Platform Brands Rising: The Competitive and Perceptual Consequences of Amazon Brand Entry

Abstract

My dissertation investigates the competitive and perceptual consequences of platform brand entry on digital marketplaces, using Amazon.com as the empirical setting. It addresses a critical gap in understanding the impact of platform brands, which, unlike traditional private labels, have unique advantages from their control over data and algorithms. The central problem is to quantify how these platform brands influence market prices, consumer engagement, and perceptions of competing products.

The empirical design contains three components. First, to define a relevant market, I propose a novel market definition approach based on rank-weighted search co-occurrence, which clusters products that frequently appear together in consumer search results. Second, I use the staggered Synthetic Difference-in-Differences (SynthDiD) model for a market-level analysis to estimate the overall effects of platform brand entry. This method addresses the staggered timing of brand launches and potential selection bias. Finally, I implemented a panel-adapted Double Machine Learning (DML) framework to uncover heterogeneous effects at the product level. This framework uses multi-modal product embeddings derived from both images and text—to measure product similarity and analyze how the treatment effects vary across different seller types and levels of similarity to the entering Amazon brands.

The preliminary findings reveal a paradoxical divergence between competitive gains and perceptual costs. The entry of Amazon brands, particularly Amazon Basics, significantly reduces market prices and increases consumer engagement, suggesting a pro-competitive ”market disciplining effect”. However, despite these gains, consumer perceptions of both product value and quality decline. This is potentially because the low-cost platform brands reset consumer expectations, making competing products seem overpriced or under-performing in comparison. The analysis also shows that these effects are not uniform. Third-party sellers and products that are moderately similar to the Amazon brand face the most pronounced negative impacts. In contrast, products with high similarity to the Amazon brand experience a price increase, suggesting they may benefit from a form of platform favoritism or protection. These results provide valuable insights for sellers, platforms, and policymakers by highlighting how platform brands can unevenly reshape competition and consumer perceptions.

Keywords: Platform brands, Platform strategy, Private label strategy, Entry effect, Digital competition

1 Introduction

Private label (or store brand) is an important element in brick-and-mortar retail in terms of product assortment and market power (Kumar and Steenkamp, 2007). Over the past decades, national brands have learned to coexist with private labels by adapting their positioning, improving quality differentiation, and defending their value (Chintagunta et al., 2002; Ailawadi et al., 2008; Pauwels and Srinivasan, 2004, 2009). In the digital era, the rise of online retail platforms such as Amazon.com is creating both new opportunities and new challenges for private labels. For reference ease, I call the private label products introduced by digital retail platforms as "platform brand". Pioneering the practice, Amazon launched its private label program in 2005 (Morrison, 2021). By the end of 2024, Amazon offered over 23,000 products across at least 406 Amazon platform brands and received \$1 billion in sales from these products (Herrera, 2022). Globally, JD.com in China launched its own platform brand, Jingzao (Technode, 2019). Flipkart in India has introduced a variety of platform brands across multiple categories (Sengupta, 2020).

Platform brands differ fundamentally from traditional private labels. In brick-and-mortar retail, private labels compete for limited shelf space against national brands, and consumers choose from a fixed set of in-store displays. In digital marketplaces, by contrast, product space is virtually unlimited, search engines dynamically shape what consumers see, and user reviews often substitute for physical displays. Competition also shifts: instead of facing primarily large national brands, platform brands contend with a fragmented base of independent sellers. Critically, as marketplace owners, platforms like Amazon not only control visibility and recommendations but also hold vast amounts of user-level data unavailable to third-party sellers. This dual role, as both referee and player, creates the potential for platforms to tilt the playing field, raising concerns about distorted competition and consumer harm[1].

Despite the growing popularity of platform brands and the increasing regulatory scrutiny about them, there is little empirical evidence on how platform brand expansion actually plays out in the marketplace. Do platforms introduce their own brands to fill unmet gaps and broaden choice in underserved markets, or do they use control over algorithms and data to target high-traffic, high-margin product spaces? For competition: when a platform brand enters, do market prices fall and dispersion narrow due to a low-cost entrant, or do prices rise and dispersion widen due to higher visibility costs and seller repositioning? How do consumers' evaluations of competing products change? Which products and sellers are most affected? In my dissertation, I quantify these competitive and perceptual consequences of platform-brand entry in search-based marketplaces.

[1] For example: *The Ending Platform Monopolies Act* introduced in the U.S. Congress in 2021 proposes to prohibit Amazon from selling private labels. Similar concerns were raised by the *Digital Markets Acts (DMA)* in Europe and the pending *American Innovation and Choice Online Act (AICOA)*.

I address these questions in the context of Amazon.com. I compile a multi-source panel of 108,846 unique products from 1,158 search result pages, tracking each product's daily price, rating, and sales rank from launch through December 2024. Using deep learning and large language models (LLMs), I convert unstructured product descriptions and images into comparable measures of product similarity and extract aspect-level consumer perceptions from review text.

My empirical strategy contains three components. First, to reflect both visibility and competitive proximity in the search space, I develop a novel market definition approach that clusters products based on search co-occurrence patterns. This approach allows us to define behaviorally grounded markets and designate the relevant sets of products that are plausibly affected by the entry of Amazon brands. Second, I employ the staggered synthetic difference-in-differences (SynthDID) strategy (Arkhangelsky et al., 2021) to estimate market-level treatment effects on price, review metrics, and consumer perceptions. Finally, to uncover heterogeneous effects, I implement a panel-adapted Double Machine Learning (DML) framework (Chernozhukov et al., 2018; Clarke and Polselli, 2025), which, through a product-level lens, traces where platform brand entry has the most pronounced impact.

The results reveal a paradoxical divergence of platform brand entry: while the entry of Amazon brands drives market prices down and boosts consumer engagements, suggesting a market disciplining effect (Desai et al., 2020; Gilbert and Shelanski, 2023), it erodes consumer perceptions of product value and quality. This indicates that the entry of platform brands impacts the market not only through prices but also by subtly altering how consumers evaluate and trust products. Furthermore, I find that third-party sellers and products that are moderately similar to platform brands would be impacted to a greater extent. These results highlight a deeper form of distortion in platform markets where market efficiency may improve, but consumer expectations can outpace what the sellers can deliver. Ongoing analysis examines the mechanisms behind these shifts.

As one of the first empirical assessments of the competitive and perceptual consequences of platform brand entry, this paper contributes to the emerging literature at the intersection of platform strategy and strategic competition. While prior work focused on offline retail or price-based outcomes (Raju et al., 1995; Chintagunta et al., 2002; Pauwels and Srinivasan, 2004; Ailawadi et al., 2008), I show that the entry of platform brands not only creates market disciplining forces but also reshapes how consumers evaluate product value and product quality. These findings shed light on the self-preferencing debate, which has focused on algorithm bias (Zou and Zhou, 2025) or listing prominence (Long and Amaldoss, 2024), by demonstrating demand-side distortions in consumer perception.

The findings generate important implications for sellers, platforms, and policymakers navigating the evolving dynamics of platform-dominated marketplaces. For third-party sellers, platform brand entry, particularly in the value tier, creates intense price pressure and elevates consumer scrutiny. For platform owners, the challenge lies in balancing market efficiency gains with the long-term sustainability of the ecosystem; while platform brands may drive lower prices and higher ratings, they also erode consumer perception of competing products and risk displacing seller diversity. Finally, for regulators, our heterogeneity analysis reveals that while platform brand entry can lower prices and improve ratings, seemingly benefiting consumers, it may also undermine consumer perceptions of quality, leading to a "perception paradox" that calls for regulatory attention to information transparency and consumer welfare beyond price alone.

The remainder of this paper is organized as follows. Section 2 provides the industry background and the empirical setting for the research. Section 3 presents the data and sample construction. Section 4 discusses the identification challenges and outlines my empirical strategy. Section 5 presents the findings. I discuss the managerial implications and future research opportunities in Section 6.

2 Research Context

As the world's largest online retail platform, Amazon operates a hybrid business model: it not only facilitates transactions for millions of third-party sellers but also acts as a first-party retailer selling products directly to consumers. Moreover, Amazon develops and sells its platform brands, such as Amazon Basics, starting in 2005 (Morrison, 2021). Figure 1 illustrates the structure of this integrated marketplace ecosystem.

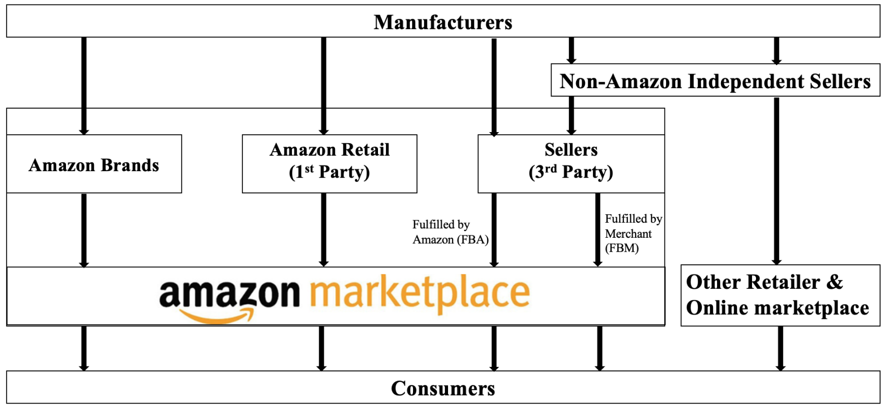
The Amazon platform brands remain part of Amazon's first-party operation. They are owned, branded, and marketed by Amazon, deeply integrated with the platform. Although these platform brands account for only about 1% of the total revenue by all sellers on Amazon.com (MobiLoud, 2024; Momentum Commerce, 2024), a smaller share compared to traditional retail chains[2], their influence is disproportionately large due to their algorithmic visibility and strategic positioning (Deng et al., 2023; Tang et al., 2024).

Intuitively, Amazon's private label strategy could resemble the traditional retail chain (Dubé, 2024), offering low-cost alternatives that enhance margins and intensify price competition (Ailawadi et al., 2001; Chintagunta et al., 2002; Pauwels and Srinivasan, 2004). However, Amazon's dual role as a marketplace operator and a seller complicates this comparison. It designs data-driven brands and controls discovery via search, promotions, and badges (Korganbekova and Korganbekov, 2024; Angwin et al., 2016; Long and Amaldoss, 2024; Momentum Commerce, 2023), potentially advantaging its own offers (Jiang et al., 2011; Wen and Zhu, 2019). At the same time, third-party sellers remain central to platform economics (about 62% of units; substantial fee revenue)[3].

[2] Mainstream retailer such as Walmart and Target attributes 23% and 15% of sales to their private label products

[3] Third-party sellers accounted for 62% of units sold in Q4 2024; roughly one-quarter of revenue derives from seller fees: Marketplace Pulse.

Figure 1: **Structure of Amazon's Hybrid Marketplace.**



This dual incentive raises the empirical questions about the net effect on market price, consumer evaluations, and incidence across sellers.

I study these dynamics in the Home & Kitchen category on Amazon.com. This category provides an ideal empirical setting: First, this is one of the largest and most competitive categories on Amazon.com[4]. Second, this category features a diverse mix of Amazon brands, from value-oriented brands like Amazon Basics to more design-oriented, higher-end brands such as Stone & Beam and Rivet. Third, consumer evaluations of home products span both vertical and horizontal dimensions: vertical quality attributes, such as performance, durability, and functionality, coexist with horizontal preferences related to aesthetics, style, and design. As a result, the Home & Kitchen category provides a rich setting to evaluate how Amazon brand entry impacts not only price and competition, but also consumer perception and product positioning.

3 Data and Sample Construction

I constructed a multi-source panel dataset from three primary sources. First, I collect search queries on Amazon.com to identify all Amazon platform brands and the competing products. For each product, I scrape product-level metadata and consumer search data. Second, to examine the impact of Amazon brand entry on prices, I obtain the historical price data for each product using the API from Keepa.com. Third, to examine the impact of the Amazon brand entry on consumer perceptions, I use Large Language Models to extract the aspect-based sentiment on each consumer review. I validate the results using a manually labeled ground truth dataset. For the heterogeneity analysis, I measure product similarity leveraging the state-of-the-art multi-modal representation

[4]In the U.S. market alone, this category generates approximately \$103.3 billion in sales in 2024 and exhibits a year-over-year growth rate of roughly 20%.See Momentum Commerce (2024), "Amazon's Home & Kitchen Category Forecasted to Grow 21.8% YoY in 2024," available at: https://www.momentumcommerce.com/amazons-home-kitchen-category-forecasted-to-grow-21-8-yoy-in-2024/, accessed on July 18, 2024.

learning. The remainder of this section provides a detailed discussion of each component.

3.1 Search Queries and Product Listings

I simulate the data-generating process of a typical consumer's shopping journey by collecting search results and product metadata from Amazon.com. Using a curated list of 1,158 high-traffic search queries, based on the most searched keywords on SEMrush.com in 2021, I scrape the top 300 search results, yielding 324,520 product-query pairs. On average, each product appears in approximately three distinct search queries, resulting in a total of 108,846 unique products. For each product, I scrape detailed information including product name, images, price, delivery terms, and the "Date First Available." I also collect all historical customer reviews associated with these products, including timestamps, star ratings, and review text. Table 1 presents a summary statistic of the search and product data. As it shows, the products belong to seven broad categories in the home and furniture market. The average price is approximately $\$400$. The number of unique product items in each category ranges from approximately $5,000$ (in Shoe Rack category) to $23,000$ (in Rug category). The proportion of these products being sold by third-party sellers (vis-a-vis first-party by Amazon) is 75%.

Table 1: Summary Statistics by Product Category

Category	NumKeyword	NumSearch	NumUniqueProd	AvgPrice	AvgRating	AvgRatingCount	ThirdPartyPerc
Chair	187	52850	21703	515.58	2.27	549.79	73.93
Desk	177	48480	19042	512.51	2.22	732.07	74.07
Dresser	190	55649	13617	525.48	1.92	337.86	74.44
LED Lights	148	39614	15222	125.13	2.15	990.23	82.61
Rug	182	49735	23498	223.60	1.80	1115.14	71.25
Shoe Rack	123	33439	5045	312.83	1.91	573.51	79.51
TV Stand	151	44753	10722	596.49	1.85	548.09	72.10
Total	1158	324520	108846	398.67	2.04	736.88	74.72

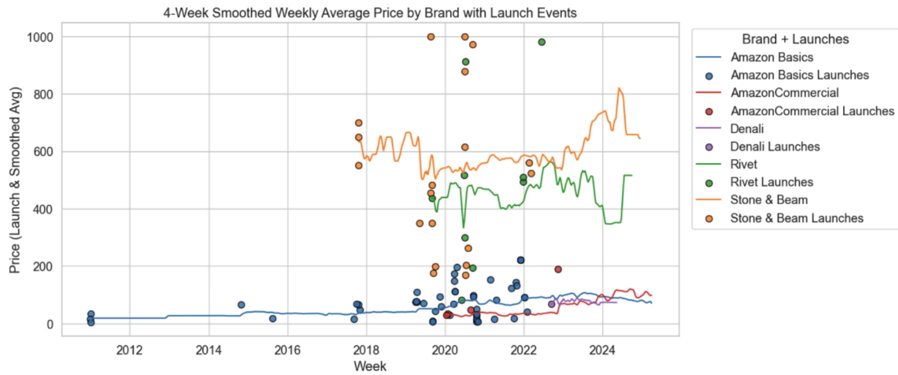
Notes: This table reports summary statistics for the seven product categories in our constructed Amazon Home & Kitchen sample. *NumKeyword* is the number of distinct search keywords collected for each category. *NumSearch* is the total number of keyword-ASIN pairs observed in search results. *NumUniqueProd* indicates the number of unique products. *AvgPrice*, *AvgRating*, and *AvgRatingCount* are averaged at the product level. *ThirdPartyPerc* refers to the percentage of products offered by third-party sellers.

These products include a total of 108 Amazon brands, which serve as the focal treatment in our analysis. Figure 2 presents the timeline of entry and the average price of these brands, aggregated by brand and time. As it shows, the majority of Amazon brands were launched between 2017 and 2022, with most priced below $\$200$. Several premium brands, such as Stone & Beam and Rivet, exhibit significantly higher average prices at approximately $\$600$ and $\$400$, respectively.

3.2 Daily Price Data from Keepa.com

I collect historical daily price data using the Keepa.com API, which is a widely used price-tracking tool that archives real-time changes in Amazon product listings (Chen and Tsai, 2023; Raval, 2023; Feng et al., 2024; Lee

Figure 2: The Entry of Amazon Brands by Time and Brand



and Musolff, 2025). Keepa tracks prices at the ASIN level across multiple seller types, including Amazon retail (i.e., first-party), fulfilled-by-Amazon (FBA) third-party sellers, and non-FBA third-party sellers. For each product, I retrieve price, Buy Box winner, seller composition, and availability information at a daily interval. This allows us to construct a high-frequency product by time panel and to more accurately detect potential changes in pricing across seller types before and after the entry of Amazon brands.

3.3 Aspect-based Consumer Perception

To gain a deeper understanding of consumer opinion across different aspects, such as price value or product satisfaction, I implement a fine-grained, aspect-based sentiment analysis (Liu and Zhang, 2012; Puranam et al., 2021). Specifically, I focus on consumer perception of price and product quality. Since product quality is widely recognized as a multidimensional construct, I define five quality dimensions based on the existing literature (Garvin, 1984; Puranam et al., 2021): performance, aesthetic, durability, assembly, and product description fit.

To measure consumer perceptions, I leverage large language models (LLMs) (Gilardi et al., 2023). Specifically, I prompt ChatGPT to evaluate each product review along the predefined dimensions. For each aspect, a ChatGPT 3.5-turbo model is instructed to identify relevant content, if present, and assign a sentiment score on a scale from -10 to 10, where negative values indicate unfavorable sentiment, positive values indicate favorable sentiment, and a score of 0 denotes that the aspect is not mentioned. The prompt includes aspect definitions, annotated examples, and rating heuristics (chain of thought) to ensure alignment and scoring consistency. The resulting aspect-level sentiment scores are then aggregated at the product by time level to construct a panel of consumer perception variables that capture the temporal dynamics of sentiment.

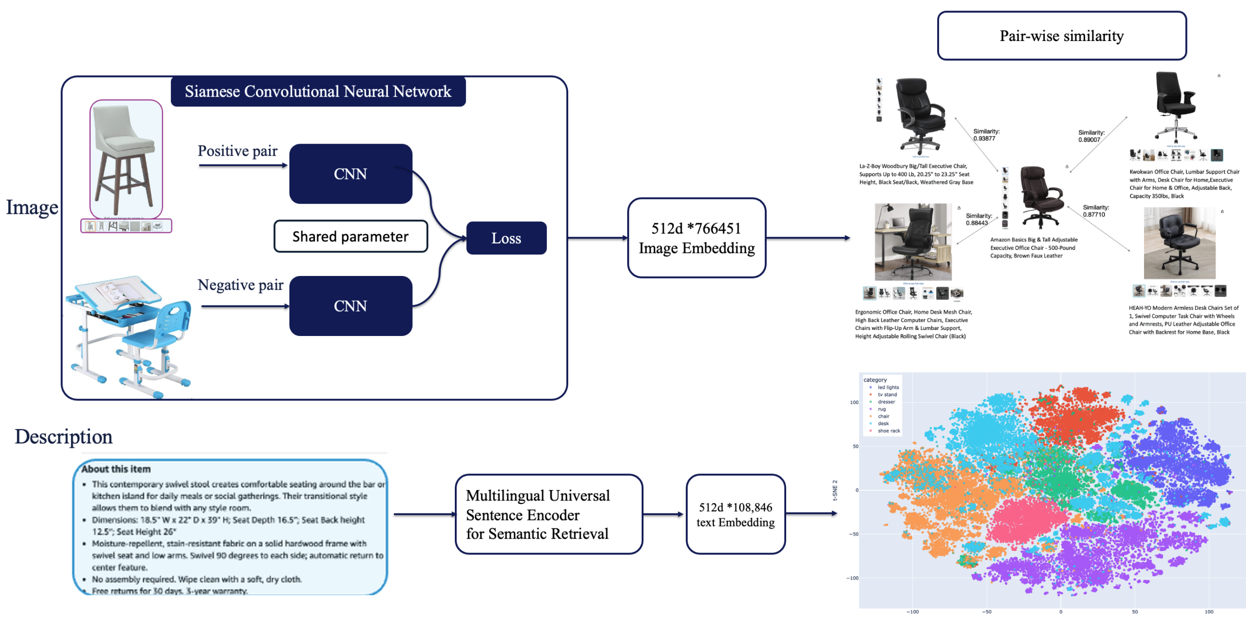
To validate the accuracy of the model-generated sentiment scores, I recruited and trained a group of marketing graduate students to manually label a random sample of 4,000 reviews. The human-labeled data serve as a ground truth benchmark for evaluating the consistency and reliability of our automated annotations. Across aspects, the

ChatGPT-based sentiment analysis achieves accuracy ranging from 0.756 to 0.942.

3.4 Multi-modal Product Embedding to Measure Similarity

Product listings are inherently multimodal, combining text that conveys functionality and branding with images that capture design and aesthetics(Liu et al., 2020). To accurately capture product-level similarity, I leverage deep learning techniques to extract high-dimensional representations of both product images and textual descriptions from Amazon product detail pages, which contain rich unstructured information that reflects a product's features, functional attributes, aesthetic design, and brand positioning. This similarity measure plays a central role in our empirical strategy, as it (i) helps validate market clustering quality, (ii) captures competitive proximity in the heterogeneity analysis, and (iii) controls for confounders in the causal estimation. I illustrate our approach in Figure 3.

Figure 3: Framework of Measuring Product similarity via Multi-modal Embedding

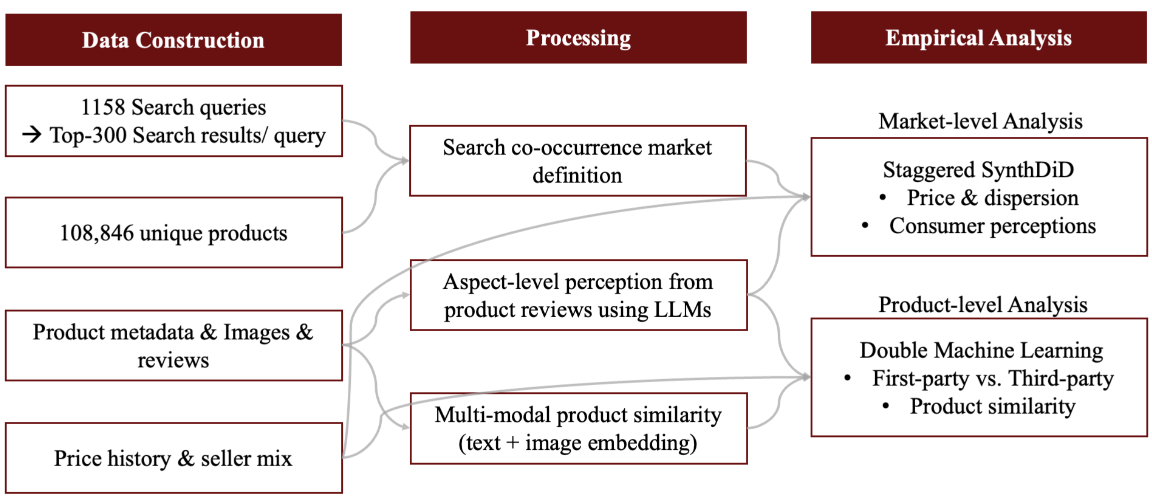


For the image encoder, I collect 766,451 product images for the 108,846 products on Amazon.com. On average, each product has 7 images. I use a Siamese Convolutional Neural Network (CNN) to generate image embeddings, mapping each image to a 512-dimensional vector in latent space. The network is trained with contrastive loss (Koch et al., 2015; Bell and Bala, 2015) to minimize distances between visually similar items and maximize distances between dissimilar ones. For the text encoder, I use a pretrained language model to convert product descriptions into 512-dimensional semantic embeddings (Yang et al., 2020). I then compute the pairwise cosine similarity between all products based on both image and text representations. These similarity measures enable us to quantify the degree of similarity between any two products in terms of appearance and messaging.

4 Empirical Strategy

My study estimates the causal effects of Amazon brand entry on market price, engagement, and consumer perceptions. The empirical strategy involves two important elements. First, I construct the proper search-based market space that facilitates the causal inference. Second, I address several identification challenges inherent in this setting by employing a combination of causal inference approaches. Following the research design road map presented in Figure 4, I discuss each part in detail below.

Figure 4: Research Design Road Map



4.1 Search Co-occurrence Market Construction

Unlike physical retail environments, where products from the same or complementary categories are placed in adjacent aisles or shelf locations, online marketplaces like Amazon are organized around search queries. As a result, products that are categorically similar may not appear together under the same search terms, while products from different categories may be displayed together in response to broad or ambiguous queries. This fragmented yet overlapping exposure creates ambiguity in defining market boundaries, which can introduce spillover effects between treated and control units within the same product category, potentially violating the stable unit treatment value assumption (SUTVA) (Goldfarb et al., 2022). To address these challenges, I propose a new method for defining product markets based on rank-weighted search co-occurrence patterns. The underlying rationale is that products frequently appearing together in search results, particularly in high-visibility positions, compete more directly for consumer attention and purchase decisions.

Figure 5: Overview of the Proposed Market Definition Method

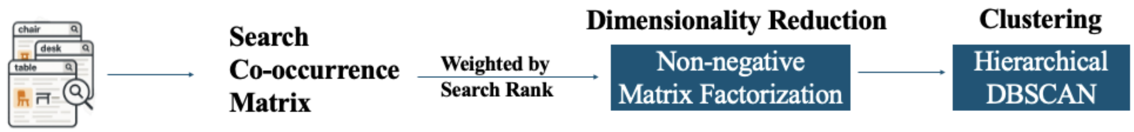


Figure 5 illustrates the overall structure and the three key components of my proposed method. First, following Decarolis and Rovigatti (2021), I construct a product-query co-occurrence matrix from search query-product pairs, where $rank_{ij}$ is the rank position if the product i appears under search results for query j, and 0 otherwise. I apply rank-weighted transformation to prioritize co-occurrences in prominent search positions, defined in Equation 1:

$$w_{ij} = \exp\left(-\frac{\text{rank}_{ij}}{\lambda}\right) \cdot \mathbb{1}\{\text{prod. } i \text{ in query } j\} \tag{1}$$

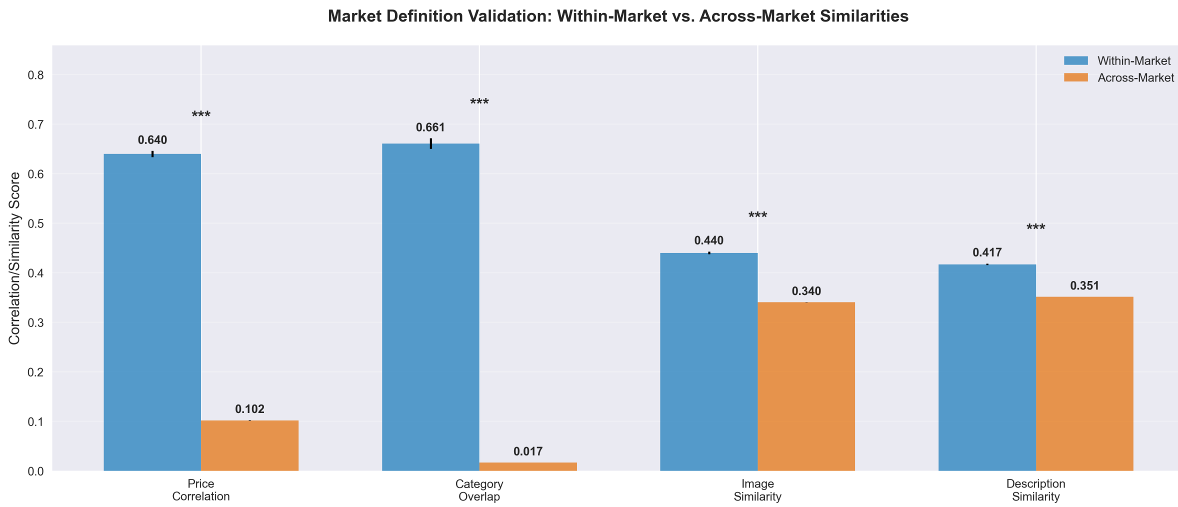
This formation assigns a weight w_{ij} to each product-query pair (i, j), where $rank_{ij}$ denotes the rank position of product i in the search results for query j and λ controls the decay rate of rank importance. In my empirical implementation, we set $\lambda = 30$. The exponential term prioritizes products in top-ranked positions, while the indicator function $\mathbb{1}\{\text{prod. } i \text{ in query } j\}$ ensures only observed product–query co-occurrences receive non-zero weight. This weighting procedure groups together products that frequently co-appear in prominent positions (i.e., higher-ranked), capturing both underlying algorithmic similarity and the critical aspect of consumer visibility in search behavior (Kumar et al., 2020).

Second, to reduce the dimensionality of the weighted co-occurrence matrix while also maintaining interpretability of the resulting market structure, I apply Non-negative Matrix Factorization (NMF) (Cichocki et al., 2009). NMF decomposes the original matrix into two lower-dimensional, non-negative matrices that capture latent product–keyword associations and market-level themes. I present the NMF visualization in Appendix **??**.

Third, I employ Hierarchical DBSCAN (Campello et al., 2013), a density-based clustering algorithm, on the reduced co-occurrence matrix. Compared to traditional methods like k-means, Hierarchical DBSCAN provides two notable advantages. First, it captures the hierarchical nature of product categorization. For instance, an "ergonomic office chair" naturally nests within the broader "office chair" category. Second, it allows for variable cluster sizes, accommodating the diversity of product markets, which may range from highly competitive and broad to narrow and niche.

Using this method, I group the 108,846 products into 790 markets. To validate my market construction, I compare key metrics within and across markets in Figure 6. These include intra-market price co-movement, which reflects shared demand and pricing pressures, overlaps in product category coverage, and product similarity based on textual descriptions and image embeddings. As Figure 6 shows, across all four metrics, intra-market performance significantly exceeds that of inter-market comparisons under my clustering.

Figure 6: Validation of Market Construction: Within-market vs Across-market Comparison



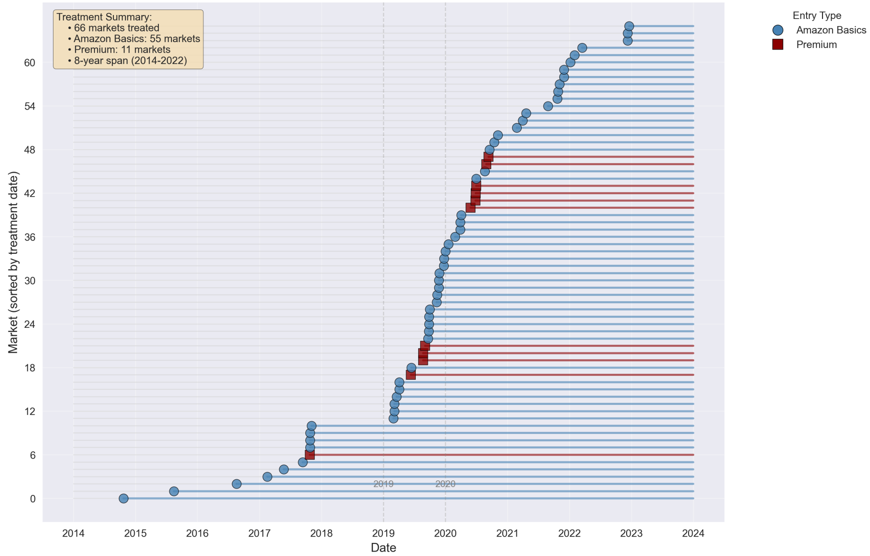
I designate the market as “treated” if it contains at least one Amazon brand product; in such cases, all products within that cluster are considered exposed to the treatment. The timing of treatment is determined by the earliest product first available date of the Amazon brand product that entered the market. This approach reflects the notion that even a single platform brand entry can reshape competitive dynamics and anchor consumer expectations within a category. Based on my market construction, 66 markets were treated between 2014 and 2022, of which 55 were entered by products under the Amazon Basics brand. The variation in treatment timing is illustrated in Figure 7.

4.2 Identification Challenges and Solutions

In this paper, I estimate the effects of Amazon brand entry at two levels. At the market level, I examine how the entry of Amazon brands influences overall market price, consumer engagement, and perception. At the product level, I examine heterogeneity in treatment effects based on seller type (first-party vs. third-party) and the similarity between competing products and Amazon brands. This dual-level approach is particularly important in platform settings, where strategies such as private label entry may directly impact a subset of products while generating broader spillovers or structural changes across the market.

The observational nature of the data presents several challenges for identifying the treatment effect. For market-level analysis, there are three challenges. First, I observe each product-market combination in only one state

Figure 7: Treatment Variation by time and brand type



(treated or untreated), making direct comparison difficult. Second, as shown in Figure 7, Amazon launches platform brands in a staggered manner, which may be correlated with evolving market characteristics. Third, its entry decisions are likely endogenous, driven by internal forecasts of profit potential, competitive dynamics, and market opportunities (Jiang et al., 2011). This selection bias threatens the identification as treated markets may differ systematically from the untreated ones along unobserved dimensions. To address these three issues, I adopt the staggered Synthetic Difference-in-Difference (SynthDiD) approach, as proposed by Arkhangelsky et al. (2021), to estimate the average treatment effect.

For the product-level analysis, products differ along a range of structured and unstructured attributes, such as textual descriptions, design features, and brand signals, that may be correlated with both treatment exposure and outcome variables. Furthermore, it may also be influenced by contemporaneous shocks, such as promotional events, platform algorithm changes, supply disruptions, or category-specific trends, that coincide with the entry of Amazon brands. Traditional empirical methods struggle to accommodate this high-dimensional heterogeneity while maintaining valid causal inference. To address these complexities and the panel nature of Amazon data, I employ the panel-adapted Double Machine Learning (DML) framework, as proposed by Chernozhukov et al. (2018) and extended in Clarke and Polselli (2025), which allows for flexible control of confounding factors in high-dimensional settings. In what follows, I discuss the identification strategy in detail.

4.2.1 Staggered Synthetic Difference-in-Difference

The SynthDiD has gained increasing traction in marketing research (Berman and Israeli, 2022; Lambrecht et al., 2024; Pachali and Datta, 2025; Qian and Xie, 2025) due to its flexibility in accommodating staggered treatments and heterogeneous trends across units. By integrating elements of synthetic control and conventional DiD, SynthDiD enhances counterfactual credibility and delivers more robust treatment effect estimates in contexts where the parallel trends assumption may be violated.

SynthDiD is well-suited for my research setting. It generates "clones" of the treated market by optimally weighting the never-treated markets that closely match the pre-entry outcome trajectories. This synthetic control market serves as a credible counterfactual for what would have happened in the absence of the platform brand entry. Additionally, by matching on pre-treatment dynamics, SynthDiD reduces the risk of treatment self-selection into inherently different markets and reduces the risk that observed effects are driven by unobserved shocks that coincide with the timing of Amazon brand entry. To further enhance the estimation effectiveness of synthDiD, I include exogenous time-varying covariates, such as the total number of products in the market and the proportion of first-party products. I do so by following the extended optimization procedures proposed by Kranz (2021) and Clarke et al. (2023).

Moreover, to address self-selection in staggered treatment timing, I adopt a block design. For each outcome variable, I apply the SynthDiD estimator separately to each market cohort defined by unique Amazon brand entry dates and aggregate the results by treatment length (Arkhangelsky et al., 2021; Callaway and Sant'Anna, 2021; Pachali and Datta, 2025). Specifically, conditional on every treatment date $a \in \mathcal{A}$, I construct a balanced panel of product markets, where $M_{co,a}$ represents never-treated control markets and $M_{tr,a}$ denotes markets first exposed to Amazon brands product at date a. Each balanced panel spans $T = T_{\text{pre}} + T_{\text{post}}$ periods. For each block a, I define $M_a = M_{co,a} \cup M_{tr,a}$ and estimate the average treatment effect using the synthetic difference-in-differences estimator proposed by Arkhangelsky et al. (2021), implemented via a weighted two-way fixed effects model:

$$(\hat{\eta}_a, \hat{\mu}_a, \hat{\gamma}_{m,a}, \hat{\gamma}_{t,a}) = \underset{\eta,\mu,\gamma_m,\gamma_t}{\arg\min} \left\{ \sum_{m=1}^{M_a} \sum_{t=1}^{T} \hat{\omega}_{m,a} \hat{\lambda}_{t,a} \cdot \right. \\ \left. (\tilde{y}_{mt} - \mu - \gamma_m - \gamma_t - \eta \cdot \text{Treated}_{mt})^2 \right\} \tag{2}$$

where $\tilde{y}_{mt}$ denotes the outcome variable (e.g., price, review volume), and Treated_{mt} is an indicator equal to one if market m is exposed to treatment by time t. The parameter η measures the causal effect of Amazon brand entry. The unit weights $\hat{\omega}_{m,a}$ align pre-treatment trends between treated and control units, while the time weights $\hat{\lambda}_{t,a}$ balance across the pre- and post-treatment periods. The intercept μ captures the baseline level of the outcome variable, γ_m accounts for the market-fixed effect, and γ_t captures the time-fixed effect. To estimate an overall treatment effect, I compute a weighted average of the cohort-specific treatment effects, using weights proportional to the number of newly treated markets in each cohort.

4.2.2 Panel-adapted Double Machine Learning

The Double Machine Learning (DML) approach has gained prominence in recent years (Dube et al., 2020; Gordon et al., 2023; Ellickson et al., 2023; Xu et al., 2024; Shi et al., 2024), as it enables valid causal inference in high-dimensional settings by combining the predictive flexibility of machine learning with the statistical rigor of semi-parametric estimation. It is especially useful in this setting to estimate treatment effects while accounting for high-dimensional control variables, and to explore how the effects vary by seller type and product similarity.

To estimate product-level heterogeneity using the DML framework in a panel setting, I address two key challenges: unobserved fixed effects and intertemporal dependence within a product. First, to account for unobserved product- and time-fixed effects, I adopt the approach proposed by Clarke and Polselli (2025), augmenting the covariate set $\tilde{X}_{it}$ with time-averaged covariate vector for product i, $\bar{X}_i$, as specified in Equation 3. Time fixed effects are captured through a saturated set of year-month dummies. Second, to address intertemporal dependence within products, I implement 5-fold cross-validation using chronologically ordered splits. Each fold contains a contiguous sequence of time periods, thereby ensuring that future observations are not used to predict past outcomes and preventing information leakage (Varian, 2014).

$$\tilde{X}_{it} = \left(X_{it}, \bar{X}_i\right) \quad \text{where} \quad \bar{X}_i = \frac{1}{T_i}\sum_{t=1}^{T_i} X_{it} \tag{3}$$

Specifically, following Dube et al. (2020) and Xu et al. (2024), I implement the DML approach in three steps:

$$\textit{Step 1:} \quad Y_{it} = g(W_{it}) + \alpha_i + \xi_t + \varepsilon_{it}, \tag{4}$$

$$\textit{Step 2:} \quad D_{it} = m(W_{it}) + \alpha_i + \xi_t + \mu_{it}, \tag{5}$$

$$\textit{Step 3:} \quad \tilde{Y}_{it} = \theta(X_i) \cdot \tilde{D}_{it} + \varepsilon_{it}, \tag{6}$$

where Y_{it} denotes the outcome variable (e.g., log price or consumer perceptions of product i at month t), D_{it} is an indicator equal to one if product i is treated by the entry of Amazon brand at month t, W_{it} are high-dimensional controls for product i at month t, such as product description embeddings, image embedding, product category dummy and brand indicators, α_i is the product-fixed effects uncorrelated with the covariates, ξ_t is the year-month dummy to control for time-fixed effects, and X_i includes the heterogeneous features of interest (e.g., seller type or similarity).

In Step 1, I model the outcome Y_{it} as a function of high-dimensional covariates W_{it}, product fixed effects α_i, and time fixed effects ξ_t, capturing baseline variation through a flexible function $g(\cdot)$. In Step 2, I similarly model the treatment variable D_{it} using a separate machine learning model $m(\cdot)$, again controlling for α_i and ξ_t. In Step 3, I compute residualized outcomes $\tilde{Y}_{it}$ and treatments $\tilde{D}_{it}$ from the first two stages and estimate the heterogeneous treatment effect $\theta(X_i)$ using a Causal Forest model (Athey and Wager, 2018; Athey et al., 2019; Guo et al., 2021), where X_i includes product-level moderators such as seller type and similarity to Amazon brands.

The intuition of this DML approach is that high-dimensional control variables, such as product descriptions, visual design features from product images, and sales ranks, may be correlated with both whether this product is affected by Amazon brand entry and the outcome variables, including price and consumer perceptions of value and quality. To address this potential selection bias, I use the residuals from the outcome and treatment models, i.e., the portions of variation in Y_{it} and D_{it} that the control variables cannot explain. By regressing the residualized outcome on the residualized treatment in Step 3, I isolate the variation in treatment that is orthogonal to the controls, thereby estimating the conditional treatment effects.

5 Findings

5.1 Effects of Platform Brand Entry on Market-level Price, Engagement and Consumer Perception

At the market level, I focus on the effects on price and engagement. For market price, I analyze two variables: average price and price range. The average price is measured as the natural logarithm of the market-level mean price to reduce skewness and facilitate percentage-based interpretation. The price range is calculated as the difference between the highest and lowest product prices available in each market-month. For engagement, I also consider two variables: product review volume and average rating. As the extensive literature indicates, review volume and rating reflect the degree of consumer engagement and have a profound effect on product sales (Chevalier and Mayzlin, 2006; Liu, 2006; Chen and Xie, 2008; Kim et al., 2022). Review volume is measured as the total number of new reviews added in each market-month. Average rating is calculated as the mean rating of all new reviews posted within each market-month.

Table 2 reports the market-level effects of Amazon brand entry using the staggered SynthDiD. To examine heterogeneity across brand tiers, I analyze three groups of Amazon brands: all brands (Panel A), Amazon Basics only, the flagship and most widely recognized Amazon brand, positioned to offer low-cost, functional alternatives (Panel B), and the Amazon Premium brands of Stone & Beam and Rivet (Panel C).

As shown in Panel A, the entry of Amazon brands leads to a significant price reduction, where the average

Table 2: Effects of Amazon brand Entry on Market Price and Product Reviews

	(1) Average Price	(2) Price Range	(3) Product Review Volume	(4) Average Rating of Reviews
Panel A: All Amazon Brands				
treatment effect	-0.093**	-38.527**	13.903*	0.307***
	(0.040)	(16.206)	(8.175)	(0.1361)
# of Observations	135,536	135,536	135,536	135,536
Panel B: Amazon Basics				
treatment effect	-0.119**	-38.708**	20.537***	0.417***
	(0.048)	(18.721)	(7.571)	(0.219)
# of Observations	133,644	133,644	133,644	133,644
Panel C: Amazon Premium Brands				
treatment effect	-0.000	-40.647	-7.808	-0.119
	(0.073)	(31.686)	(11.506)	(0.248)
# of Observations	126,764	126,764	126,764	126,764

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

price decreases by 8.880% (column 1: -0.093)[5]. Furthermore, panels B and C show that the entry of Amazon Basics primarily drives these effects. Moreover, I find a significant increase of 13.903 reviews in product review volume (column 3: 13.903), after controlling for the number of products per market. This could signal higher sales or a higher propensity for consumers to leave reviews, both of which increase the visibility of competing products. Furthermore, the average rating also increases by 0.307 (column 4: 0.307) overall, suggesting that sellers may be improving their product offerings to remain competitive. Overall, rather than harming consumer experience (The Markup, 2021), my analysis shows that the entry of platform brands steers the market toward more affordable prices, increased visibility, and higher product standards, suggesting a market disciplining effect which may benefit consumers.

To assess the effect of Amazon brand entry on consumer perceptions, I focus on value perception and quality perception. These two aspects are central to consumer decision-making and brand evaluations (Chen et al., 2020; Puranam et al., 2021). Both measures are derived from aspect-level sentiment analysis of review texts. Quality perception is calculated as the average sentiment across five dimensions: performance, aesthetics, durability, ease of assembly, and product-description fit. For each perception construct, I examine both the average sentiment score to capture shifts in consumer attitudes and the percentage of reviews mentioning the aspect to capture changes in consumer attention. All measures are aggregated at the market-month level.

Table 3 shows a surprising divergence that, despite the positive market-level outcomes, consumer perceptions of value and quality decline. Although market prices drop and average ratings rise, the average perceived value score falls by 0.458 and perceived quality drops by 0.888. This could be because the entry of low-cost products, like Amazon Basics, establishes a new, more critical benchmark for what consumers expect from a product at a

[5] A coefficient of -0.093 corresponds to a price change of $exp(-0.093) - 1 = 0.0888$ or an 8.88% decrease in average price.

Table 3: Effect of Amazon Brand Entry on Value and Quality Perception

	Value Perception		Quality Perception	
	AvgScore (1)	% Mentions (2)	AvgScore (3)	% Mentions (4)
Panel A: All Amazon Brands				
treatment effect	-0.458**	0.075***	-0.652***	-0.081***
	(0.186)	(0.027)	(0.223)	(0.032)
Observations	31,165	31,165	31,165	31,165
Panel B: Amazon Basics				
treatment effect	-0.607***	0.082**	-0.888***	-0.103**
	(0.119)	(0.037)	(0.249)	(0.045)
Observations	29,900	29,900	29,900	29,900
Panel C: Amazon Premium Brands				
treatment effect	0.012	-0.031	0.104	-0.027
	(0.299)	(0.046)	(0.361)	(0.053)
Observations	27,255	27,255	27,255	27,255

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Value Perception* refers to consumers' evaluations of product price and overall value. *Quality perception* is a composite measure based on the average of all dimensions of perceived product quality. *AvgScore* indicates the valence of consumer perception for the corresponding aspect toward non-platform competing products; *% Mentions* refers to the proportion of reviews that mention the corresponding aspect relative to the total review volume.

given price. Alternatively, it might reflect a shift in the consumer base toward more price-sensitive and critical shoppers. More importantly, consumer attention shifts from quality to value. The share of reviews mentioning value increases by 7.5%, while those mentioning quality fall by 8.1%. This suggests that as price competition intensifies, consumers may find it easier to compare prices than to assess complex quality attributes, making their evaluation criteria more uniform and focused on ”value for money.”

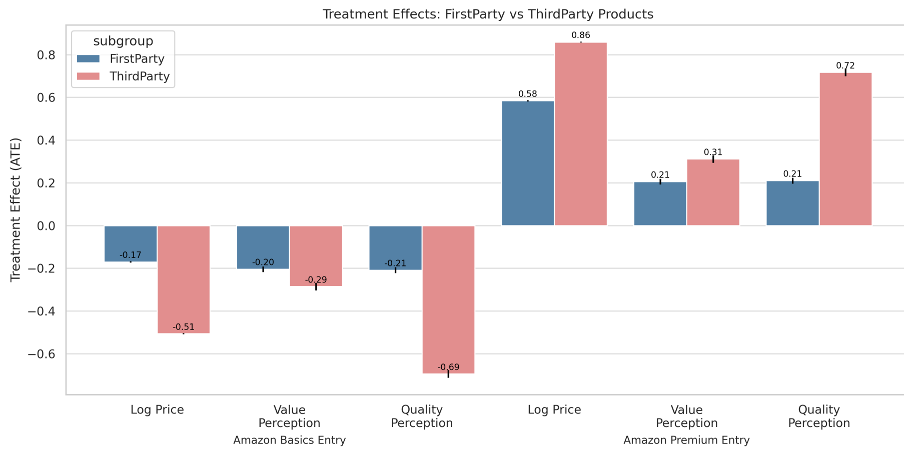
In summary, while platform brand entry improves market efficiencies through lower prices and higher ratings, it also leads to a perception cost, where consumers become more critical and less satisfied with product value and quality. This paradox suggests that market-wide improvements may mask concentrated dissatisfaction among specific product subsets. The next step is to examine how these effects vary by seller type and product similarity to the entering brands.

5.2 Heterogeneity Treatment Effect by Seller and Product

The product-level analysis uncovers significant heterogeneous treatment effects across different seller types and product characteristics. First, product-level estimates largely confirm the market-level pattern: prices decline and consumer perceptions worsen. However, while market-level ratings rise slightly, within-product ratings are essentially unchanged. This suggests the aggregate uptick reflects composition effects—exit of lower-rated items and entry of higher-rated ones—rather than improvements in existing products.

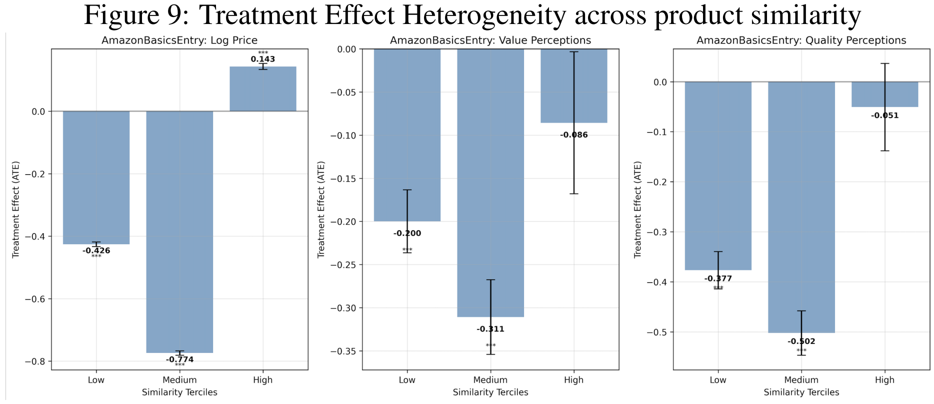
Second, as shown in Figure 8, I find that third-party sellers face a disproportionately greater competitive impact from Amazon Basics entry, with their products experiencing significantly larger price reductions and a more pronounced drop in perceived quality compared to first-party products. Conversely, the entry of premium Amazon brands has a positive effect on competing third-party products.

Figure 8: Treatment Effect Heterogeneity across Seller Types



Third, as shown in Figure 9, I also observe a non-linear relationship across product similarity to the Amazon brand. Contrary to typical substitution effects, products with high similarity to the Amazon brand experience a price increase, while those with medium similarity face the steepest price declines. This suggests that highly similar products may inadvertently receive some form of platform favoritism protection, whereas medium-similarity products, which lack both differentiation and protection, are the most vulnerable to competitive pressure.

Figure 9: Treatment Effect Heterogeneity across product similarity



In summary, these heterogeneous effects not only shed light on the underlying mechanisms of Amazon brand entry but also offer strategic guidance for sellers, highlighting the importance of positioning and differentiation, and provide regulators with evidence on how platform brands may unevenly reshape competition across seller types. We discuss these implications further in the next section.

6 Discussion

The rapid rise of platform brands, such as Amazon Basics, JD Jingzao, and Flipkart SmartBuy, has transformed the competitive dynamics of digital marketplaces. As platforms grow in scale and influence, their own brands become powerful strategic levers. In many cases, even the threat of platform brand entry can alter the behavior of third-party sellers (Jiang et al., 2011), creating a strategic uncertainty: should they compete on price, differentiate through design, or exit the category entirely?

Consumers, too, are adapting. As they become more familiar with platform brands and more attuned to price–quality tradeoffs, the standards for perceived value continue to rise, placing additional pressure on non-platform sellers to manage expectations and maintain credibility. At the same time, regulators are increasingly questioning the long-term impact of platform self-preferencing on innovation, competition, and consumer welfare. Against this backdrop, our findings offer managerial implications for sellers, platforms, and policymakers navigating this evolving landscape.

In this study, I show that platform brand entry acts as a powerful market disciplining force, leading to pro-competitive outcomes such as lower prices and increased consumer engagement. However, this market efficiency comes with a significant and often overlooked perception cost, where consumers become more critical, even as objective quality metrics improve. This paradoxical divergence highlights a potential market failure that should concern all stakeholders. For third-party sellers, the key takeaway is that differentiation is paramount; a pure price-matching strategy is unsustainable and leads to disproportionate negative effects, particularly for those with products of moderate similarity to the platform brand. Platform owners, in turn, must recognize that while their brands can enhance market efficiency, they risk undermining the ecosystem's diversity if they do not actively support third-party sellers and improve information transparency. Finally, for regulators, my work demonstrates the need to look beyond aggregate market metrics. A nuanced approach that considers the distributional effects of platform brand entry, including the uneven pressure on third-party sellers and the paradoxical consumer perceptions, is essential to craft effective policies that truly benefit consumer welfare and foster a healthy competitive environment.

As the first chapter of my dissertation, this study also opens several promising avenues for future research. First, while I focus on pricing, future research could use sales rank data to explore other competitive consequences, such as shifts in market dominance, changes in product visibility, and potential reductions in assortment diversity. Second, future studies could leverage advanced natural language processing to analyze consumer sentiment toward the platform itself, beyond individual products, to understand how the introduction of platform brands influences platform trust and reputation. Third, a deeper understanding of long-term competitive dynamics requires examining the strategic exit of both platform brands and third-party sellers. Research on product survival rates, especially for new or marginal sellers, would offer insights into the long-term effects of self-preferencing. Ultimately, a key area for future inquiry is the tension between consumer welfare and seller opportunity as platforms increasingly shift from neutral intermediaries to active market participants. Exploring this balance can inform both managerial strategy and regulatory policy regarding the governance of digital marketplaces.

References

Ailawadi, K. L., Neslin, S. A., and Gedenk, K. (2001). Pursuing the value-conscious consumer: Store brands versus national brand promotions. *Journal of Marketing*, 65(1):71–89.

Ailawadi, K. L., Pauwels, K., and Steenkamp, J.-B. E. (2008). Private-label use and store loyalty. *Journal of Marketing*, 72(6):19–30.

Angwin, J., Mattu, S., Larson, J., and Kirchner, L. (2016). Amazon's new competitive advantage: Putting its own products first. Accessed on 2024-07-15.

Arkhangelsky, D., Athey, S., A.Hirshberg, D., W.Imbens, G., and Wager, S. (2021). Synthetic difference-in-differences. *American Economic Review*, 111(12):4088–4118.

Athey, S., Tibshirani, J., and Wager, S. (2019). Generalized random forests. *The Annals of Statistics*, 47(2):1148–1178.

Athey, S. and Wager, S. (2018). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, 113(523):1228–1242.

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

Berman, R. and Israeli, A. (2022). The value of descriptive analytics: Evidence from online retailers. *Marketing Science*, 41(6):1074–1096.

Callaway, B. and Sant'Anna, P. H. C. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230. Preprint available at arXiv:1803.09015.

Campello, R. J. G. B., Moulavi, D., and Sander, J. (2013). Density-based clustering based on hierarchical density estimates. *Advances in Knowledge Discovery and Data Mining (PAKDD)*, 7819:160–172.

Chen, N. and Tsai, H. T. (2023). Price competition under information (dis)advantage. Working Paper 21, Platform Strategy Working Paper Series, Boston University. SSRN 4420175.

Chen, Y., Fay, S., and Wang, Q. A. (2020). Opinions as facts: The role of feedback in online retailing. *Journal of Marketing*, 84(1):102–120.

Chen, Y. and Xie, J. (2008). Does online word-of-mouth matter? (evidence from online book reviews). *Journal of Marketing Research*, 45(3):436–445.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., and Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, 21(1):C1–C68.

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

Chintagunta, P. K., Bonfrer, A., and Song, I. (2002). Investigating the effects of store-brand introduction on retailer demand and pricing behavior. *Management Science*, 48(10):1242–1267.

Cichocki, A., Zdunek, R., Phan, A. H., and ichi Amari, S. (2009). *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-Way Data Analysis and Blind Source Separation*. John Wiley & Sons, Chichester, UK.

Clarke, D., Pailañir, D., Athey, S., and Imbens, G. W. (2023). Synthetic difference-in-differences estimation. Technical report, Technical Report, UC Davis / Stanford / UC Berkeley / Harvard. Implements the SDID estimator of Arkhangelsky et al. (2021).

Clarke, P. S. and Polselli, A. (2025). Double machine learning for static panel models with fixed effects. *The Econometrics Journal*, 28(2).

Decarolis, F. and Rovigatti, G. (2021). From mad men to maths men: Concentration and buyer power in online advertising. *American Economic Review*, 111(10):3299–3327.

Deng, Y., Tang, C. S., Wang, W., and Yoo, O. S. (2023). Can third-party sellers benefit from a platform's entry to the market? *Service Science*, 15(4):265–284.

Desai, S., Eklund, J. E., and Lappi, E. (2020). Entry regulation and persistence of profits in incumbent firms. *Review of Industrial Organization*, 57(3):537–558.

Dube, A., Jacobs, J., Naidu, S., and Suri, S. (2020). Monopsony in online labor markets. *American Economic Review: Insights*, 2(1):33–46.

Dubé, J. H. (2024). Amazon private brands: Self-preferencing versus traditional retailing. *Antitrust Law Journal*, 86(1):115–148.

Ellickson, P. B., Kar, W., and Reeder, J. C. (2023). Estimating marketing component effects: Double machine learning from targeted digital promotions. *Marketing Science*, 42(4):704–728.

Feng, X. F., Liu, X., Zhang, S., and Srinivasan, K. (2024). Sustainability and competition on amazon. Working Paper 4958107, SSRN.

Garvin, D. A. (1984). What does "product quality" really mean? *MIT Sloan Management Review*, 26(1):25–43.

Gilardi, F., Alizadeh, M., and Kubli, M. (2023). Chatgpt outperforms crowd workers for text-annotation tasks. *Proceedings of the National Academy of Sciences*, 120(30):e2305016120.

Gilbert, R. and Shelanski, H. (2023). Understanding dynamic competition. *Antitrust Law Journal*, 86(3).

Goldfarb, A., Tucker, C., and Wang, Y. (2022). Conducting research in marketing with quasi-experiments. *Journal of Marketing*, 86(3):1–20.

Gordon, B. R., Moakler, R., and Zettelmeyer, F. (2023). Close enough? a large-scale exploration of non-experimental approaches to advertising measurement. *Marketing Science*, 42(4):768–793.

Guo, T., Sriram, S., and Manchanda, P. (2021). The effect of information disclosure on industry payments to physicians. *Journal of Marketing Research*, 58(1):115–140.

Herrera, S. (2022). Amazon has been slashing private-label selection amid weak sales. Wall Street Journal, July 14, 2022. Accessed: 2025-07-15.

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

Kim, E., Ding, M. A., Wang, X. S., and Lu, S. (2022). Does topic consistency matter? a study of critic and user reviews in the movie industry. *Journal of Marketing*, 86(5):1–23.

Koch, G., Zemel, R., and Salakhutdinov, R. (2015). Siamese neural networks for one-shot image recognition. In *ICML Deep Learning Workshop*.

Korganbekova, M. and Korganbekov, A. (2024). Regulating data usage by dual platforms. Working paper, Booth School of Business and Boston University.

Kranz, S. (2021). Synthetic difference-in-differences with time-varying covariates. Technical report, Ulm University. Vignette for the R package `xsynthdid`.

Kumar, M., Eckles, D., and Aral, S. (2020). Scalable bundling via dense product embeddings. *arXiv preprint arXiv:2002.00100*.

Kumar, N. and Steenkamp, J.-B. (2007). *Private label strategy: How to meet the store brand challenge*. Harvard Business Review Press.

Lambrecht, A., Tucker, C. E., and Zhang, X. (2024). Tv advertising and online sales: A case study of intertemporal substitution effects for an online travel platform. *Journal of Marketing Research*, 61(2):248–270.

Lee, K. H. and Musolff, L. (2025). Two-sided markets shaped by platform-guided search. Working paper, Working Paper, National University of Singapore; Wharton School, University of Pennsylvania. Proprietary Amazon sales and BuyBox prominence data; structural model of intermediation power; consumer surplus analysis.

Liu, B. and Zhang, L. (2012). A survey of opinion mining and sentiment analysis. In Aggarwal, C. C. and Zhai, C., editors, *Mining Text Data*, pages 415–463. Springer.

Liu, L., Dzyabura, D., and Mizik, N. (2020). Visual listening in: Extracting brand image portrayed on social media. *Marketing Science*, 39(4):669–686.

Liu, Y. (2006). Word of mouth for movies: Its dynamics and impact on box office revenue. *Journal of marketing*, 70(3):74–89.

Long, F. and Amaldoss, W. (2024). Self-preferencing in e-commerce marketplaces: The role of sponsored advertising and private labels. *Marketing Science*, 43(5):925–952.

MobiLoud (2024). 50+ amazon statistics you should know in 2024. Accessed on 2024-07-15.

Momentum Commerce (2023). The relationship between amazon's choice badges and private label products. Accessed on 2024-07-15.

Momentum Commerce (2024). Amazon's private label market share shrinks by 6% year-over-year in q1 2024. Accessed on 2024-07-15.

Morrison, S. (2021). Amazon's strategy to squeeze marketplace sellers and maximize its own profits is evolving. Accessed on 2024-07-15.

Pachali, M. J. and Datta, H. (2025). What drives demand for playlists on spotify? *Marketing Science*, 44(1):54–64. Prepublished as SSRN 4079693 (revised March 2024);.

Pauwels, K. and Srinivasan, S. (2004). Who benefits from store brand entry? *Marketing Science*, 23(3):364–390.

Pauwels, K. and Srinivasan, S. (2009). Pricing of national brands versus store brands: Market power components, findings and research opportunities. *Handbook of Pricing Research in Marketing*, pages 258–282.

Puranam, D., Kadiyali, V., and Narayan, V. (2021). The impact of increase in minimum wages on consumer perceptions of service: A transformer model of online restaurant reviews. *Marketing Science*, 40(5):985–1004.

Qian, K. and Xie, Y. (2025). Express: The power of star creators: Evidence from the live streaming industry. *Journal of Marketing Research*, 0(ja):00222437251359163.

Raju, J. S., Sethuraman, R., and Dhar, S. K. (1995). The introduction and performance of store brands. *Management science*, 41(6):957–978.

Raval, D. R. (2023). Steering in one click: Platform self-preferencing in the amazon buy box. Working paper, Harvard Law School. First draft April 2021; revised December 31, 2023.

Sengupta, R. (2020). How the amazon and flipkart private labels affect other sellers on the ecommerce platforms. *YourStory*. Accessed: 2025-07-17.

Shi, B., Mao, X., Yang, M., and Li, B. (2024). What, why, and how: An empiricist's guide to double/debiased machine learning. Forthcoming in Information Systems Research.

Tang, C. S., Deng, Y., Wang, W., and Yoo, O. S. (2024). Can an e-commerce platform and its third-party sellers benefit from each other's market entry? *Production and Operations Management*, 33(1):69–86.

Technode (2019). Jd moves into c2m model with made-to-measure shirts. *TechNode*. Accessed: YYYY-MM-DD.

The Markup (2021). Amazon puts its own "brands" first above better-rated products. https://themarkup.org/amazons-advantage/2021/10/14/amazon-puts-its-own-brands-first-above-better-rated-products.

Varian, H. R. (2014). Big data: New tricks for econometrics. *Journal of Economic Perspectives*, 28(2):3–28.

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

Xu, Y., Ghose, A., and Xiao, B. (2024). Mobile payment adoption: An empirical investigation of alipay. *Information Systems Research*, 35(2):807–828.

Yang, Y., Cer, D., Ahmad, A., Guo, M., Law, J., Constant, N., Hernandez Abrego, G., Yuan, S., Tar, C., Sung, Y., Strope, B., and Kurzweil, R. (2020). Multilingual universal sentence encoder for semantic retrieval. In Celikyilmaz, A. and Wen, T., editors, *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pages 87–94, Online. Association for Computational Linguistics.

Zou, T. and Zhou, B. (2025). Self-preferencing and search neutrality on online retail platforms. *Management Science*, 71(5):4087–4107.