Private Labels and Product Imitation: Evidence from Amazon*

Malika Korganbekova — *The University of Chicago*

Alibek Korganbekov — *Indiana University*

July 22, 2026

CLICK HERE FOR THE MOST RECENT VERSION

Abstract

Platforms that operate a marketplace while selling on it can use marketplace data to compete with the sellers they host. Scraping over 770 million product pages from Amazon's four largest marketplaces, we find that 73% of Amazon's private-label products are near-visually identical to a specific pre-existing third-party product, though imitation is widespread among third-party sellers as well. Consumers who buy private labels save 42% at the median. However, exploiting cross-country variation in the timing of private-label entry within categories, we find no category-wide price benefit: entry compresses the price distribution from both tails, increases seller exit and out-of-stock rates, and worsens incumbent sales ranks, while product quality is unchanged. These effects grow with private-label penetration, and in the most penetrated categories the margin of adjustment shifts from seller exit to deterred entry – the pattern predicted by "kill zone" theories of platform competition.

*We thank conference participants at the Kellogg School of Management, Yale University AI Center, Marketing Science 2025, Digital Competition and Tech Regulation Conference HBS Conference for their useful feedback and comments. Corresponding authors: malika.korganbekova@chicagobooth.edu and akorganb@iu.edu.

1 Introduction

A growing number of digital platforms operate in a dual role. They run a marketplace that connects consumers and third-party sellers, and they sell their own products on that same marketplace. Amazon, Apple, and Walmart Marketplace all introduce private-label products that compete directly with the sellers on their platforms. This dual role of platforms has attracted intense regulatory scrutiny, most notably by the U.S. House Judiciary Committee investigation (U.S. House Committee on the Judiciary, 2020) and the European Commission's probe into Amazon.[1] The central concern in both cases is that platforms exploit individual-level seller data to introduce competing products, a practice known as "sherlocking." Amazon denies this. Its 2014 Seller Data Protection Policy prohibits the use of non-public, seller-specific data, and its lawyers testified to Congress that only aggregate data informs private-label decisions.

The regulatory concern rests on an economic premise: that private-label entry harms competition, third-party sellers and, ultimately, consumers. Whether this premise holds is not obvious. Existing research finds that private labels can benefit consumers because they offer comparable products at lower prices (Chintagunta et al., 2002; Dubé, 2024; Farronato et al., 2022). The long-run effect, however, depends on what happens to sellers. On the one hand, platform entry may create a "kill zone" and divert demand from third-party sellers to the platform's own products (Kamepalli et al., 2020). Sellers that lose revenue may leave the platform, decreasing competition, which would harm consumers over time. On the other hand, entry may validate the category and attract new consumers, expanding the market for all sellers (Rochet and Tirole, 2003). Which force dominates is an empirical question, and systematic evidence remains scarce.

This paper addresses two questions. First, is there systematic evidence that Amazon uses *individual* third-party sellers' data to introduce its private-label products? Second, what is the impact of Amazon's private-label entry on consumers and on affected third-party

[1] See https://audiovisual.ec.europa.eu/en/media/video/I-198672.

sellers? While we do not estimate total welfare, our reduced-form estimates speak to the key components of the tradeoff - prices, product variety, and seller survival - that are essential for informing the ongoing policy debate.

To answer these questions, we assemble a novel dataset of unprecedented scale. We scrape over 770 million product pages from Amazon's four largest marketplaces - the U.S., Germany, the U.K., and Canada - covering both private-label and third-party seller products from 2018 to 2026. For each product, we observe prices, ratings, review counts and the distribution of review ratings (e.g., the number of 1-star, 2-star ratings, etc.), product images, textual descriptions, sales ranks, introduction dates, and product taxonomy information. This panel structure allows us to track product entry, measure similarity between products, and estimate the causal effects of private-label entry on incumbent sellers over time. We identify the effects of Amazon entry using two sources of variation. First, we use similarity-based matching to understand the impact of private label entry at the product-level. Second, we use a cross-country staggered difference-in-differences design that exploits variation in the timing of private-label entry across countries within the same product category to get category-level effects.

To answer the first question - what data informs Amazon's private-label entry - we proceed as follows. We measure the visual and textual similarity between each entering Amazon private-label product and the incumbent third-party products in its category. To get visual similarity, we train a Siamese neural network that maps product images into a 512-dimensional embedding space and compute the cosine similarity between the resulting vectors. For textual similarity, we construct sentence embeddings from product titles and descriptions and compare them in the same way. High similarity to a specific incumbent product - rather than to the category as a whole - is difficult to reconcile with the use of aggregate data alone.

Calculating pairwise cosine similarities across millions of products is computationally prohibitive. For the similarity and product-level analysis, we therefore focus on 14 product

categories, which contain 1,673 private-label products and 5.5 million third-party products. We document substantial visual similarity between Amazon's private labels and pre-existing third-party products. Of the 1,673 private-label products, 1,222 (73%) have at least one third-party predecessor with a visual cosine similarity of 0.9 or above - a threshold corresponding to near-identical appearance. Some pairs exceed 0.98, effectively duplicates of specific incumbent products. This pattern is difficult to reconcile with the use of aggregate data alone and suggests that Amazon relied on individual-level product data in its entry decisions. However, imitation is not unique to the platform. We also find substantial, though less pronounced, similarity among third-party products themselves: sellers copy one another, and some products are practically identical. Copying is thus a widespread practice on the marketplace rather than an Amazon-specific behavior.

Evidence that Amazon targets specific products when producing its private label products does not by itself establish harm. We therefore turn to our second question: the impact of Amazon's private-label entry on consumers and third-party sellers.

Amazon's private-label products are, on average, $24 (35%) cheaper than their most similar third-party counterparts. This means that consumers get practically same product at a lower price. In certain categories, we document that the price advantage is not driven by lower quality inputs: using our textual similarity measures, we can hold product materials fixed - for example, comparing cast-iron dutch ovens to cast-iron dutch ovens - and private labels remain substantially cheaper. We also document that consumers give higher ratings and reviews to the private label products rather than their most-similar third-party counterparts, which suggests that consumers are satisfied with the products.

On the supply side, we estimate the causal impact of private-label entry on third-party sellers at both the product and category levels. At the product level, third-party products directly imitated by Amazon experience a statistically significant deterioration in sales rank (4.5%), downward price adjustments (1.1%), an elevated probability of discontinuation. Amazon's entry thus displaces the very products it imitates.

At the category level, using the Callaway and Sant'Anna (2021) difference-in-differences estimator with not-yet-treated controls and exploiting within-category cross-country variation in the timing of private-label entry, we find that treated category–country cells experience a significant 31% increase in seller exits, a 4.4 percentage point increase in out-of-stock rates, and a 14% deterioration in third-party sellers' sales ranks, while product quality as measured by star ratings is largely unaffected. Thus, categories that receive Amazon's private-label products experience negative competitive outcomes. These effects scale with treatment intensity: as the share of private-label products in a category increases, the sales ranks of incumbent sellers deteriorate more sharply. Moreover, in the categories with a larger share of the private label products, we find a significant decrease in seller entry compared to not-yet-treated same category in other countries. Since most of the products that are sold and get imitated by Amazon belong to smaller sellers, they are significantly impacted by the private label entry.

Overall, our results indicate that although consumers may currently be happier with Amazon products, there are negative long-term consequences on the supply-side of the market, which can harm consumers in the long-term.

This paper contributes to several strands of the literature. A large theoretical literature examines platform self-preferencing and private-label introduction (Farrell and Katz, 2000; Hagiu and Spulber, 2013; Hagiu et al., 2015, 2022; Jiang et al., 2011; Anderson and Bedre Defolie, 2024; Hagiu and Wright, 2015; Madsen and Vellodi, 2022; Zhou and Zou, 2023). Our paper provides large-scale empirical evidence on several mechanisms these models highlight, including the role of data in guiding entry decisions and the heterogeneous effects on incumbent sellers.

On the empirical side, our work is most closely related to studies of Amazon's private-label strategy (Chen and Tsai, 2024; Farronato et al., 2022; Gutierrez, 2021; Lam, 2021; Lee and Musolff, 2021; Raval, 2022; Crawford et al., 2022). We make two primary contributions. First, we provide the first systematic, large-scale analysis of product imitation using image

recognition methods applied to millions of products to document the prevalence and characteristics of imitation across a broad set of categories. Second, we provide direct evidence on competitive effects at both the category and product levels, including seller exit and entry, price compression, product quality, and individual product discontinuation.

The remainder of the paper is organized as follows. Section 2 describes the empirical setting and the data. Section 3 presents the evidence on product imitation. Section 4 examines consumer-side outcomes. Section 5 estimates product-level effects on targeted third-party sellers. Section 6 estimates category-level effects of private-label entry using a cross-country staggered design. Section 7 concludes.

2 Empirical Setting

Private-label products - goods manufactured by or on behalf of a retailer and sold under the retailer's own brand - have been a feature of retail since the 1950s. Grocery chains pioneered the practice, and general merchandise retailers followed: Target's "Up & Up," Costco's "Kirkland Signature," and Walmart's "Great Value" all compete directly with branded products on their shelves. By eliminating brand marketing costs and leveraging shelf-space control, retailers offer comparable products at lower prices while earning higher margins (Dubé, 2024).

Digital marketplaces introduce a fundamentally different structure, and the distinction matters for understanding why platform private labels raise concerns that traditional store brands do not. Three features set the digital platform setting apart.

First, the *informational advantage* is qualitatively different. A traditional retailer observes wholesale costs and point-of-sale data for the products on its shelves. Amazon observes granular data on every dimension of seller and consumer behavior across its entire marketplace: search queries, click-through rates, conversion rates, return rates, pricing histories, and inventory levels. The platform can identify precisely which products are

commercially successful, at what margin, and why - down to metrics that no traditional retailer has access to and that third-party sellers themselves cannot observe about their competitors. This asymmetry creates the potential for what regulators call *sherlocking*: the platform uses proprietary marketplace data to identify successful products and introduce competing private labels.

Second, the platform *controls the competitive infrastructure.* A traditional retailer controls shelf space; Amazon controls the algorithm - search rankings, the Buy Box, "Amazon's Choice" badges, and product recommendations. These algorithmic levers can impact which products consumers pay attention to. They can be used to benefit the platform's own offerings in ways that are opaque and difficult to verify from the outside. Crucially, this means that a platform private label does not need to be a superior product to displace incumbents - algorithmic placement alone can shift demand. This distinguishes the platform setting from traditional retail, where shelf-space allocation is observable and often contractually negotiated.

Finally, third-party sellers *lack comparable outside options.* If Procter & Gamble faces competition from Walmart's "Great Value," it can shift volume to Target, Costco, Kroger, or any number of alternative retailers. If a third-party seller on Amazon faces competition from Amazon Basics, the outside option is dramatically worse: Amazon accounts for roughly 40–50% of U.S. e-commerce, and no alternative marketplace offers comparable customer reach. This concentration amplifies the competitive harm of private-label entry and creates a hold-up problem: sellers invest in product development, listing optimization, and review accumulation on Amazon's platform, and the platform can observe which investments succeed and then enter with its own product. The *threat* of expropriation may distort sellers' ex ante incentives to invest and innovate, even before entry actually occurs.

These three features - informational advantage, algorithmic control, and the absence of outside options - interact to make platform private-label entry economically distinct from traditional store brands. Reflecting these concerns, the U.S. House Committee on

the Judiciary's 2020 investigation specifically targeted large platforms such as Amazon and Apple, raising the concern that Amazon may use proprietary marketplace data to inform the design of its private-label offerings (U.S. House Committee on the Judiciary, 2020). Additional concerns include Amazon's ability to influence product visibility through the Buy Box algorithm and search rankings (Lee and Musolff, 2021; Farronato et al., 2022). These concerns are not unique to Amazon: Apple's App Store has faced similar scrutiny for replicating popular third-party apps. The European Union's Digital Markets Act restricts "gatekeepers" from using non-publicly available business-user data to compete against those users. Our first set of results directly addresses whether Amazon uses individual-level data.

2.1 Data

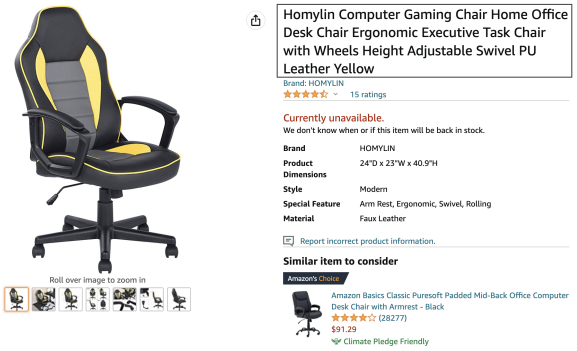
We construct a novel panel dataset by scraping product pages from Amazon's four largest marketplaces: the U.S., Germany, the U.K., and Canada. Our data collection covers 2018 to 2026 and encompasses over 770 million product pages: approximately 350 million from the U.S., 226 million from Canada, 105 million from Germany, and 91 million from the U.K., spanning both Amazon's private-label brands and third-party sellers across the full product taxonomy. Figure 1 shows a screenshot of the top of a product page, along with the detailed product description.

For each product, we observe the product title, textual description, product images, sale price, manufacturer name, star rating, the number and distribution of reviews by star level, customer questions and answers, the product's position within Amazon's hierarchical category taxonomy (e.g., Home & Kitchen → Furniture → Living Room Furniture → TV & Media Furniture → Television Stands & Entertainment Centers, from broadest to most granular), sales rank within each level, image count, Buy Box status, and fulfillment method. For the majority of products, we observe the exact introduction date; for the remainder, we approximate entry using the earlier of the first customer review date and the first scraping date. These introduction dates ("Date First Available" in Figure 1b) are crucial because we

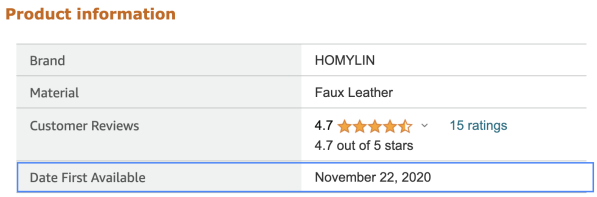
use them to identify incumbent products relative to each private-label entry.

Figure 1: Amazon Product Page

(a) Top of the product page



(b) Detailed description



Notes: Screenshots of a representative product page from the U.S. marketplace. Panel (a) shows the top of the page, including price, images, ratings, and availability. Panel (b) shows the product information section.

Our scraping infrastructure visits each product page approximately once per month, generating a panel that tracks changes in prices, ratings, review accumulation, sales ranks, and product availability over time. We identify Amazon's private-label products using a comprehensive list of confirmed Amazon-owned brands;[2] all remaining products are classified as third-party.

2.2 Amazon's Private-Label Program

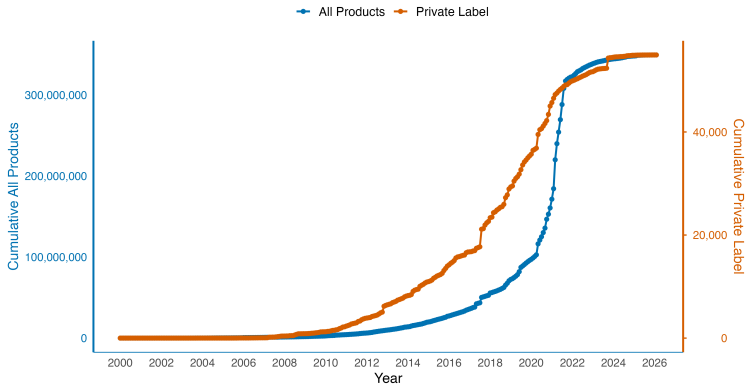
Amazon operates one of the largest private-label programs among digital platforms, marketing products under dozens of proprietary brands - including Amazon Basics, Amazon Essentials, Solimo, Presto!, Goodthreads, Lark & Ro, and Stone & Beam - spanning consumer electronics accessories, office supplies, apparel, home furnishings, pet supplies, and grocery.

Figure 2 plots the cumulative number of private-label and third-party seller products over time for Amazon US. The platform's product catalog grew rapidly after 2015, exceeding

[2]See https://www.pattern.com/blog/amazons-private-label-brands-complete-guide.

300 million products by the early 2020s. Amazon's private-label portfolio, while orders of magnitude smaller, expanded steadily in parallel and has surpassed 40,000 cumulative products.

Figure 2: Cumulative number of private-label and third-party seller products over time



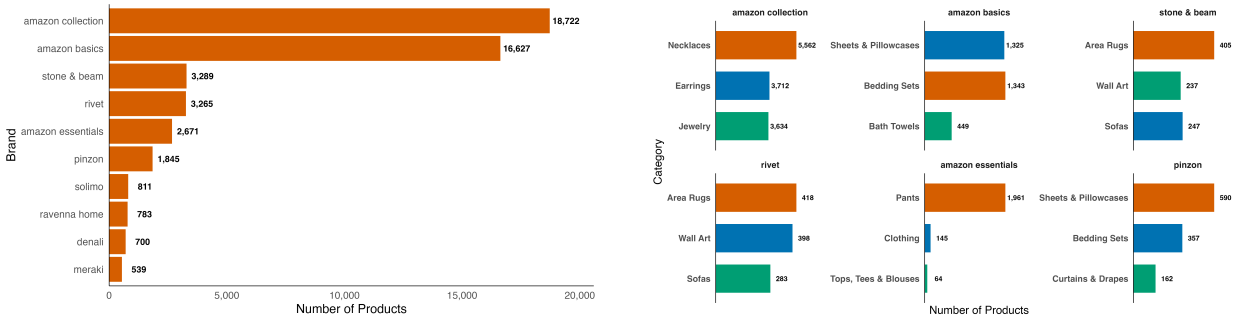
Notes: Cumulative number of products by first-availability date, U.S. marketplace. The left axis (blue) plots all products; the right axis (orange) plots Amazon private-label products, identified using the brand list in Appendix Table A1.

Figure 3a displays the top 10 private-label brands by product count.

Figure 3: Amazon private-label brand composition

(a) Top 10 brands by product count

(b) Category distribution for top six brands



Notes: Panel (a) shows the ten largest Amazon private-label brands by number of products, U.S. marketplace. Panel (b) shows the three largest product categories for each of the top six brands.

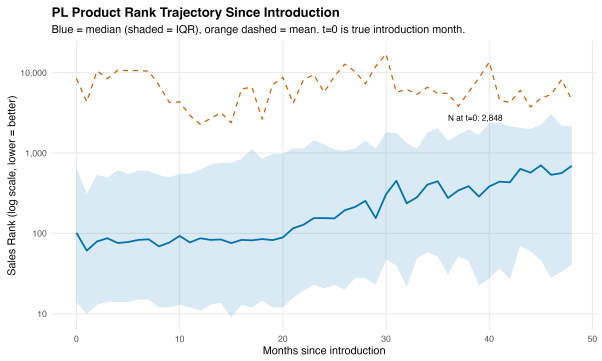
Amazon Collection and Amazon Basics together account for the majority of private-label products, with over 18,000 and 16,000 products respectively. The remaining brands are considerably smaller, with Stone & Beam, Rivet, and Amazon Essentials each contributing between 2,600 and 3,300 products. Figure 3b shows how the top six brands distribute

their products across categories. Amazon Collection concentrates in jewelry and accessories, Amazon Basics focuses on bedding and home goods, and Stone & Beam specializes in furniture. This category specialization suggests that Amazon tailors each brand identity to a specific market segment rather than deploying a single umbrella brand across all categories.

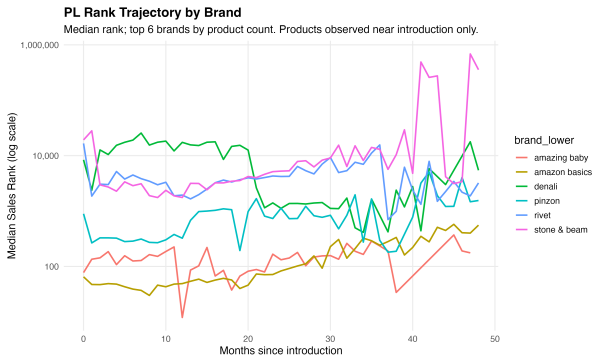
Although private labels account for a small share of the catalog, they occupy disproportionately prominent positions in the sales ranks. Figure 4a shows that the median private-label product ranks between 100 and 1,000 within its subcategory. The larger brands perform even better: as Figure 4b shows, Amazon Basics and Amazing Baby products typically rank in the top 100 to 150 of their categories.

Figure 4: Sales rank trajectories of Amazon private-label products

(a) All private-label products



(b) By brand



Notes: Sales rank (log scale; lower = better) by months since product introduction, U.S. marketplace. Panel (a): solid line is the median across all private-label products, shaded region is the interquartile range, dashed line is the mean. Panel (b): median rank for the six largest brands by product count for which we had complete sales rank history.

3 Evidence on Product Imitation

To understand whether private label products resemble existing third-party sellers, we focus on 14 product categories with substantial private-label presence. Computing pairwise visual similarity across the full catalog of hundreds of millions of products is computationally prohibitive, requiring us to focus on a subset of categories. We select categories where

products should exhibit meaningful visual heterogeneity: pet supplies (dog beds and mats), office chairs, cat condos and cat trees, dinnerware sets, dutch ovens, bookcases, desks, suitcases, baby blankets, bath towels, accent chairs, end tables, ottomans, and food storage containers. The heterogeneity should be in distinct shapes, colors, and design features that allow our similarity model to discriminate between products. Categories such as cables, adapters, or standardized electronic accessories, where nearly all products look alike regardless of manufacturer, would yield uninformative visual similarity scores. Within the above mentioned 14 categories, our sample contains 1,673 Amazon private-label products and over 5.5 million third-party products.

A note on sample scope: the imitation analysis, consumer outcomes (Section 4), and product-level regressions (Section 5) are conducted using the U.S. marketplace only, as this is the largest market with the most comprehensive data coverage. The category-level difference-in-differences analysis (Section 6) pools all four marketplaces and exploits within-category cross-country variation in the timing of private-label entry for identification.

3.1 Visual Similarity Measurement

To measure visual resemblance between products, we train a Siamese neural network - a model designed for learning similarity metrics (Bromley et al., 1993; Chopra et al., 2005). The network consists of two identical weight-sharing subnetworks, each mapping a product image to a $d = 512$-dimensional embedding. Let $f_\theta(\cdot)$ denote the embedding function. The training objective minimizes:

$$L(\theta) = \sum_{x_a, x_b} \underbrace{L_b(x_a, x_b)}_{\substack{\text{Penalty if similar images} \\ \text{are far apart}}} + \sum_{x_b, x_c} \underbrace{L_c(x_b, x_c)}_{\substack{\text{Penalty if dissimilar images} \\ \text{are nearby}}} \tag{1}$$

where L_b is a contrastive loss that penalizes large distances between embeddings of visually similar products, and L_c penalizes small distances between dissimilar products. We construct training pairs using products sharing the same manufacturer and narrowest category (similar)

and products from different categories (dissimilar).

We measure similarity between products i and j as:

$$\text{sim}(i,j) = \frac{f_\theta(x_i) \cdot f_\theta(x_j)}{\|f_\theta(x_i)\| \, \|f_\theta(x_j)\|} \tag{2}$$

A cosine similarity of 0.9 or above corresponds to near-identical appearance - for example, office chairs of the same design differing only in branding. We compute cosine similarity for each private-label product against all third-party products within the same category. We complement visual similarity with textual similarity based on transformer-based sentence embeddings of product titles and descriptions. See Appendix E for more details.

We classify a third-party product as "potentially imitated" if there exists a private-label product in the same category with visual cosine similarity ≥ 0.9 that was introduced after the third-party product.

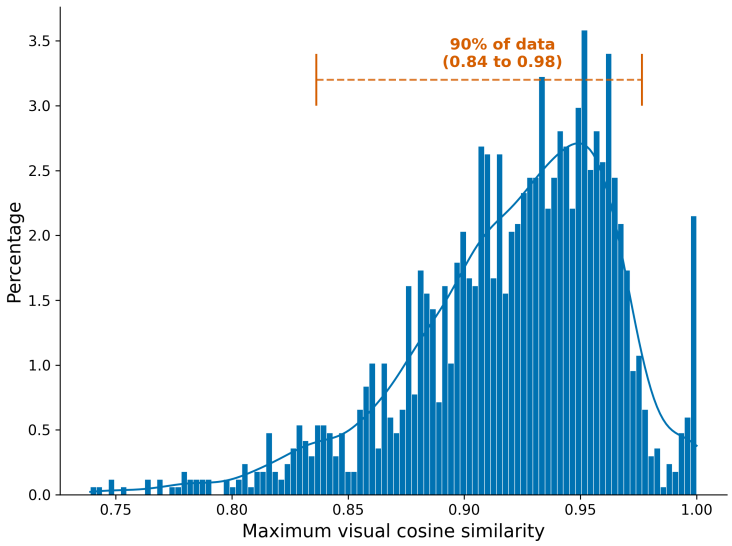
Visual similarity among products may be consistent with deliberate imitation, but it may also reflect convergent design within a category. However, extremely high similarity of 0.9 and above is more consistent with deliberate product imitation, as will become clear from the analysis below.

3.2 Prevalence of Visual Similarity

For each of the 1,673 private-label products in our focal categories, we compute the maximum visual cosine similarity relative to any pre-existing third-party product. Figure 5 plots this distribution.

Out of 1,673 private-label products, 1,222 (73%) have at least one pre-existing match with cosine similarity of 0.9 or above. The distribution is notably right-skewed, with a pronounced mode near 0.93 and 90% of observations falling between 0.84 and 0.98. The median maximum similarity exceeds 0.92, and even the 25th percentile exceeds 0.85. Only 9% of private-label products have a maximum similarity below 0.80, indicating that virtually all private-label

Figure 5: Distribution of maximum visual cosine similarity between each private-label product and pre-existing third-party products



Notes: For each of the 1,673 Amazon private-label products in the 14 focal categories, we compute the maximum visual cosine similarity relative to any third-party product introduced earlier in the same category. U.S. marketplace. Dashed lines mark the range containing 90% of observations (0.84 to 0.98).

entries bear meaningful visual resemblance to at least one pre-existing third-party product.

At 0.9 and above, products share the same silhouette, color palette, and design elements, differing only in branding. The pervasiveness of high-similarity matches across categories - office chairs, dutch ovens, dinnerware, end tables, suitcases, dog beds - indicates a broad strategy of introducing private labels that closely resemble successful existing offerings.

Figures 6a and 6b display representative product pairs; additional examples appear in Appendix B.1. In the dog beds example (Figure 6a), Amazon Basics offers a product at $37.98 that closely resembles a third-party product priced at $29.99, with a cosine similarity of 0.96. In the dutch ovens example (Figure 6b), the Amazon Basics product ($46.28) is visually near-identical to a third-party product priced at $422.99, achieving a perfect cosine similarity of 1.00 - yet the private-label version is priced at roughly one-tenth of the original. These examples illustrate two distinct patterns: in some cases Amazon may provide a more expensive item, while in others it offers dramatically cheaper alternatives to premium products.

Figure 6: Examples of visual similarity between Amazon private-label products and third-party predecessors

(a) Dog beds (cosine similarity: 0.96)



(b) Dutch ovens (cosine similarity: 1.00)



Notes: Each panel shows an Amazon private-label product (left) and its most visually similar pre-existing third-party product (right), with prices and seller names as listed at the time of data collection. Cosine similarity computed from the Siamese network embeddings described in Section 3.1.

Table 1: Characteristics of potentially imitated vs. non-imitated third-party products

	Mean Imitated	Mean Non-Imitated	Difference	p-value	
Sale Price ($)	71.7	112.1	−40.4	0.000	***
Star Rating	4.1	4.1	−0.0	0.000	***
Review Count	918.2	666.4	251.8	0.000	***
Image Count	3.8	5.1	−1.3	0.000	***
Buy Box Rank	1.0	1.0	0.0	0.000	***
One-Star Reviews	43.8	33.0	10.8	0.400	
Two-Star Reviews	29.5	20.4	9.0	0.300	
Three-Star Reviews	62.7	42.7	20.0	0.010	**
Four-Star Reviews	119.8	84.7	35.1	0.003	***
Five-Star Reviews	662.4	485.6	176.7	0.000	***
$N_{\text{imitated}} = 1{,}222$ products;		$N_{\text{non-imitated}} = 5{,}506{,}926$ products;		14 categories	

Notes: Stars indicate statistical significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Imitated products are defined as third-party products with visual cosine similarity ≥ 0.9 relative to at least one subsequently introduced Amazon private-label product. U.S. marketplace, 14 focal categories. The star rating difference rounds to −0.0 but is statistically significant due to the large sample size; the unrounded difference is −0.04 stars.

Table 1 reports mean comparisons between the product characteristics of imitated versus non-imitated third-party products. Imitated products are substantially cheaper ($71.7 versus $112.1), consistent with Amazon targeting mass-market price points. They have significantly more reviews (918 versus 666), with the difference driven by four- and five-star reviews (35 and 177 more on average, respectively) - signals of commercial success that are publicly observable but may also correlate with internal metrics available only to the platform. Table 2 extends this analysis to sales rank and seller characteristics. Imitated products hold significantly better (lower) sales ranks at every taxonomy level. At the broadest level (Category 1, e.g., Home & Kitchen), imitated products rank approximately 187,000 positions higher; at the intermediate level (Category 2, e.g., Furniture), 8,000 positions higher; and at a more granular level (Category 3, e.g., TV Stands), 22,000 positions higher.

Table 2: Sales ranks and seller characteristics: imitated vs. non-imitated products

	Mean Imitated	Mean Non-Imitated	Difference	p-value	
Panel A: Sales Rank					
Category 1 Rank	3,039,457	3,225,970	−186,513	0.000	***
Category 2 Rank	7,887	16,005	−8,119	0.000	***
Category 3 Rank	21,183	43,165	−21,982	0.000	***
Panel B: Seller Characteristics					
Number of Products	5,057	1,531	3,526	0.000	***
Seller Tenure (days)	3,703	2,968	735	0.000	***
$N_{\text{imitated}} = 1{,}222$ products; $N_{\text{non-imitated}} = 5{,}506{,}926$ products; 14 categories					

Notes: Lower sales rank indicates higher sales. Seller tenure is measured as the number of days between the seller's first product listing and the date of the private-label entry. U.S. marketplace, 14 focal categories. Same sample as Table 1. *** $p < 0.01$.

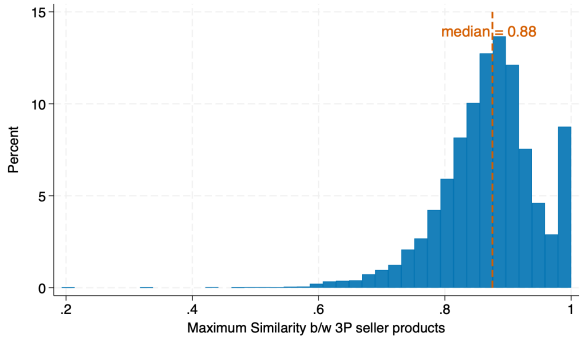
Imitated products are sold by larger sellers (5,057 versus 1,531 products) who have been on the platform roughly two years longer (3,703 versus 2,968 days). Taken together, Tables 1 and 2 show that Amazon's private-label entry disproportionately targets cheaper, more popular, better-ranked products sold by larger, more experienced sellers. The extreme similarity of some products (a large share with cosine similarity ≥ 0.9) suggests that Amazon targets *specific* products rather than aggregate category patterns alone.

3.3 Third-Party Imitation Patterns

Imitation is not unique to the platform: third-party sellers routinely introduce products that closely resemble existing offerings from other third-party sellers. Figure 7a plots the distribution of maximum cosine similarity among third-party products; the median is 0.88, compared to 0.92 for Amazon's private labels. Figure 7b shows an example of two nearly identical chairs sold by different third-party sellers (cosine similarity: 0.97). However, the characteristics of products targeted by third-party imitators differ systematically from those targeted by Amazon.

Figure 7: Visual similarity among third-party seller products

(a) Maximum visual cosine similarity between third-party seller products



(b) Example: chairs sold by two third-party sellers (cosine similarity: 0.97)



Notes: Panel (a) plots, for each third-party product, the maximum visual cosine similarity relative to any subsequently introduced third-party product (Amazon private labels excluded); the dashed line marks the median (0.88). Panel (b) shows an example pair of near-identical office chairs sold by two different third-party sellers. U.S. marketplace, 14 focal categories.

The price gap between potentially imitated and non-imitated products is much smaller for third-party imitation ($7.9 versus $40.4 for Amazon). The review count difference is statistically insignificant for third-party imitation, whereas it is large and highly significant for Amazon. The five-star review surplus that distinguishes Amazon's targets is absent for third-party imitators. Third-party imitators also show weaker selection on granular sales rank (see Appendix Table A4).

Two interpretations of these asymmetries merit consideration. First, Amazon's

Table 3: Third-party imitation: characteristics of potentially imitated vs. non-imitated products

	Mean Imitated	Mean Non-Imitated	Difference	p-value	
Sale Price ($)	51.0	58.9	−7.9	0.000	***
Star Rating	4.1	4.1	−0.1	0.000	***
Review Count	1,650.8	1,320.6	330.3	0.400	
Image Count	5.4	5.2	0.2	0.000	***
Buy Box Rank	1.0	1.0	0.0	0.000	***
One-Star Reviews	109.9	115.6	−5.7	0.900	
Two-Star Reviews	69.8	72.3	−2.6	1.000	
Three-Star Reviews	140.2	135.6	4.6	0.900	
Four-Star Reviews	247.9	245.1	2.8	1.000	
Five-Star Reviews	1,431.6	1,407.6	23.9	0.900	

Notes: Imitated products are defined as third-party products with visual cosine similarity ≥ 0.9 relative to at least one subsequently introduced third-party product (excluding Amazon private labels). U.S. marketplace, 14 focal categories, $N \approx 5.5$ million third-party products. The sample differs from Tables 1–2 because the comparison here is among third-party products only (Amazon private labels excluded from the imitator definition). Stars indicate statistical significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

systematic selection on sales rank, review volume, and seller size could reflect access to proprietary marketplace data - such as detailed conversion rates, search analytics, and inventory turnover - unavailable to ordinary sellers. Second, the differences may reflect distinct competitive strategies rather than differential information: third-party imitators may enter at lower price points to undercut incumbents, while Amazon leverages its brand recognition and distribution infrastructure to target the most commercially established products. Our evidence cannot definitively distinguish between these channels. The systematic selection on granular sales rank - a dimension more closely tied to internal platform data than to publicly observable signals - is suggestive, but we acknowledge that a well-resourced firm could plausibly replicate this targeting through careful analysis of public information alone.

Overall, although Amazon appears to target specific products when developing private labels, similar practices are also common among third-party sellers. As a result, such conduct would be difficult to regulate, since other sellers with sufficient resources could, in principle, do the same. We next turn to how private labels affect consumer and seller outcomes.

4 Consumer Outcomes

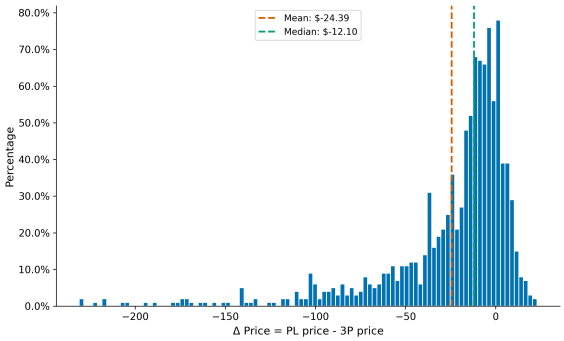
The comparisons in this section are descriptive - they characterize the differences in product prices and proxies to consumer satisfaction outcomes (e.g., ratings and reviews). We compare *private-label products directly to their closest third-party matches.*

4.1 Price Differences

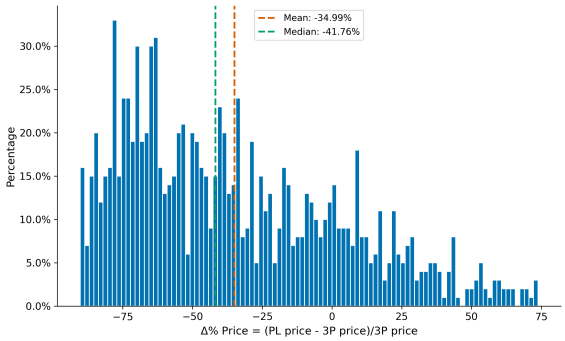
For each private-label product with a highly similar third-party predecessor (cosine similarity ≥ 0.9), we compute the price gap at the time of private-label introduction. Figure 8a plots the distribution of $\Delta\text{Price} = P^{PL} - P^{3P}$.

Figure 8: Price differences between private-label products and their most similar third-party predecessors

(a) Absolute differences ($)



(b) Percentage differences



Notes: Sample restricted to the U.S. marketplace. Each observation is a private-label product paired with its most similar third-party predecessor (cosine similarity ≥ 0.9).

On average, Amazon's private-label products are $24 cheaper than their closest third-party matches, with the vast majority priced below their counterparts (Figure 8a). The distribution is sharply left-skewed: most price differences are concentrated between −$50 and $0, with a small number of product pairs exhibiting gaps exceeding −$100, typically in higher-priced categories such as furniture and kitchen equipment. In relative terms, the median private-label product is approximately 42% cheaper (Figure 8b), and

fewer than 15% of private-label products are priced above their most similar third-party predecessor. A consumer who switches to the near-identical private-label alternative thus saves more than two-fifths of the purchase price.

In some categories, we can directly test whether the lower prices are associated with lower quality production material. We use textual similarity measures to do this analysis. For example, in durtch ovens, Table 4 shows that private labels enter most dutch-oven material types. For the same material of the dutch oven, private label prices are roughly one-third to two-fifths of what third-party sellers charge.

Table 4: Average price by material: Amazon PL vs. third-party (cookware)

Material	PL Avg Price	3P Avg Price	Ratio
Enameled cast iron	$69	$211	0.33
Cast iron (plain)	$67	$200	0.33
Non-stick	$57	$163	0.35
Ceramic	-	$173	-
Aluminum	$42	$98	0.43
Stainless steel	$65	$159	0.41
Clay / Terracotta	-	$170	-
Porcelain	-	$156	-
Copper	-	$265	-

Notes: Average sale price among products classified under each material. Ratio = PL price / 3P price. " - " indicates fewer than 3 PL products in that material. Amazon PL products are priced at roughly one-third of the third-party average within enameled cast iron, their primary material.

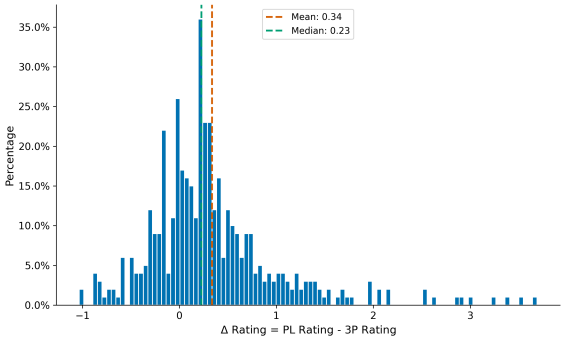
4.2 Ratings and Reviews

Figure 9a plots rating differences between private-label products and their most similar third-party predecessors. The mean rating advantage is 0.34 stars and the median is 0.23 stars. Despite entering the market an median of 960 days after their most similar third-party predecessor (Appendix A7), Amazon's private-label products also accumulate far more customer reviews (Figure 9b). Existing literature shows that even small rating differences have meaningful demand implications: Magnusson (2022) estimates that a half-star increase

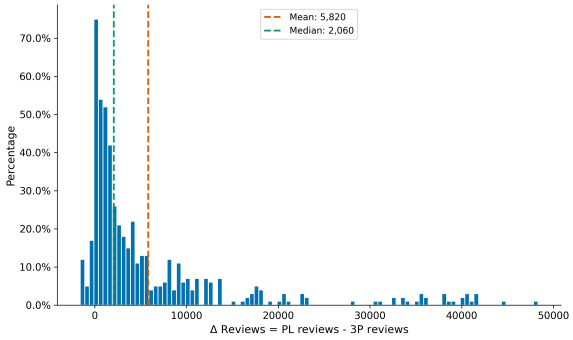
on Wayfair causes a 5% increase in demand. If this result holds in Amazon setting too, that would be equivalent to a 3.4% average increase in revenue for private label products.

Figure 9: Rating and review differences between private-label products and their most similar third-party predecessors

(a) Star rating differences



(b) Review count differences



Notes: Sample restricted to the U.S. marketplace. Each observation is a private-label product paired with its most similar third-party predecessor (cosine similarity ≥ 0.9).

The review count distribution is heavily right-skewed, with the mean difference exceeding 5,800 reviews and the median exceeding 2,000. Table 5 decomposes this advantage by star rating: the surplus is overwhelmingly concentrated in five-star reviews (mean: 4,475 more; median: 1,431), while the one-star surplus is much smaller (mean: 157; median: 50). Thus, purchasing consumers leave better reviews for the private label products.

Table 5: Review count differences by star rating

Star Rating	Mean Difference	Median Difference
1-Star	156.9	50.0
2-Star	125.8	34.0
3-Star	321.5	95.0
4-Star	740.5	254.0
5-Star	4,475.2	1,431.0

Notes: Differences computed as $N_s^{PL} - N_s^{3P}$ for each star level $s \in \{1, 2, 3, 4, 5\}$. Positive values indicate the private-label product has more reviews. Based on 1,222 matched PL–3P product pairs. U.S. marketplace.

The large review volume advantage implies substantially higher sales volumes, higher review propensity, or both. Contributing factors may include lower prices, prominent

algorithmic placement (Farronato et al., 2022), the "Amazon's Choice" badge, consumer trust in Amazon-branded products, selection bias (Amazon may launch private labels only for products it expects to perform well), and Amazon's Vine early reviewer program. The short-run consumer implications appear favorable, but the long-run welfare consequences depend on how private-label entry affects third-party sellers' incentives - a question we address next.

5 Product-Level Effects

We now turn to the effects of private-label entry on third-party sellers. We begin with product-level estimates that isolate the effect on individual products directly targeted by Amazon's private labels, before examining broader category-level outcomes in Section 6.

Treatment is defined at the product level: for each third-party product matched to an Amazon private label (visual cosine similarity ≥ 0.9), the treatment date is the week in which the similar private-label product first appears. The control group consists of third-party products in the same category that closely resemble the treated product ($\text{sim}(i,j) \geq 0.9$) but are dissimilar to any private-label product ($\max_k \text{sim}(j,k) < 0.9$). The idea is similar in spirit to a regression discontinuity: control products closely resemble the eventually imitated ones but retain features that distinguish them from the private-label products. A potential concern is that private-label entry may eventually affect *all* products in a category, including our controls; the category-level analysis in Section 6 addresses this by estimating effects on the full category.

We estimate the following difference-in-differences specification at the product–week level:

$$Y_{it} \;=\; \beta\,\text{Post}_{it} \;+\; \alpha_i \;+\; \gamma_{g(i),t} \;+\; \varepsilon_{it}, \tag{3}$$

where Y_{it} is an outcome of product i in week t (log sales rank, log granular sales rank, log price, or an indicator for discontinuation), and Post_{it} is an indicator equal to one if

product i is treated and week t falls on or after the entry week of the matching private-label product. The product fixed effects α_i absorb all time-invariant product characteristics, and the group-year-by-week fixed effects $\gamma_{g(i),t}$ - where $g(i)$ denotes the cohort of products treated in the same period - flexibly control for common shocks within cohorts, so that β is identified from comparisons between treated products and their matched controls within the same cohort and week. Standard errors are clustered at the product level. Table 6 reports the results.

Two features of the sample sizes in Table 6 merit explanation. First, the sample of 4,141 matched third-party products exceeds the 1,222 directly imitated products identified in Section 3 because a single treated product can have multiple matches: several third-party products in the same category may satisfy the control criteria of being highly similar to the treated product while remaining dissimilar to any private label. Second, although all outcomes are measured on the same set of 4,141 matched products, sales rank histories are incomplete for a subset of them due to intermittent gaps in our scraping infrastructure. The sales-rank regressions are therefore estimated on the 1,831 products with observed rank data, while the price and discontinuation regressions use the full matched sample. Because these gaps arise from our data collection process rather than from product characteristics or performance, the missingness is plausibly unrelated to the outcomes of interest.

After Amazon introduces a visually similar private-label product, matched third-party products experience a statistically significant worsening of sales rank: 4.5% at the broad taxonomy level and 4.0% at the granular level. Because sales rank is a monotonic function of recent sales velocity, the rank deterioration implies a decline in relative sales performance. However, the mapping from rank to sales is nonlinear, so these coefficients should be interpreted as directional evidence of competitive harm rather than as precise estimates of sales losses. Matched products also reduce their prices by 1.1%, consistent with competitive pressure from the lower-priced private-label entrant. This price decline is modest relative to the 42% median price gap documented in Section 4, suggesting that incumbents adjust

Table 6: Product-level effects of Amazon private-label entry on matched third-party products

	Log(Sales Rank)	Log(Sales Rank, Granular)	Log(Price)	Discontinued
After Amazon Entry	0.045***	0.040***	−0.011***	0.003**
	(0.010)	(0.009)	(0.001)	(0.001)
Baseline Mean	13.337	7.470	5.126	0.001
Product FE	✓	✓	✓	✓
Group-Year × Week FE	✓	✓	✓	✓
# Matched 3P Products	1,831	1,831	4,141	4,141
Observations	108,446	108,446	292,897	292,897

Notes: Each column reports the coefficient on a post-treatment indicator from a product-level difference-in-differences regression. Treated products are third-party products in the U.S. marketplace (14 focal categories) with visual cosine similarity ≥ 0.9 to a subsequently introduced Amazon private-label product; control products are similar third-party products in the same category that are dissimilar to any private label (see text). The number of matched products varies across specifications because sales rank data is available for fewer products than price and discontinuation data. Standard errors clustered at the product level in parentheses. ** $p < 0.05$, *** $p < 0.01$.

prices only partially rather than fully matching the private-label price.

The probability of product discontinuation increases by 0.3 percentage points. While the magnitude appears small in absolute terms, the baseline discontinuation rate in our sample is only 0.1% per period, so the effect represents a several-fold increase in the likelihood of exit. The discontinuation result thus provides direct evidence that private-label entry contributes to product exit.

These product-level results demonstrate that the competitive effects of private-label entry extend directly to the performance and survival of specific targeted incumbents. In the next section, we complement these estimates with category-level analysis to assess the broader competitive restructuring induced by private-label entry.

6 Category-Level Effects

6.1 Identification

We estimate the effect of Amazon's private-label entry on category-level outcomes using a cross-country staggered difference-in-differences design. The unit of observation is a category–country–month cell (c, o, t), with $o \in \{\text{US}, \text{DE}, \text{UK}, \text{CA}\}$, and treatment is the first appearance of an Amazon private-label product in category c in country o. Identification comes from the fact that Amazon launches private labels in the same category at different times across countries: if a category is treated in the U.S. in January 2021 but in the U.K. not until September 2021, the U.K. cell serves as a control for the U.S. cell during the intervening months. We accordingly restrict the sample to the 747 categories with such identifying variation - staggered treatment timing across at least two countries, or a mix of treated and not-yet-treated country cells - so that estimates rely on within-category comparisons rather than comparisons across categories.

Because treatment is staggered, the standard two-way fixed effects estimator may be biased under heterogeneous treatment effects (de Chaisemartin and D'Haultfœuille, 2020; Goodman-Bacon, 2021; Sun and Abraham, 2021). We therefore implement the Callaway and Sant'Anna (2021) estimator with not-yet-treated controls:

$$ATT(g,t) = \mathbb{E}\big[Y_t(g) - Y_t(0) \mid G = g\big] - \mathbb{E}\big[Y_t(g) - Y_t(0) \mid C = 1\big], \tag{4}$$

where g indexes the cohort (month of first treatment for a category–country cell) and $C = 1$ denotes cells not yet treated at time t - including, critically, cells for the same category in other countries. The overall ATT aggregates across cohorts and post-treatment periods. Standard errors are clustered at the category level, and we report bootstrap confidence intervals.

The identifying assumption is parallel trends: conditional on group and time effects,

treated cells would have evolved like not-yet-treated cells absent entry. The cross-country design strengthens this assumption, since within-category comparisons absorb global category-level shocks (e.g., a worldwide technology shift in a product category) and within-country comparisons absorb aggregate country shocks (e.g., macroeconomic conditions or regulatory changes).

Finally, for the intensive-margin outcomes (prices, star ratings, out-of-stock rates), category aggregates are computed over incumbent products only - those first listed before private-label entry in the cell - isolating the effect on the pre-existing competitive landscape. The extensive-margin outcomes (entries and exits) include all products.

6.2 Treatment Variation

The cross-country design requires that treatment timing varies meaningfully across countries within the same category. Of the 1,015 categories containing at least one Amazon private-label product, 761 are listed in at least two countries. Of these, 747 exhibit identifying variation for the cross-country design: 222 have staggered treatment timing (private-label entry occurs in different months across countries) and 653 include a mix of treated and never-treated country cells; the two groups overlap, and the analysis sample is their union ($222 + 653 - 128 = 747$). The remaining 14 multi-country categories are treated simultaneously in every country in which they appear, leaving no untreated control, and are dropped (Table 7).

Table 8 reports pairwise treatment timing across country pairs. No country systematically moves first: the U.S. precedes Germany in 60% of shared categories but trails the U.K. (40% versus 49%), and the remaining pairs are similarly mixed, with a nontrivial share of categories treated in the same month. This pattern matters for identification. If one country always led private label entry, lagging countries' entries might be mechanically triggered by the leader's - for instance, a successful U.S. launch prompting rollout elsewhere - which would tie treatment timing to category performance and threaten parallel trends. The absence of

Table 7: Sample construction: cross-country CS analysis

	Categories
Total private-label categories	1,015
Listed in 1 country only (dropped)	254
Listed in ≥2 countries	761
Sources of identifying variation (761 multi-country categories)	
Staggered treatment timing across countries	222
Mix of treated and never-treated countries	653
of which satisfy both conditions	*128*
Treated simultaneously everywhere, no control (dropped)	14
Analysis sample (with identifying variation)	**747**
Category–country cells	2,138

a consistent leader argues against such a channel.

Table 8: Pairwise cross-country treatment timing (CS analysis sample)

Country pair	N (categories)	Median gap (months)	Mean gap (months)	% first (col. 1)	% first (col. 2)	% same month
US–DE	42	5	6.0	59.5%	31.0%	9.5%
US–UK	83	6	10.4	39.8%	49.4%	10.8%
US–CA	130	4	9.9	37.7%	36.9%	25.4%
DE–UK	50	7	8.3	36.0%	46.0%	18.0%
DE–CA	38	7	9.0	39.5%	50.0%	10.5%
UK–CA	79	6	11.7	49.4%	41.8%	8.9%

Notes: Categories treated in both countries of the pair. CS analysis sample. "Gap" = $|t_{\text{first},c_1} - t_{\text{first},c_2}|$ in months. "% first" columns report the share of categories in which each country of the pair is treated first.

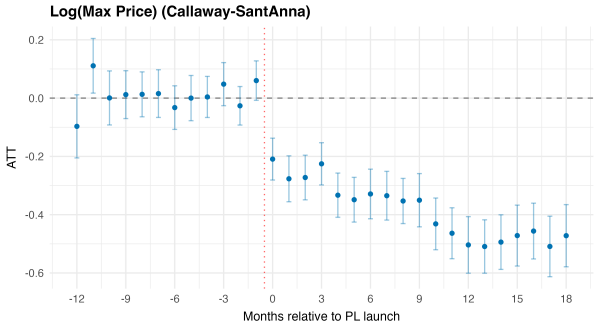
6.3 Category-level results

Private-label entry compresses the price distribution within a category from both tails. Figure 10 plots the event studies for maximum, minimum, and average prices. Maximum prices decline sharply at treatment ($ATT = -0.58$, SE = 0.054), corresponding to a 44% decline - the highest-priced products either reduce their prices or exit. Simultaneously, minimum prices *rise* ($ATT = +0.84$, SE = 0.059), a 133% increase - the cheapest products exit or raise prices. Average prices begin to drift upward around the 16th month post-treatment, but the overall effect is statistically insignificant ($ATT = +0.01$,

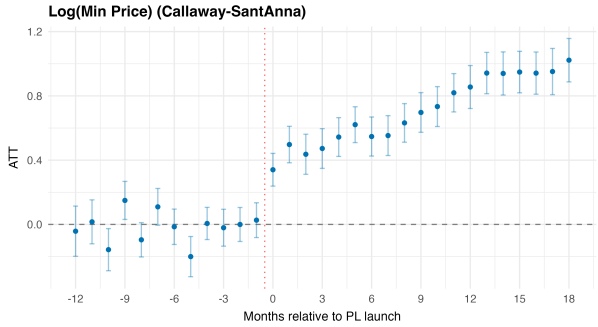
SE = 0.025). The effect is thus purely distributional: private-label entry compresses the product space toward the middle of the price distribution, eliminating both the cheapest and the most expensive offerings.

Figure 10: Effect of private-label entry on prices

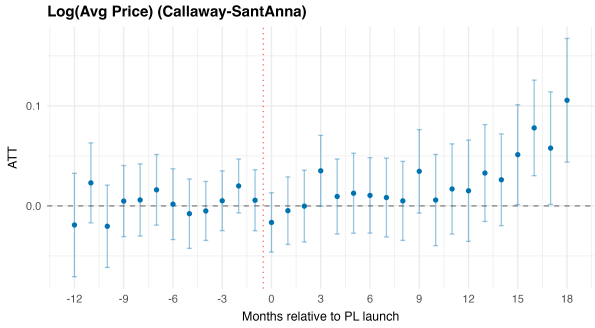
(a) Log maximum price



(b) Log minimum price



(c) Log average price



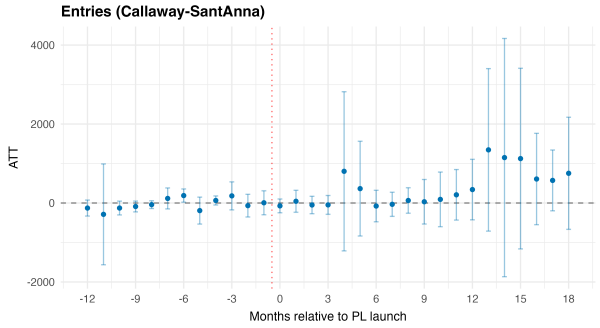
Notes: Callaway and Sant'Anna (2021) event studies with not-yet-treated controls on the pooled cross-country panel (US, DE, UK, CA). The unit of observation is a category–country–month cell; outcomes are computed over incumbent third-party products (products first listed before private-label entry in the cell). Month 0 is the first Amazon private-label entry in the category–country cell. Vertical bars are 95% confidence intervals from a multiplier bootstrap (1,000 iterations), clustered at the category level.

Private-label entry does not appear to affect seller entry: the effects are statistically insignificant throughout the entire 18-month post-treatment horizon (Figure 11a). There is, however, a significant delayed effect on exits: following private-label entry, approximately 489 additional sellers exit a treated category–country cell on average (Figure 11b). Closer to the 18-months the effect is too noisy with only some months reaching significance, but there is an increased trajectory.

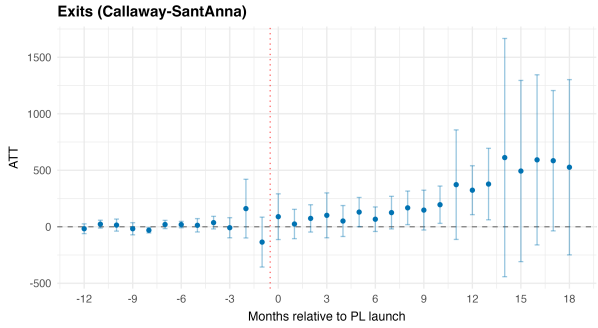
We then investigate the impact on out-of-stock rates. Out-of-stock rates also rise

Figure 11: Effect of private-label entry on third-party seller entry and exit

(a) Third-party seller entry



(b) Third-party seller exit



Notes: Callaway and Sant'Anna (2021) event studies with not-yet-treated controls on the pooled cross-country panel (US, DE, UK, CA). Outcomes are the monthly number of third-party seller entries (a) and exits (b) in a category–country cell; these extensive-margin outcomes include all products. Month 0 is the first Amazon private-label entry in the cell. Vertical bars are 95% confidence intervals from a multiplier bootstrap (1,000 iterations), clustered at the category level.

significantly (Figure 12a): the aggregate ATT is 4.4 percentage points. On Amazon, sellers often do not formally discontinue a product; instead, the listing permanently goes out of stock and receives a "Currently Unavailable" badge. This outcome therefore likely captures an additional margin of seller withdrawal beyond the exits documented above. This is consistent with the product-level results in Section 5, where targeted sellers discontinue their products.

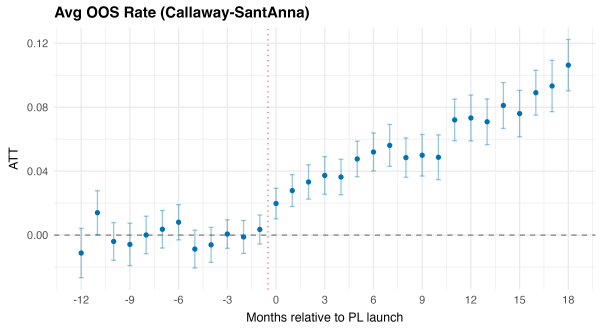
Figure 13 shows that incumbent third-party products experience an increase in both average and median sales ranks over time. Because a higher rank indicates lower sales, incumbent products perform systematically worse after private-label entry.

Figure 12b shows that average star ratings are essentially unaffected by private-label entry ($ATT = -0.03$, SE $= 0.013$). While statistically significant, the magnitude is economically negligible - a 0.03-star decline on a 5-point scale.

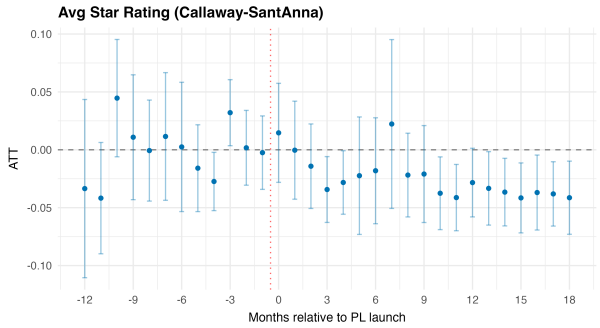
Table 9 collects the aggregate ATT estimates for all outcomes. The dominant pattern is competitive contraction - accelerated seller exit, rising out-of-stock rates, deteriorating incumbent sales ranks, and price compression toward the middle of the distribution - with magnitudes that are economically large. Product quality as measured by star ratings is

Figure 12: Effect of private-label entry on third-party seller out-of-stock rates and star ratings

(a) Out-of-stock rate



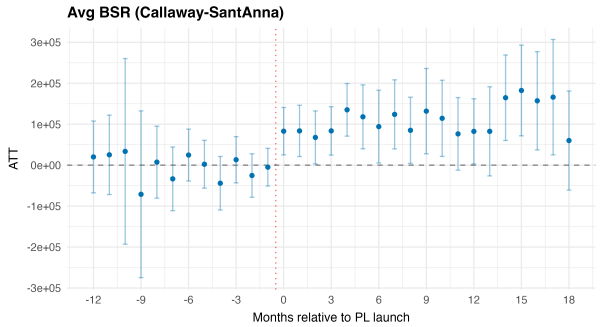
(b) Star rating



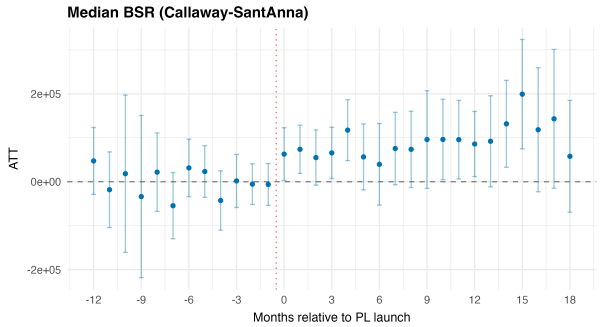
Notes: Callaway and Sant'Anna (2021) event studies with not-yet-treated controls on the pooled cross-country panel (US, DE, UK, CA). Outcomes are the share of incumbent third-party products listed as out of stock (a) and the average star rating of incumbent products (b) in a category–country cell. Month 0 is the first Amazon private-label entry in the cell. Vertical bars are 95% confidence intervals from a multiplier bootstrap (1,000 iterations), clustered at the category level.

Figure 13: Effect of private-label entry on third-party seller ranks

(a) Average rank



(b) Median rank



Notes: Callaway and Sant'Anna (2021) event studies with not-yet-treated controls on the pooled cross-country panel (US, DE, UK, CA). Outcomes are the average (a) and median (b) best-seller rank among incumbent third-party products in a category–country cell; a higher rank indicates worse sales performance. Month 0 is the first Amazon private-label entry in the cell. Vertical bars are 95% confidence intervals from a multiplier bootstrap (1,000 iterations), clustered at the category level.

essentially unchanged.

Table 9: Effects of private-label entry: Callaway–Sant'Anna ATT (cross-country panel)

Outcome	Baseline mean	ATT	SE	% Δ
Panel A: Competition				
Entries	558	−37	(160)	−6.6
Exits	1,577	489***	(150)	31
Avg OOS Rate	0.41	0.044***	(0.006)	10.7
Panel B: Prices				
Log(Max Price)	4,162	−0.576***	(0.054)	−43.8
Log(Min Price)	6.55	0.844***	(0.059)	133
Log(Avg Price)	142	0.009	(0.025)	0.9
Panel C: Seller Rank				
Avg BSR	872,290	126,690***	(42,624)	14.5
Median BSR	612,296	104,013**	(46,798)	17.0
Panel D: Product Quality				
Avg Star Rating	4.10	−0.034**	(0.013)	−0.8
Categories		747		
Category–country cells		2,138		

Notes: Each row reports the aggregate *ATT* from the Callaway and Sant'Anna (2021) estimator with not-yet-treated controls on the pooled cross-country panel (US, DE, UK, CA). The unit of observation is a category–country–month cell. Treatment is first Amazon private-label entry in a given category and country. Pre-mean is the average of the outcome variable across treated units in pre-treatment periods (level means shown for log outcomes). % Δ is the implied percentage change: $ATT/\overline{Y}_{\text{pre}} \times 100$ for level outcomes, $(e^{ATT} - 1) \times 100$ for log outcomes. Standard errors from multiplier bootstrap (1,000 iterations). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

6.4 Heterogeneity by Private Label Intensity

Table 10 reports the Callaway and Sant'Anna (2021) estimates separately for three groups of treated cells, defined by the maximum share of private-label products among all active products observed post-treatment: Low (<2%, $N = 301$ cells), Medium (2–8%, $N = 152$), and High (>8%, $N = 614$).

The competitive effects intensify monotonically with private-label penetration across most outcomes. Out-of-stock rates rise by 2.6 percentage points in low-share cells but by roughly twice as much – 5.0 to 5.1 percentage points – in medium- and high-share cells.

Price compression steepens with intensity: the decline in maximum prices grows from 43% in low-share cells ($ATT = -0.57$) to 53% in medium-share cells ($ATT = -0.76$) and 59% in high-share cells ($ATT = -0.88$), while the increase in minimum prices is largest in high-share cells (157%, $ATT = 0.95$). Listing investment follows the same gradient, with the decline in average image count growing from 0.71 to 1.01 images per product.

The extensive-margin results reveal that the *composition* of the seller decline shifts with intensity. In low-share cells, the adjustment operates through exit: approximately 1,396 additional sellers exit, while entry is unaffected. In high-share cells, by contrast, entry deterrence dominates: new entry falls by roughly 258 sellers per cell, a statistically significant decline, while the exit effect is small and insignificant. (The declining exit coefficient across bins partly reflects category size: because exits are measured in levels and high-share cells are mechanically smaller categories, the same proportional exit response produces a smaller absolute count.) This pattern suggests that modest private-label presence displaces existing sellers, whereas intensive private-label presence additionally deters prospective sellers from entering at all – consistent with the "kill zone" logic of Kamepalli et al. (2020).

The rank results sharpen this interpretation. Incumbent sellers' average sales ranks are unaffected in low-share cells but deteriorate by approximately 191,000 positions in high-share cells – a large and statistically significant worsening that indicates incumbents lose sales volume precisely where Amazon's private-label footprint is largest. Product quality, by contrast, is essentially flat across all three bins: even in high-share cells, average star ratings decline by only 0.05 stars, an economically negligible magnitude.

6.5 Discussion

Four takeaways emerge from the category-level analysis.

First, the harm from private-label entry is category-wide rather than confined to the products Amazon imitates. Treated category–country cells experience a 31% increase in seller exit, a 4.4 percentage point rise in out-of-stock rates, and deteriorating incumbent

Table 10: Dose-response by PL share: Callaway–Sant'Anna ATT

	Low (<2%)		Medium (2–8%)		High (>8%)	
Outcome	ATT	SE	ATT	SE	ATT	SE
Panel A: Competition						
OOS Rate	0.026***	(0.008)	0.050***	(0.011)	0.051***	(0.008)
Exits	1,396***	(479)	745***	(233)	147	(103)
Entries	311	(419)	−2	(211)	−258**	(124)
Panel B: Prices						
Log(Max Price)	−0.570***	(0.048)	−0.755***	(0.091)	−0.883***	(0.062)
Log(Min Price)	0.839***	(0.099)	0.822***	(0.145)	0.945***	(0.080)
Panel C: Seller Rank						
Avg BSR	−103,789	(69,592)	114,081	(83,925)	191,415***	(47,949)
Panel D: Product Quality						
Avg Star Rating	−0.005	(0.014)	−0.027	(0.032)	−0.050**	(0.022)
Avg Image Count	−0.711***	(0.048)	−0.921***	(0.066)	−1.005***	(0.048)
Treated cells	301		152		614	

Notes: Each cell reports the aggregate *ATT* from the Callaway and Sant'Anna (2021) estimator with not-yet-treated controls, estimated on the subset of treated cells in the indicated PL share bin. PL share = max(cumulative PL products / total active products) observed post-treatment. Standard errors from multiplier bootstrap (1,000 iterations). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

sales ranks – effects on the full set of incumbents, not only on directly targeted products.

Second, the consumer price benefit does not spill over to the category. Average prices are unchanged after entry; the 42% saving accrues only to consumers who buy the private-label product itself. Meanwhile, the price distribution compresses from both tails – maximum prices fall 44% and minimum prices rise 133% – so the cheapest options disappear and product variety shrinks at both extremes.

Third, the harm scales with private-label penetration, and its margin shifts. At low penetration, adjustment operates through the exit of existing sellers while incumbent ranks are largely unaffected; at high penetration, incumbent ranks deteriorate sharply and new entry is significantly deterred – the pattern predicted by "kill zone" theories (Kamepalli et al., 2020).

Fourth, product quality is flat throughout: star ratings barely move in any penetration bin. The restructuring therefore reflects competitive displacement of comparable products,

not the weeding out of inferior ones.

7 Conclusion

This paper asks whether a dominant platform uses its dual role – operating a marketplace while selling on it – to the detriment of the sellers it hosts, and what this means for consumers. Two answers emerge.

On the first question, Amazon's private-label entry is not guided by aggregate category trends alone: the majority of its private-label products are near-identical copies of specific, commercially successful third-party products, sold by the largest and most established sellers. At the same time, close imitation is a marketplace-wide practice – third-party sellers routinely copy one another – and our data cannot establish whether Amazon's more precise targeting reflects non-public seller data or sophisticated analysis of publicly observable signals. The regulatory debate, which has centered on data misuse, may therefore be framed too narrowly: the conduct at issue is common to the marketplace, even if the platform executes it with greater precision.

On the second question, the welfare tradeoff is asymmetric. The consumer benefit is real but localized: it accrues only to those who buy the private-label product, does not spill over into category-wide price competition, and comes with a loss of variety as the cheapest and most expensive options disappear. The seller harm, by contrast, is broad and causal, and it compounds with Amazon's footprint: where private labels are a small presence, existing sellers exit; where the footprint is large, prospective sellers stop entering at all. Because product quality is unaffected throughout, this is the displacement of comparable products, not a cleansing of inferior ones. The entry-deterrence margin is, in our view, where the long-run harm accumulates: a marketplace that new sellers avoid loses variety and competition in ways that no contemporaneous price comparison captures.

These findings bear on regulation in three ways. First, banning platform private labels

outright is not supported by the evidence: the products deliver genuine savings to the consumers who buy them, and their entry does not degrade quality. Second, data-use restrictions of the kind embodied in the EU Digital Markets Act address only one possible channel behind Amazon's targeting – a channel our evidence is consistent with but cannot confirm – and would leave untouched the imitation that pervades the marketplace among sellers themselves. Third, and most directly supported by our estimates, policy should attend to entry incentives in heavily penetrated categories, where the deterrence of new sellers is the margin on which competition erodes.

Our reduced-form estimates characterize prices, variety, and seller survival but do not aggregate to welfare; a structural model of platform competition with endogenous entry and exit could deliver those estimates. Extending the analysis to other dual-role platforms, such as Apple's App Store or Walmart's marketplace, would test whether these patterns are general features of platform private-label entry or specific to Amazon.

References

Anderson, Simon and Özlem Bedre Defolie, "Hybrid platform model: Monopolistic competition and a dominant firm," *The RAND journal of economics*, 2024, *55* (4), 684–718.

Bell, Sean and Kavita Bala, "Learning visual similarity for product design with convolutional neural networks," *ACM Transactions on Graphics*, jul 2015, *34* (4), 1–10.

Bottou, Léon, Frank E Curtis, and Jorge Nocedal, "Optimization methods for large-scale machine learning," *SIAM review*, 2018, *60* (2), 223–311.

Bromley, Jane, Isabelle Guyon, Yann LeCun, Eduard Säckinger, and Roopak Shah, "Signature Verification Using a "Siamese" Time Delay Neural Network," in "Advances in Neural Information Processing Systems," Vol. 6 1993.

Callaway, Brantly and Pedro H. C. Sant'Anna, "Difference-in-Differences with Multiple Time Periods," *Journal of Econometrics*, 2021, *225* (2), 200–230.

Cer, Daniel, Yinfei Yang, Sheng yi Kong, Nan Hua, Nicole Limtiaco, Rhomni St John, Noah Constant, Mario Guajardo-Cespedes, Steve Yuan, Chris Tar et al., "Universal sentence encoder," *arXiv preprint arXiv:1803.11175*, 2018.

Chen, Nan and Hsin-Tien Tsai, "Steering via algorithmic recommendations," *The RAND Journal of Economics*, 2024, *55* (4), 501–518.

Chintagunta, Pradeep K, André Bonfrer, and Inseong Song, "Investigating the effects of store-brand introduction on retailer demand and pricing behavior," *Management Science*, 2002, *48* (10), 1242–1267.

Chopra, Sumit, Raia Hadsell, and Yann LeCun, "Learning a Similarity Metric Discriminatively, with Application to Face Verification," in "Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition" 2005, pp. 539–546.

Crawford, Gregory S, Matteo Courthoud, Regina Seibel, and Simon Zuzek, "Amazon entry on Amazon marketplace," Technical Report, Centre for Economic Policy Research 2022.

de Chaisemartin, Clément and Xavier D'Haultfœuille, "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects," *American Economic Review*, 2020, *110* (9), 2964–2996.

Dubé, Jean-Pierre, "Amazon Private Brands: Self-Preferencing Versus Traditional Retailing," *Antitrust Law Journal*, 2024, *86* (1), 115–148.

Farrell, Joseph and Michael L. Katz, "Innovation, Rent Extraction, and Integration in Systems Markets," *Journal of Industrial Economics*, 2000, *48* (4), 413–432.

Farronato, Chiara, Andrey Fradkin, and Alexander MacKay, "Self-Preferencing at Amazon: Evidence from Search Rankings," *AEA Papers and Proceedings*, 2022, *112*, 239–243.

Goodman-Bacon, Andrew, "Difference-in-Differences with Variation in Treatment Timing," *Journal of Econometrics*, 2021, *225* (2), 254–277.

Gutierrez, Germán, "The Welfare Consequences of Regulating Amazon," *Working Paper*, 2021.

Hadsell, Raia, Sumit Chopra, and Yann LeCun, "Dimensionality reduction by learning an invariant mapping," in "2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)," Vol. 2 IEEE 2006, pp. 1735–1742.

Hagiu, Andrei and Daniel Spulber, "First-Party Content and Coordination in Two-Sided Markets," *Management Science*, 2013, *59* (4), 933–949.

_ and Julian Wright, "Marketplace or reseller?," *Management Science*, 2015, *61* (1), 184–203.

_ , Bruno Jullien, and Julian Wright, "Creating Platforms by Hosting Rivals," *Management Science*, 2015.

_ , Tat-How Teh, and Julian Wright, "Should Platforms Be Allowed to Sell on Their Own Marketplaces?," *RAND Journal of Economics*, 2022, *53* (2), 297–327.

Jiang, Baojun, Kinshuk Jerath, and Kannan Srinivasan, "Firm Strategies in the "Mid Tail" of Platform-Based Retailing," *Marketing Science*, 2011, *30* (5), 757–775.

Kamepalli, Sai Krishna, Raghuram G. Rajan, and Luigi Zingales, "Kill Zone," Working Paper 27146, National Bureau of Economic Research May 2020.

Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E Hinton, "ImageNet classification with deep convolutional neural networks," *Communications of the ACM*, 2017, *60* (6), 84–90.

Lam, H Tai, "Platform search design and market power," *Job Market Paper, Northwestern University*, 2021.

Lee, Kwok Hao and Leon Musolff, "Entry into Two-Sided Markets Shaped by Platform-Guided Search," *Working Paper*, 2021.

Madsen, Erik and Nikhil Vellodi, "Insider Imitation," *Working Paper*, 2022.

Magnusson, Evan, "Unboxing the causal effect of ratings on product demand: Evidence from wayfair. com," *The Journal of Industrial Economics*, 2022, *70* (3), 525–564.

Raval, Devesh, "Steering in one click: Platform self-preferencing in the amazon buy box," *Unpublished manuscript*, 2022, *18*.

Rochet, Jean-Charles and Jean Tirole, "Platform competition in two-sided markets," *Journal of the european economic association*, 2003, *1* (4), 990–1029.

Sun, Liyang and Sarah Abraham, "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects," *Journal of Econometrics*, 2021, *225* (2), 175–199.

U.S. House Committee on the Judiciary, "Investigation of Competition in Digital Markets: Majority Staff Report and Recommendations," Staff Report, Subcommittee on Antitrust, Commercial and Administrative Law, Washington, DC 2020.

Yang, Yinfei, Daniel Cer, Amin Ahmad, Mandy Guo, Jax Law, Noah Constant, Gustavo Hernandez Abrego, Steve Yuan, Chris Tar, Yun-Hsuan Sung et al., "Multilingual universal sentence encoder for semantic retrieval," in "Proceedings of the 58th annual meeting of the Association for Computational Linguistics: system demonstrations" 2020, pp. 87–94.

Zhou, Bobby and Tianxin Zou, "Competing for Recommendations: The Strategic Impact of Personalized Product Recommendations in Online Marketplaces," *Marketing Science*, 2023, *42* (2), 360–376.

Appendix

A Additional Details on Data Construction

A.1 Private-Label Brand Identification

We identify Amazon's private-label products using a comprehensive list of Amazon-owned brands compiled from public reporting, Amazon's own disclosures, and manual verification. Table A1 lists the biggest private-label brands in our data.

Table A1: Biggest Amazon private-label brands in our sample

Brand Name	# Products	Primary Categories
Amazon Collection	18,722	Jewelry, accessories
Amazon Basics	16,627	Electronics accessories, office, home
Stone & Beam	3,289	Furniture, home décor
Rivet	3,265	Furniture, lighting
Amazon Essentials	2,671	Apparel, basics
Pinzon	1,845	Bedding, bath
Solimo	811	Grocery, personal care, household
Ravenna Home	783	Furniture
Denali	700	Outdoor, home
Meraki	539	Personal care

B Additional Results on Product Similarity

B.1 Visual Similarity Examples

Figures A1–A4 display additional side-by-side comparisons of Amazon private-label products and their most similar third-party predecessors.

Figure A1: Office chairs: Amazon private-label product and most similar third-party predecessor



B.2 Similarity Scores over Time

Figure A5 plots the average maximum cosine similarity for private-label products introduced in each year of our sample. The figure reveals a declining trend: more recently introduced private-label products tend to have lower maximum similarity scores relative to pre-existing third-party products.

B.3 Robustness to Alternative Similarity Thresholds

Figure A2: Dinnerware sets: Amazon private-label product and most similar third-party predecessor



Table A2: Prevalence of similarity under alternative thresholds

Similarity Threshold	# PL Products with Match	Percent of Total
≥ 0.80	1,521	90.9%
≥ 0.85	1,408	84.2%
≥ 0.90	1,222	73.1%
≥ 0.95	687	41.1%

Notes: Total number of private-label products in the 14 focal categories is 1,673 ($N_{3P} \approx 5.5$ million). U.S. marketplace.

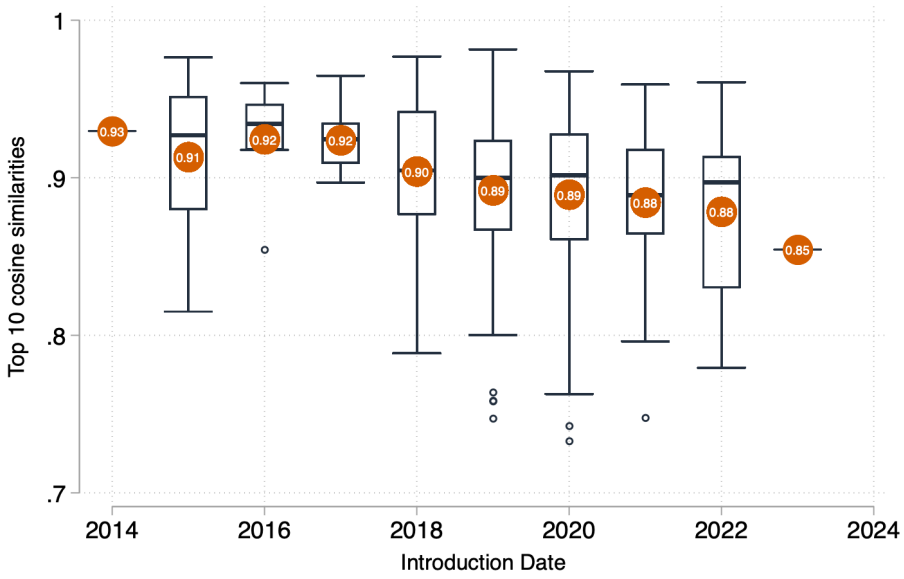
Figure A3: End tables: Amazon private-label product and most similar third-party predecessor



Figure A4: Suitcases: Amazon private-label product and most similar third-party predecessor



Figure A5: Average maximum visual cosine similarity between Amazon private-label products and pre-existing third-party products, by year of private-label introduction



C Category-Level Descriptive Statistics

Amazon does not enter categories at random. Table A3 compares pre-treatment characteristics of treated and never-treated categories in the U.S. marketplace. Treated categories are substantially larger: they have more sellers (6.1 vs. 1.9 log sellers), higher maximum prices (7.1 vs. 4.6 log max price), and a larger review stock (8.7 vs. 4.9 log reviews). They also have slightly fewer images per listing and lower star ratings.

Table A3: Balance: treated vs. never-treated categories (U.S. marketplace)

	Treated	Never-treated	Difference	p-value
Log(Total Sellers)	6.136	1.949	4.187***	0.000
Log(Max Price)	7.146	4.649	2.497***	0.000
Log(Avg Price)	4.273	3.829	0.444***	0.000
Avg Star Rating	4.031	4.182	-0.151***	0.000
Log(Total Reviews)	8.704	4.889	3.815***	0.000
Avg Images	3.150	3.448	-0.298***	0.000
Avg Videos	0.044	0.109	-0.066***	0.000

Notes: Pre-treatment means computed over the 12 months before first private-label entry (treated) or the first 12 months of data (never-treated). U.S. marketplace. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Amazon is significantly more likely to enter categories with more sellers, lower star ratings, and fewer images and videos (which could be a proxy of sellers' investment to the product page). This selection is *not* a threat to our identification: the cross-country design compares the same category across countries at the same time, absorbing all category-level characteristics (including those that predict entry). It would be a threat only if Amazon's entry *timing* across countries within a category were correlated with country-specific trends in outcomes - a scenario addressed by the flat pre-trends.

D Additional Results on Seller Outcomes

D.1 Third-Party-to-Third-Party Imitation: Sales Rank Comparison

Table A4 reports sales rank comparisons for third-party-to-third-party imitation.

Table A4: Third-party imitation: sales rank comparisons

Taxonomy Level	Mean Imitated	Mean Non-Imitated	Difference	p-value	
Category 1	496,649	995,456	−498,807	0.000	***
Category 2	7,342	14,134	−6,792	0.003	***
Category 3	23,428	32,910	−9,482	0.600	

Notes: Imitated products are defined as third-party products with visual cosine similarity ≥ 0.9 relative to at least one subsequently introduced third-party product (excluding Amazon private labels). Lower sales rank indicates higher sales. U.S. marketplace, 14 focal categories. $N \approx 5.5$ million third-party products. ** $p < 0.05$, *** $p < 0.01$.

D.2 Imitation Timing by Seller Size

Figure A6 disaggregates the distribution of imitation lags by seller revenue quintile. Smaller sellers experience substantially longer lags between their product's introduction and the entry of a similar Amazon private label.

Figure A6: Distribution of imitation lag by seller revenue quintile

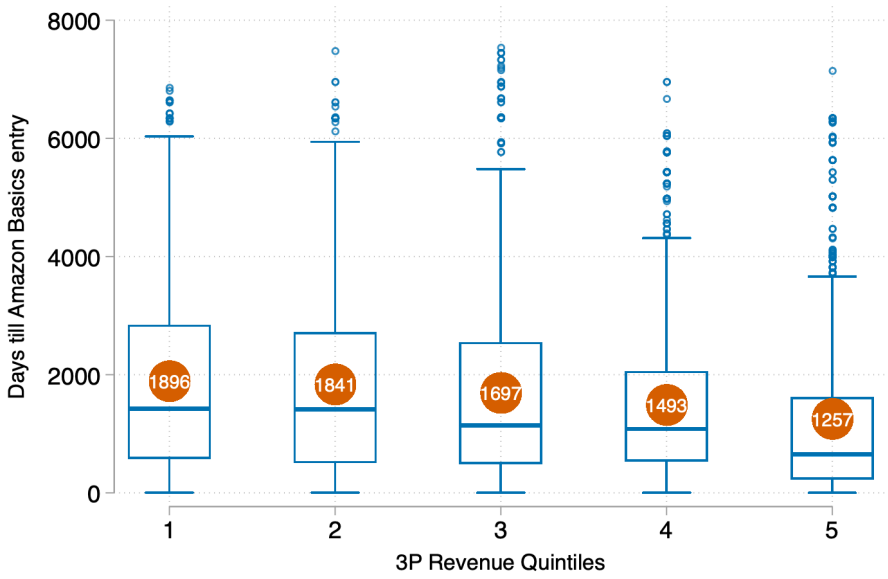
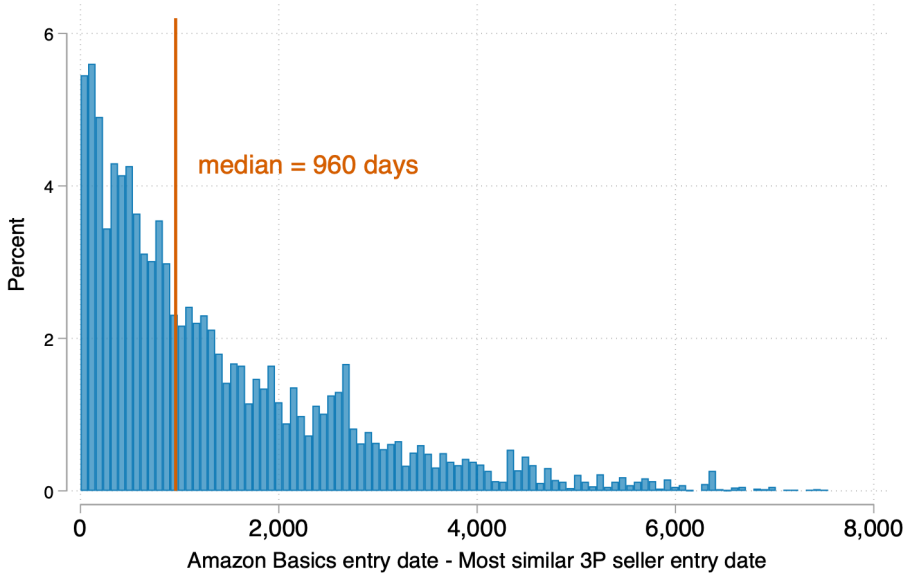


Figure A7: Distribution of days between third-party product introduction and similar private-label entry



E Creating embeddings

E.1 Image embeddings

We use convolutional neural networks (CNNs) to embed each product image as a 512×1 vector in a lower-dimensional space. The model is trained on all available images from Amazon. Since each product has multiple images (Figure A8), we construct labels without manual annotation: pairs from the same product are labeled one, and pairs from different products are labeled zero. The trained CNN places images of the same product close together and images of different products far apart.

Figure A8: Different images of the same product



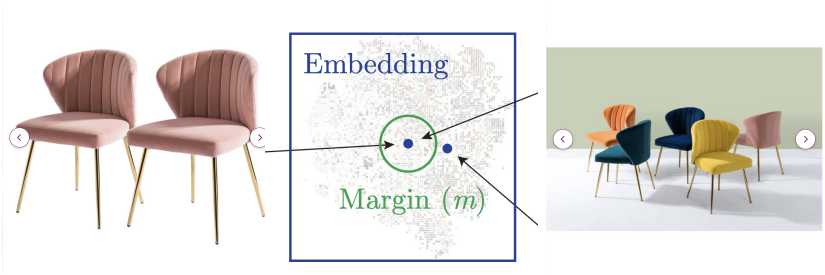
Figure A9 illustrates the goal: learn a representation in which the image in Panel (A) and the matched image of the same product in Panel (B) map to nearby points, while the image in Panel (C), a completely different product, is separated from them by at least a margin m.

Formally, a convolutional neural network is a function f_θ that maps each image I into an embedding x given parameters θ: $f_\theta : I \to x$ (see Hadsell et al. (2006) and Bell and Bala (2015) for details). The goal is to solve for θ so that the resulting embeddings place similar images nearby and dissimilar images far apart.

Concretely, suppose we have three images I_a, I_b, I_c. Images I_a and I_b are two views of the same object (Panels (A) and (B) of Figure A9), while I_b and I_c are different objects (Panels (B) and (C)). The CNN maps them to embeddings x_a, x_b, x_c. With well-trained parameters,

Figure A9: Idea behind the embeddings

(a) Image I_a *(b)* Image I_b



(c) Image I_c



Notes. The figure illustrates the idea behind image embeddings. The products in Panels (a) and (b) are images of the same item, and the model's task is to place them close together in the embedding space. The product in Panel (c) is a different item and should be more distant from (a) and (b).

the positive pair (x_a, x_b) should be close together and the negative pair (x_b, x_c) far apart.

The objective for this task is the contrastive loss of Hadsell et al. (2006). We minimize

$$L(\theta) = \sum_{x_a, x_b} \underbrace{L_b(x_a, x_b)}_{\substack{\textit{Penalty if similar images} \\ \textit{are far away}}} + \sum_{x_b, x_c} \underbrace{L_c(x_b, x_c)}_{\substack{\textit{Penalty if dissimilar images} \\ \textit{are nearby}}} \tag{5}$$

where $L_b(x_a, x_b) = ||x_a - x_b||_2^2$ and $L_c(x_b, x_c) = \max\{0,\, m^2 - ||x_b - x_c||_2^2\}$. The first term penalizes a positive pair whose embeddings are too far apart; the second penalizes a negative pair whose embeddings are closer than the margin m.

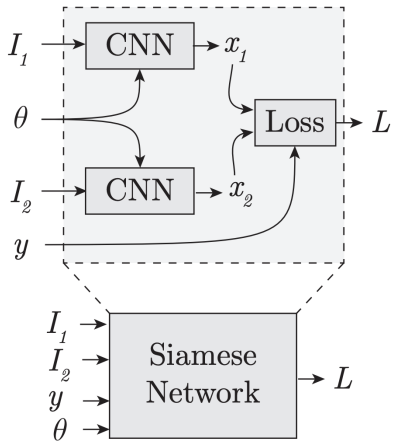
We minimize Equation 5 over θ using stochastic gradient descent with momentum (Krizhevsky et al., 2017; Bottou et al., 2018):

$$v^{(t+1)} \leftarrow \mu \cdot v^{(t)} - \alpha \cdot \frac{\partial L}{\partial \theta}(\theta^{(t)}) \tag{6}$$

$$\theta^{(t+1)} \leftarrow \theta^{(t)} + v^{(t+1)} \tag{7}$$

where v is the momentum sequence, $\mu \in [0, 1)$ is the momentum, and $\alpha \in [0, \infty)$ is the learning rate. To compute L and $\partial L / \partial \theta$ efficiently, we follow Hadsell et al. (2006) and use a siamese network: two copies of the CNN that share the parameters θ. The network takes two images I_1, I_2, the parameters θ, and an indicator for whether the images are a positive pair ($y = 1$) or a negative pair ($y = 0$), and returns the loss L (Figure A10).

Figure A10: Siamese network



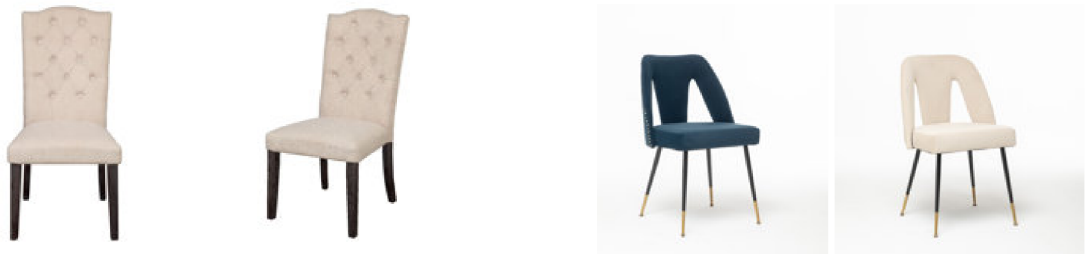
Notes. This figure illustrates the architecture of the deep neural network that we use to create image embeddings. Source: Bell and Bala (2015).

Different images of the same product form a positive pair, and the model should map them to similar embeddings under a distance metric such as cosine similarity. For negative examples, we sample, for each product, other products from both the same category and other categories.

The model output is a 512×1 vector for each product. We then compute cosine similarities between products' vectors to quantify how alike they are. Figures A11 and A12 show examples of products with high cosine similarity (> 0.9) and low cosine similarity (< 0.1), computed from the model output.

Figure A11: Examples of products with high cosine similarity

(a) Cosine similarity: 0.9294 *(b)* Cosine similarity: 1.0000



Notes. This figure shows examples of products with high cosine similarity. Products with cosine similarity above 0.9 are typically visually very similar.

Figure A12: Examples of products with low cosine similarity

(a) Cosine similarity: -0.0001 *(b)* Cosine similarity: 0.0074



Notes. This figure shows examples of products with low cosine similarity. Products with cosine similarity below 0.1 are typically visually very different.

E.2 Textual Embeddings

We create textual embeddings of the product details, e.g., specifications, from the product page and the product reviews. We use the Universal Sentence Encoder developed by Google (Cer et al., 2018; Yang et al., 2020). The encoder maps raw English text into a 512-dimensional vector space. Unlike word-level embeddings (e.g., Word2Vec, GloVe), sentence-level encoders model the meaning of word *sequences* and are trained on multiple tasks simultaneously, resulting in more robust semantic representations.