering the negative of the treatment effect on the treated for that specific outcome level y. By averaging this conditional effect over the distribution of observed outcomes for the treated products, γ recovers the unconditional Average Treatment Effect on the Treated (ATT).

It is straightforward to include additional control variables in the above argument. Assumption 1 and Lemma 1 would then apply *conditional on observables*. Substantial changes in the estimated treatment effect after adding additional control suggests that Assumption 1 does not hold unconditionally. We therefore show estimation results with and without various control variables when analyzing the effect of recommendations, and having a box, respectively.

We favor the ANCOVA-style specification in Equation (1) over a Difference-in-Differences (DiD) approach to account for Regression to the Mean. If top performance in the US is partially driven by transient positive shocks ("luck"), treated products (which are likely selected based on their strong sales performance) would naturally revert to the mean in Canada. A DiD specification implicitly restricts $\delta = 1$, which forces a one-to-one mapping and would likely overestimate the treatment effect. Our specification leaves δ unrestricted, allowing the model to flexibly adjust for the correlation between the US and Canadian outcomes.[10]

If treatment rates in Canada were non-zero, identification would require

[10] It is straightforward to show that modifying the left-hand side of (2) to be a difference ($Y_i^{CA} - Y_i^{US}$) yields the same result on the right-hand side, but the regression implementation differs in how it handles δ.

stronger assumptions. Specifically, we would have to assume that the treatment effect is the same for all products.[11] By focusing on non-AB products (which never have the box in Canada) in Section 4.2 and AB products (which never show competitor boxes in Canada) in Section 4.3, we avoid these additional requirements.

4.2 The Effect of Recommendations on Non-AB Product Sales

As foreshadowed in the previous section, in this section, we restrict the sample to non-AB products and focus on the effect of recommendations. Recall that Y_i^m in Equation (1) denotes either the log sales rank or log sales of product i in domain m.[12] Tailoring the regression specification to the focus of this section, the treatment variable D_i^m is an indicator for whether product i is recommended in domain m; for clarity, we label this variable as Recommended (US) in the regression tables. We also include category fixed effects to account for baseline demand differences across categories.[13]

Table 5 presents the estimation results. Column (1) establishes the baseline: non-AB products that are recommended in the US have, on average and all else equal, sales ranks in Canada that are more than 50% higher (worse) than those of non-recommended products with similar US sales performance. This implies that the recommendation provided a substantial boost in the US, which, when lost in Canada, results in a sharp drop in sales relative to other products in the same category. When we control for other US characteristics – including whether the US product page features the box ("Has Box

[11] In this case, observing the same probability of treatment conditional on the US outcome for products that are treated in the US versus untreated, i.e. $\mathbb{E}[D_i^{CA}|Y_i^{US}(1), D_i^{US} = 1] = \mathbb{E}[D_i^{CA}|Y_i^{US}(0), D_i^{US} = 0]$, is sufficient for obtaining the causal effect. If $\mathbb{E}[D_i^{CA}|Y_i^{US}(1), D_i^{US} = 1] \neq \mathbb{E}[D_i^{CA}|Y_i^{US}(0), D_i^{US} = 0]$, we would have to correct the estimate by the difference in treatment rates but would still be able to recover the causal effect.

[12] Following Chevalier and Goolsbee (2003), we use the observed sales together with the sales rank to predict sales for the entire sample. This approach has been applied to books (Chevalier and Mayzlin, 2006; Reimers and Waldfogel, 2021) as well as a broad range of products (Gutierrez, 2021; He, Hollenbeck and Proserpio, 2022) sold on Amazon. See A.2 in the Appendix for details.

[13] For instance, products in the category "Beach Towels" may face systematically lower demand in Canada than in the US. This would violate Assumption 1 ii) since it implies a correlation between recommendation status in the US and idiosyncratic preferences of Canadians. The fixed effects, along with controls for price and review count, ensure we are comparing products with similar observable profiles within the same category. This approach strengthens our case that US box assignment is independent of Canadian outcomes in our application, conditional on observables.

Dependent Variables:	log(sales rank) (CA)		log(sales) (CA)	
Model:	(1)	(2)	(3)	(4)
Variables				
Recommended (US)	0.44***	0.34***	-0.15**	-0.12*
	(0.13)	(0.12)	(0.07)	(0.07)
log(sales rank) (US)	0.39***	0.25***		
	(0.02)	(0.02)		
log(sales) (US)			0.25***	0.14***
			(0.03)	(0.03)
Has Box (US)		-0.10**		0.05
		(0.04)		(0.03)
Amazon's Choice (US)		0.16***		-0.04
		(0.06)		(0.04)
1P (US)		-0.31***		0.16***
		(0.06)		(0.04)
log(price) (US)		0.03		-0.02
		(0.06)		(0.02)
log(review count) (US)		-0.30		0.14
		(0.23)		(0.13)
rating (US)		0.08		-0.006
		(0.22)		(0.13)
log(review count) × rating (US)		0.02		-0.0003
		(0.05)		(0.03)
Fixed-effects				
category	Yes	Yes	Yes	Yes
Fit statistics				
Observations	16,411	16,411	14,527	14,527
R^2	0.661	0.675	0.312	0.345
Within R^2	0.142	0.176	0.105	0.148

Clustered (category) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 5: Effect of a box recommendation on non-AB sales.

(US)”), the presence of the Amazon's Choice badge, first-party status, price, reviews, and ratings—the estimated sales loss settles at approximately 40% (Column 2). Recall that our estimate recovers the Average Treatment Effect on the Treated (ATT): it implies that the recommendations boost (lower) the sales rank for recommended products in the US by 29%.

The role of recommendations is corroborated when using log sales as the outcome variable. In our view, using *absolute* sales may be more meaningful, as it is easier to interpret. All else equal, products recommended in the US generate approximately 14% lower sales in Canada compared to their non-recommended counterparts (column (3)). This gap remains fairly stable as we add further controls. Based on the results in column (4), being recommended in the US is associated with a decrease in Canadian sales by 11% relative to products not recommended in the US. This result implies that being recommended causes an increase in sales by 13%.

Two additional results in Table 5 warrant discussion. First, the coefficient on US sales rank (and US sales) is substantially smaller than one in all specifications. This may indicate regression to the mean and supports our decision justification for using an ANCOVA specification instead of a standard difference-in-differences approach.

Second, the coefficient on “Has Box (US)” requires a cautious interpretation. Unlike recommendations, which are completely suppressed for non-AB products in Canada, the box features on some non-AB product pages in the Canadian domain. As discussed in Section 4.1, this violates the strict missing treatment condition required for causal identification established in Lemma 1. Moreover, having a box in the US is positively correlated with having a box in Canada. This further complicates identification. Recall, however, from the earlier discussion that the coefficient on “Has Box,” is likely attenuated when assuming a uniform treatment effect.[14]

4.3 The Effect of Having the Box on AB Product Sales

To estimate the effect of having the box on sales, we restrict the analysis to AB products. As established in Section 4.1, none of the AB product pages in Canada display a box recommending competing products. This implies that in the Canadian domain, AB products enjoy a “shielded” environment where

[14] If $\tau_i = \tau \forall i$, Equation (7) in Appendix A.1 implies that the estimated effect is scaled by $1 - (\mathbb{E}[D_i^{CA}|D_i^{US} = 1] - \mathbb{E}[D_i^{CA}|D_i^{US} = 0])$. Since the share of products having a box in Canada is larger among products having a box in the US than among products not having a box in the US, i.e., $\mathbb{E}[D_i^{CA}|D_i^{US} = 1] > \mathbb{E}[D_i^{CA}|D_i^{US} = 0]$, the estimated effect is biased towards zero.

consumers are not diverted to competitors – a protection from competition that is often missing in the US.

While the outcome variables Y_i^m remains as defined in the previous section, the treatment D_i^m is now an indicator for whether product i has the box in domain m. For transparency, we rename the indicator "Has Box (US)" in the regression tables. We again include category fixed effects in the regression specification from Equation (1).

This specification identifies the sales effect of being exposed to a competitor's recommendation. By controlling for Y_i^{US}, we compare AB products that achieve the same sales performance in the US, but under different competitive conditions. If the presence of a competitor box siphons off demand in the US, then γ should capture the "recovery" in sales rank that occurs when the box is removed in the Canadian market.

The logic of our identification argument parallels that for recommendations in the previous section, albeit with the opposite sign. Consider two AB products that achieve the same sales rank in the US. One is subject to a competitor box, while the other is not. To achieve sales comparable to those of the product without the box, the product with the box must have compensated for the effect of the box through comparatively stronger consumer appeal. In Canada, where the box is systematically removed for all AB products, these products face lower competitive pressure. We therefore expect the sales rank of the product that have a box in the US to do better in terms of their sales in Canada relative to its peers. A negative γ coefficient would therefore confirm that exposure to competitors through the box in the US hurts the product's sales performance there.

Table 6 presents the estimation results. Directionally, the point estimates are consistent with our above argument, suggesting that AB products exposed to competitor boxes in the US perform better in Canada where that exposure is removed. Column (2) indicates that, conditional on the US sales rank and adding similar controls as in the previous section, the sales rank in Canada is 17% lower (better) for products having a box in the US compared to products without a box in the US. Again, the result is corroborated when using log sales as the outcome: Products having a box in the US generate around 6% higher sales in Canada compared to their peers without a box. Put differently, the presence of competitor recommendations on a product page dampens demand by approximately 6% (Column 4). However, we note that unlike the results for recommendations, these estimates do not reach conventional levels of statistical significance, likely due to the relatively small sample of AB products. Nevertheless, we contend that the direction of the estimates suggest that, by removing these boxes in Canada, Amazon protects its own branded products from the same competitive pressures they

Dependent Variables:	log(sales rank) (CA)		log(sales) (CA)	
Model:	(1)	(2)	(3)	(4)
Variables				
Has Box (US)	-0.14	-0.19	0.04	0.06
	(0.16)	(0.16)	(0.14)	(0.13)
log(sales rank) (US)	0.65***	0.56***		
	(0.06)	(0.08)		
log(sales) (US)			0.44***	0.30***
			(0.06)	(0.08)
Recommended (US)		0.06		0.010
		(0.15)		(0.10)
Amazon's Choice (US)		0.01		0.06
		(0.17)		(0.12)
log(price) (US)		0.29**		-0.23***
		(0.12)		(0.07)
log(review count) (US)		2.0**		-0.76
		(0.82)		(0.79)
rating (US)		3.7**		-1.2
		(1.6)		(1.4)
log(review count) × rating (US)		-0.47**		0.18
		(0.18)		(0.17)
Fixed-effects				
category	Yes	Yes	Yes	Yes
Fit statistics				
Observations	675	675	511	511
R^2	0.844	0.848	0.637	0.656
Within R^2	0.531	0.543	0.351	0.386

Clustered (category) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 6: Effect of having a box on AB sales.

experience in the US market and thereby increases their sales.

5 Effect Heterogeneity

This section provides further evidence on the heterogeneity of the treatment effects documented above. To ensure sufficient statistical power for these granular splits, we focus on non-AB products in this section.

5.1 The Effect of the Number of Recommendations on Non-AB Sales

In our main specification, we treated recommendations as binary, a product was either recommended or not. However, the the intensity of the treatment varies across products: a product may appear in the boxes of multiple different products simultaneously. If the treatment effect operates through increased visibility, we would expect a monotonic relationship: products that lose a high volume of recommendations in Canada should suffer a sharper sales decline than those that lost only a few.

Table 7 tests this hypothesis by replacing the binary treatment indicator with a discrete variable counting the number of times a product was recommended in the US.[15] The results confirm that the intensity of exposure matters. The coefficient implies that for each additional recommendation in the US, the product's sales rank in Canada deteriorates (increases) by approximately 3%, and its sales volume drops by 1%. This suggests that the recommendation boost scales with the frequency of exposure, and that the aggregate effect documented in Section 4.2 is driven by the cumulative impact of multiple recommendation instances rather than a simple extensive margin effect.

5.2 Heterogeneity by baseline popularity

The value of a recommendation may also depend on the product's baseline popularity. Intuitively, recommendations might serve as a crucial discovery mechanism for niche or "long-tail" products, whereas established best-sellers may already enjoy sufficient organic visibility, rendering the recommendation redundant.

[15] We calculate the number of recommendations among the products scraped, i.e., among products that appear on the same search results page and are linked through boxes. This is likely a lower bound since the product might also be recommended on product pages that we did not scrape.

Dependent Variables:	log(sales rank) (CA)		log(sales) (CA)	
Model:	(1)	(2)	(3)	(4)
Variables				
Recommendation count (US)	0.04***	0.03**	-0.01**	-0.01*
	(0.010)	(0.01)	(0.006)	(0.006)
log(sales rank) (US)	0.38***	0.25***		
	(0.02)	(0.02)		
log(sales) (US)			0.25***	0.14***
			(0.03)	(0.03)
Has Box (US)		-0.12***		0.05*
		(0.04)		(0.03)
Amazon's Choice (US)		0.17***		-0.04
		(0.06)		(0.04)
1P (US)		-0.31***		0.17***
		(0.06)		(0.04)
log(price) (US)		0.03		-0.02
		(0.05)		(0.02)
log(review count) (US)		-0.29		0.14
		(0.23)		(0.13)
rating (US)		0.10		-0.01
		(0.22)		(0.13)
log(review count) $\times$ rating (US)		0.01		0.0006
		(0.05)		(0.03)
Fixed-effects				
category	Yes	Yes	Yes	Yes
Fit statistics				
Observations	16,411	16,411	14,527	14,527
R^2	0.661	0.675	0.311	0.345
Within R^2	0.140	0.175	0.104	0.148

Clustered (category) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 7: Effect of number of recommendations on non-AB sales.

Dependent Variables:	log(sales rank) (CA)		log(sales) (CA)	
Model:	(1)	(2)	(3)	(4)
Variables				
log(sales rank) (US)	0.38***	0.25***		
	(0.02)	(0.02)		
Recommended ×	0.71***	0.65***		
(sales rank > median) (US)	(0.21)	(0.20)		
Recommended ×	0.33**	0.21		
(sales rank < median) (US)	(0.13)	(0.14)		
log(sales) (US)			0.25***	0.14***
			(0.03)	(0.03)
Recommended ×			-0.35***	-0.34***
(sales < median) (US)			(0.11)	(0.11)
Recommended ×			-0.08	-0.04
(sales > median) (US)			(0.09)	(0.08)
Has Box (US)		-0.11**		0.05*
		(0.04)		(0.03)
Amazon's Choice (US)		0.16***		-0.04
		(0.06)		(0.04)
1P (US)		-0.30***		0.16***
		(0.06)		(0.04)
log(price) (US)		0.03		-0.02
		(0.06)		(0.02)
log(review count) (US)		-0.31		0.15
		(0.23)		(0.13)
rating (US)		0.07		-0.001
		(0.22)		(0.13)
log(review count) × rating (US)		0.02		-0.0008
		(0.05)		(0.03)
Fixed-effects				
category	Yes	Yes	Yes	Yes
Fit statistics				
Observations	16,411	16,411	14,527	14,527
R^2	0.662	0.675	0.312	0.345
Within R^2	0.142	0.176	0.105	0.149

Clustered (category) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 8: Heterogeneity of the effect of box recommendations on non-AB sales by baseline popularity in the US.

To investigate this, we split the sample along the median US sales rank. Table 8 reveals a striking asymmetry: the treatment effect is concentrated entirely among less popular products. We find that non-AB products with below-median sales (worse rank) suffer disproportionately when they lose their recommendation status in Canada. Specifically, we find that, among non-AB products in the upper half of the US sales rank distribution, US-recommended products experience an increase in their sales rank by roughly 90% compared to products that are not recommended in the US, conditional on controls (Column (2)). According to our estimates, this translates into about 29% lower sales. In contrast, for non-AB products in the more popular half, the lack of the recommendation leads to a statistically negligible change in sales (rank). These results suggests to us that the "Similar Items to Consider" box may serve as an informational tool that reduces search frictions for less well-known, and thus lower-selling, products.

5.3 Removing the box for non-AB products

While our main results show that non-AB products generally lose when Amazon engages in self-preferencing, this aggregate effect may conceal the presence of non-AB "winners" from self-preferential box assignment in Canada. Specifically, non-AB products that hosted a competitor box on their product page in the US may benefit from the absence of the box in the Canadian domain.

Consider a non-AB product A that is never recommended itself, but whose product page in the US features a box recommending a competitor B. In Canada, Amazon's self-preferencing algorithm removes the recommendation for competitor product B. This effectively shields product A from competitive pressure in the Canadian domain. If the box on product A's product page diverts sales to product B, we should therefore expect product A to perform relatively better in Canada than a product with comparable US performance whose product page was not subject to a box in the US.

To test this hypothesis, we conduct an analogous analysis on non-AB products to Table 6, which we presented in Section 4.3 for the AB product sample. The treatment group consists of non-AB products whose page displayed a competitor box in the US, while the control group consists of those that did not. Since the treatment group experiences a reduction in competitive pressure in Canada, we expect a recovery in their sales rank.

Table 9 confirms this prediction. Non-AB products that hosted a competitor box in the US experience a 29% improvement (decrease) in their sales rank and a 22% increase in sales volume in Canada relative to the control group.

Dependent Variables:	log(sales rank) (CA)		log(sales) (CA)	
Model:	(1)	(2)	(3)	(4)
Variables				
Has Box (US)	-0.37***	-0.34***	0.20***	0.20***
	(0.08)	(0.07)	(0.06)	(0.05)
log(sales rank) (US)	0.38***	0.25***		
	(0.02)	(0.03)		
log(sales) (US)			0.24***	0.12***
			(0.04)	(0.04)
Recommended (US)		0.39**		-0.10
		(0.17)		(0.07)
Amazon's Choice (US)		0.11*		-0.004
		(0.06)		(0.04)
1P (US)		-0.34***		0.17***
		(0.06)		(0.04)
log(price) (US)		0.001		-0.01
		(0.06)		(0.02)
log(review count) (US)		-0.29		0.22
		(0.30)		(0.16)
rating (US)		0.02		0.15
		(0.30)		(0.16)
log(review count) × rating (US)		0.01		-0.02
		(0.07)		(0.04)
Fixed-effects				
category	Yes	Yes	Yes	Yes
Fit statistics				
Observations	10,429	10,429	9,167	9,167
R^2	0.659	0.674	0.300	0.339
Within R^2	0.144	0.181	0.103	0.153

Clustered (category) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 9: Effect of having a non-AB box on non-AB products.

We interpret these magnitudes as a lower bound. As discussed in Section 4.1, although the algorithm suppresses non-AB product recommendations in Canada, it sometimes replaces them with recommendations for AB products. Indeed, in our sample, the probability that a box containing a non-AB product is displayed is 0% in Canada. By contrast, the probability that a box containing an AB product is displayed is 24% for the treatment group (i.e., non-AB products whose page features a box in the US), compared with 15% for the control group (i.e., those without a box in the US). Because some treated products therefore remain subject to recommendations of a competitor product, albeit now a AB product, the coefficient on Has Box (US) is likely downward biased and smaller in magnitude than the true sales effect of having the box. Nevertheless, the fact that we observe a significant positive effect substantiates the role of the box assignment in diverting consumers and sales: removing competitor recommendations substantially boosts the host product's sales, regardless of whether the host is Amazon or not.

6 Discussion

In this section, we explore the reasons behind the observed cross-border differences in self-preferencing and attempt to quantify the aggregate economic implications of our findings.

Naturally, we do not know the exact rationale for Amazon's decision to engage in self-preferencing in this way in Canada, but not in the US. Nevertheless, two considerations are likely to contribute to the more aggressive self-preferencing observed in Canada relative to the US. First, the platform may be using self-preferencing as a compensatory mechanism for the weaker organic performance of its private labels in the Canadian market. As shown in Figure 2, AB products are significantly under-represented in the top percentiles of the Canadian sales rank distribution compared to the US. Confronted with idiosyncratic local preferences for established brands, the platform may have incentives to distort the box-assignment algorithm to favor its own products.

Second, the disparity may reflect regulatory arbitrage. In recent years, antitrust scrutiny of digital markets has been considerably more intense in the US. Following reports that Amazon used proprietary data to launch competing products, the platform reportedly scaled back private label promotions to mitigate political pressure.[16] In contrast, Canadian competition enforce-

[16] In 2019, NBC News reported that Amazon removed many promotions of its own products after US policymakers advocated for breaking up large technology

ment has historically operated under stricter evidentiary standards for abuse of dominance, often requiring explicit demonstration of anti-competitive effects, and has capped financial penalties.[17] This differential in the legal risk associated with self-preferential conduct may partially explain differences in practices across jurisdictions.

To gauge the economic significance of these practices, we perform a back-of-the-envelope calculation estimating the counterfactual impact if the US market were to adopt the Canadian regime (i.e., if the "Similar item to consider" box were reserved exclusively for AB products). In our data, non-AB products that currently benefit from these recommendations account for approximately 20% of the total non-AB sales volume in the US. Applying our treatment effect estimate from Table 5, removing these recommendations would depress the sales of the affected products by roughly 11%.

This implies that a Canadian-style self-preferencing policy would reduce total non-AB sales in the US by approximately 2.2%. To quantify this in unit terms, we estimate the total volume of non-AB products sold on Amazon US. In 2023, independent sellers sold more than 4.5 billion items on Amazon.com.[18] As third-party sellers accounted for approximately 61% of total units sold on the platform in 2023,[19] this suggests a total US sales volume of roughly 7.4 billion units. Since non-AB products (including both third-party and first-party vendor sales) constitute approximately 99% of this volume, the relevant base is approximately 7.3 billion units.[20] Applying the 2.2% decline to this base implies a loss of more than 160 million unit sales annually.[21]

companies. See https://www.nbcnews.com/tech/tech-news/amazon-quietly-removes-promotions-its-own-products-calls-tech-regulation-n990666.

[17]It is important to note that both the FTC (see https://www.ftc.gov/legal-library/browse/cases-proceedings/1910129-1910130-amazoncom-inc-amazon-ecommerce) and the Competition Bureau Canada (see https://www.canada.ca/en/competition-bureau/news/2025/07/competition-bureau-advances-investigation-of-amazons-marketplace-fair-pricing-policy.html) are currently litigating against Amazon for alleged anti-competitive conduct. However, Canadian law has historically limited private access to the Competition Tribunal and required high standards of proof for anti-competitive acts (see https://laws.justice.gc.ca/eng/acts/C-34/page-14.html#h-89563).

[18]See Amazon's *2023 U.S. Small Business Empowerment Report*. Available at https://assets.aboutamazon.com/38/7a/3642a4d4459ba4b093132041a228/amazonsmallbusinessempowermentrepo

[19]See Amazon's Q4 2023 Earnings Release. The share of paid units sold by third-party sellers averaged 61% in 2023.

[20]See https://www.momentumcommerce.com/amazons-private-label-market-share-shrinks-by-6-year-over-year-in-q1-2024/.

[21]We abstract from the corresponding sales gain for AB products. Since AB products make up at most 1% of total sales, including them would have a negligible impact on the aggregate figures. Moreover, our identification strategy does not allow for a causal

It is important to note that the counterfactual presented here is limited to self-preferencing with respect to the “Similar Items to Consider” widget occurring in the US domain to the same extent as in the Canadian domain. Our estimate effectively assumes that the competitive playing field remains level across all other dimensions of the website, making it a strictly conservative lower bound.

7 Conclusion

This paper provides empirical evidence of platform self-preferencing by exploiting a stark institutional discontinuity: the systematic suppression of recommendations for products that are not sold under the Amazon Basics brand (i.e., non-AB products) in the Canadian marketplace. While Amazon’s US algorithm frequently recommends non-AB products in the “Similar item to consider” box, we document that the Canadian algorithm restricts this visibility exclusively to Amazon’s own private label.

Leveraging this cross-domain variation, we identify the causal impact of algorithmic exclusion on market outcomes. We find that non-AB products that merit a recommendation based on their US performance but are denied one in Canada suffer a 11% reduction in sales volume. This effect is heterogeneous, disproportionately harming less popular “long-tail” products that rely on algorithmic discovery to reach consumers. Conversely, we find suggestive evidence that Amazon shields its own brands from competitive pressure in Canada, further entrenching their market position.

The economic magnitude of this conduct is substantial: if the US market were subject to the same degree of exclusionary self-preferencing observed in Canada, our estimates imply a loss of over 160 million unit sales annually for non-AB sellers. These results underscore the material distortion caused by gatekeeper dual-role platforms and highlight the critical importance of monitoring algorithmic behavior across jurisdictions.

Finally, we emphasize that these estimates likely represent a lower bound on the total welfare impact of self-preferencing. Our analysis isolates the effect of a single website feature—the “Similar item to consider” widget—which constitutes only one touchpoint in the consumer journey. To the extent that Amazon employs similar preferential logic in search rankings, the “Buy Box” assignment, or other promotional widgets, the aggregate distortion of competition is likely significantly larger than what we document here. These results underscore the material distortion caused by gatekeeper dual-role platforms.

estimate of the effect of recommendations on AB products, so we exclude this channel.

References

Chen, Nan and Hsin-Tien Tsai, "Steering via algorithmic recommendations," *The RAND Journal of Economics*, 2024.

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

Chevalier, Judith and Austan Goolsbee, "Measuring prices and price competition online: Amazon. com and BarnesandNoble. com," *Quantitative marketing and Economics*, 2003, *1*, 203–222.

Corniere, Alexandre De and Greg Taylor, "A model of biased intermediation," *The RAND Journal of Economics*, 2019, *50* (4), 854–882.

Dendorfer, Florian, "First-party selling and self-preferencing," *International Journal of Industrial Organization*, 2024, *97*, 103098.

Etro, Federico, "e-Commerce platforms and self-preferencing," *Journal of Economic Surveys*, 2024, *38* (4), 1516–1543.

Farronato, Chiara, Andrey Fradkin, and Alexander MacKay, "Self-preferencing at Amazon: evidence from search rankings," in "AEA Papers and Proceedings," Vol. 113 American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203 2023, pp. 239–243.

_ , _ , and _ , "Vertical Integration and Consumer Choice: Evidence from a Field Experiment," Technical Report, National Bureau of Economic Research 2025.

Gutierrez, German, "The welfare consequences of regulating Amazon," *Job Market Paper, New York University*, 2021.

Hagiu, Andrei, Tat-How Teh, and Julian Wright, "Should platforms be allowed to sell on their own marketplaces?," *The RAND Journal of Economics*, 2022, *53* (2), 297–327.

He, Sherry, Brett Hollenbeck, and Davide Proserpio, "The market for fake reviews," *Marketing Science*, 2022, *41* (5), 896–921.

Lee, Kwok Hao and Leon Musolff, "Entry into two-sided markets shaped by platform-guided search," *Job Market Paper, Princeton University*, 2021.

Padilla, Jorge, Joe Perkins, and Salvatore Piccolo, "Self-Preferencing in Markets with Vertically Integrated Gatekeeper Platforms," *The Journal of Industrial Economics*, 2022, *70* (2), 371–395.

Peitz, Martin, "How to apply the self-preferencing prohibition in the DMA," *Journal of European Competition Law & Practice*, 2023, *14* (5), 310–315.

Raval, Devesh, "Steering in one click: Platform self-preferencing in the amazon buy box," *Unpublished manuscript*, 2022.

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), 1944–1971.

Tirole, Jean and Michele Bisceglia, "Fair gatekeeping in digital ecosystems," *TSE Working Paper*, 2023.

Waldfogel, Joel, "Amazon Self-preferencing in the Shadow of the Digital Markets Act," Technical Report, National Bureau of Economic Research 2024.

Zennyo, Yusuke, "Platform encroachment and own-content bias," *The Journal of Industrial Economics*, 2022, *70* (3), 684–710.

Zou, Tianxin and Bo Zhou, "Self-preferencing and search neutrality in online retail platforms," *Management Science*, 2024.

Appendix

A.1 Proof of Lemma 1

We begin by expanding the observed difference in conditional expectations. Let Δ denote the LHS of Equation (2). Using the definition of observed outcomes $Y_i^{CA} = D_i^{CA} Y_i^{CA}(1) + (1 - D_i^{CA}) Y_i^{CA}(0)$, we write:

$$\begin{aligned}\Delta = \mathbb{E}[D_i^{CA} Y_i^{CA}(1) + (1 - D_i^{CA}) Y_i^{CA}(0) \mid Y_i^{US}(1) = y, D_i^{US} = 1] \\ - \mathbb{E}[D_i^{CA} Y_i^{CA}(1) + (1 - D_i^{CA}) Y_i^{CA}(0) \mid Y_i^{US}(0) = y, D_i^{US} = 0]\end{aligned} \tag{3}$$

Based on Assumption 1(i), we can denote the treatment effect as $\tau_i = Y_i^{CA}(1) - Y_i^{CA}(0) = Y_i^{US}(1) - Y_i^{US}(0)$. Using this to rewrite the above equation, we get:

$$\begin{aligned}\Delta = \mathbb{E}[D_i^{CA} \tau_i + Y_i^{CA}(0) \mid Y_i^{US}(1) = y, D_i^{US} = 1] \\ - \mathbb{E}[D_i^{CA} \tau_i + Y_i^{CA}(0) \mid Y_i^{US}(0) = y, D_i^{US} = 0]\end{aligned} \tag{4}$$

Moreover, using Assumption 1(i) to replace $Y_i^{CA}(0)$, we get:

$$\begin{aligned}\Delta = \mathbb{E}[D_i^{CA} \tau_i + Y_i^{US}(0) + t_i^{CA} \mid Y_i^{US}(1) = y, D_i^{US} = 1] \\ - \mathbb{E}[D_i^{CA} \tau_i + Y_i^{US}(0) + t_i^{CA} \mid Y_i^{US}(0) = y, D_i^{US} = 0]\end{aligned} \tag{5}$$

By Assumption 1(ii), the Canadian taste shifter t_i^{CA} cancels out across the two terms. We are left with:

$$\begin{aligned}\Delta = \mathbb{E}[D_i^{CA} \tau_i + Y_i^{US}(0) \mid Y_i^{US}(1) = y, D_i^{US} = 1] \\ - \mathbb{E}[D_i^{CA} \tau_i + Y_i^{US}(0) \mid Y_i^{US}(0) = y, D_i^{US} = 0]\end{aligned} \tag{6}$$

We can now use the fact that we condition on the US outcome being equal to y. For the first term, we substitute $Y_i^{US}(0) = Y_i^{US}(1) - \tau_i$, which implies $\mathbb{E}[Y_i^{US}(0) | Y_i^{US}(1) = y, D_i^{US} = 1] = y - \mathbb{E}[\tau_i | Y_i^{US}(1) = y, D_i^{US} = 1]$. Substituting into Equation (6) yields:

$$\begin{aligned}\Delta = \mathbb{E}[D_i^{CA} \tau_i \mid Y_i^{US}(1) = y, D_i^{US} = 1] - \mathbb{E}[D_i^{CA} \tau_i \mid Y_i^{US}(0) = y, D_i^{US} = 0] \\ - \mathbb{E}[\tau_i \mid Y_i^{US}(1) = y, D_i^{US} = 1]\end{aligned} \tag{7}$$

In our setting, no products are treated in Canada ($D_i^{CA} = 0$ for all i). Thus, the first two expectations vanish. The expression simplifies to:

$$\Delta = -\mathbb{E}[\tau_i \mid Y_i^{US}(1) = y, D_i^{US} = 1] \tag{8}$$

This recovers the negative of the Average Treatment Effect on the Treated (ATT) for the US domain.

Note that Equation (7) represents the general case. If treatment rates in Canada were non-zero, identification would require additional assumptions. For instance, assuming $\tau_i = \tau \forall i$ would allow to recover the causal estimate.

A.2 Estimating Sales

We assume that sales ranks follow a Pareto distribution so that the natural logarithm of sales for a product is a linear function of the natural logarithm of the product's sales rank minus one.

$$\ln(\text{sales}_{cim}) = a_{cm} + b_{cm} \ln(\mathbf{x}_{cim}) + u_{cim} \tag{9}$$

We estimate Equation (9) with $\mathbf{x}_{cim} = (\text{rank}_{cim} - 1)$ for each category and domain, provided there are at least three distinct sales observations and excluding products with a sales rank of one. After excluding category-market pairs with a poor model fit (i.e., $R^2 \leq 0.1$), the mean R^2 across 354 regressions is 0.42. To improve our prediction, we include the number of reviews as a covariate and re-estimate Equation (9) using $\mathbf{x}_{cim} = (\text{rank}_{cim} - 1, \text{review count}_{im})$. Since reviews accumulate from past sales, they are predictive of sales. As a result, the mean R^2 (across 466 category-domain pairs) improves to 0.52. We use these predicted sales values for our analysis. Figure A.1 provides a graphical illustration of how the estimates fit the data, using a specific category-domain example.

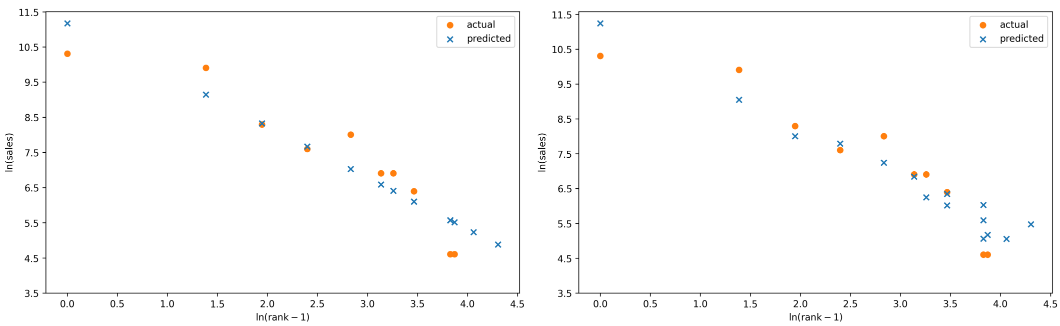


Figure A.1: Model fit for category "Rotating Power Toothbrushes" on Amazon.com using $\ln(\text{rank} - 1)$ (left) and $\ln(\text{rank} - 1)$ as well as $\ln(\text{review count})$ (right) as predictors

A.3 Additional Tables and Figures

Category	**Product Count (All)**	**Product Count (AB)**
Automotive	619	46
Baking & Cookie Sheets	264	34
Bedding Comforter Sets	341	66
Cell Phone Screen Protectors	477	14
Clothing	375	42
Electronics	929	94
Extension Cords	320	44
Floor Lamps	306	4
LED Bulbs	253	54
Mixing Bowls	268	16
Mouse Pads	276	24
Office Products	2236	460
Patio	918	56
Pet Supplies	545	108
Self-Stick Note Pads	241	60
Sports & Outdoors	1306	158
Throw Pillow Covers	278	18
Tools & Home Improvement	2530	222
USB Flash Drives	248	28
Video Games	256	4

Table A.1: (AB) product counts for the top 20 largest categories

Dependent Variables:	Recommended (US)		Has Box (US)	
Model:	(1)	(2)	(3)	(4)
Variables				
Amazon Basics	0.58***	0.56***	-0.22***	-0.19***
	(0.04)	(0.04)	(0.06)	(0.06)
Recommended(CA)	0.13***	0.13***	0.01	0.010
	(0.04)	(0.04)	(0.05)	(0.05)
Has Box (CA)	-0.02***	-0.02***	0.30***	0.30***
	(0.006)	(0.005)	(0.03)	(0.03)
log(sales rank) (US)		-0.02***		0.007*
		(0.003)		(0.004)
Fixed-effects				
category	Yes	Yes	Yes	Yes
Control				
X_{US}	No	Yes	No	Yes
X_{CA}	No	Yes	No	Yes
Fit statistics				
Observations	17,086	17,086	17,086	17,086
R^2	0.298	0.323	0.291	0.298
Within R^2	0.241	0.268	0.093	0.101

Clustered (category) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A.2: Regression of being recommended / having the box in US on being recommended / having the box in CA and product type (AB vs non AB)