Price Competition Under Information (Dis)Advantage

Nan Chen Hsin-Tien Tsai*

October, 2023

Abstract

We examine the impact of asymmetric data access within a vertically integrated e-commerce platform. Using a unique daily panel, we find that both the platform owner and third-party sellers base pricing on past sales, but only the owner exploits competitors' sales data. We estimate a price competition model with heterogeneous seller learning. We find that (1) eliminating the owner's access to competitors' data increases social welfare by 0.33%, negatively impacting the owner through reduced first-party sales and (2) providing third-party sellers with equal access to competitors' data as the owner increases social welfare by 1.65%, benefiting the owner through increased referral fees and yielding a Pareto improvement.

Keywords: Price competition; e-commerce; digital platform; data access; information asymmetry

*Chen: Department of Economics, National University of Singapore, nan@nus.edu.sg. Tsai: Department of Economics, National University of Singapore, ttsai@nus.edu.sg. For their helpful comments, we thank David Byrne, Tat Chan, Liang Guo, Jasmine Hao, Ginger Zhe Jin, Martin Peitz, Ralf van der Lans, Joel Waldfogel, Kevin Williams, and Dongni Zhu. We thank Mengyang Su and Xinbo Wang for their research assistance. This research has received financial support from the tier-1 grant FY2021-FRC2-005 provided by the Ministry of Education in Singapore.

1 Introduction

Classic theories on the role of information in determining competitive outcomes (e.g., Kühn and Vives, 1995) have gained increased relevance in today's digital era. Data, although characterized as the "non-rival" new oil (e.g., Varian, 2018), exhibits potential unequal access for firms. We investigate the implications of such disparities within a dominant, vertically integrated platform where the platform owner competes with hosted third parties. Specifically, we examine the impact of asymmetric access to historical sales data on equilibrium payoffs and welfare in the context of price competition between Amazon and third-party sellers.[1]

This empirical context is chosen for two primary reasons. First, the distinct market structure of vertically integrated platforms offers a unique setting to examine the consequences of asymmetric data access. In our case, the e-commerce platform owner manages all transactions and maintains extensive sales data, while third-party sellers likely possess limited information beyond their own sales. This context, ex-ante, serves as a suitable starting point. Second, data access design within these vertically integrated platforms holds inherent importance for maximizing social welfare (e.g., Crémer et al., 2019).

We assemble a daily panel dataset for products in the Home & Kitchen category on Amazon.com over a 7-month period. The dataset includes over 122,000 products, each with two sellers offering the same product in new condition, allowing us to investigate price competition in a duopoly market, which is the most popular oligopoly market structure on the platform. We collect daily information, such as price and inventory, for each product listing (i.e., a seller of a given product). We employ inventory data to approximate daily sales at the listing level. This measurement of listing-level sales, combined with the unique design of information access on the platform, facilitates our analysis.

[1]Recent empirical research (e.g., Bajari et al., 2019) highlights the significance of past sales data for demand forecasting in e-commerce.

We begin by presenting simple reduced-form evidence on how sellers set prices based on past sales. The high-frequency pricing and sales data enable us to obtain a plausibly causal understanding of sellers' learning behaviors. First, we demonstrate that sellers tend to gradually decrease their prices when experiencing consecutive days without sales, aligning with learned pessimistic beliefs about the demand state (e.g., Mason and Välimäki, 2011). Correspondingly, we discover that sellers are likely to increase their prices immediately following a sale event, suggesting that they may adjust their prices in response to beliefs about stronger demand. Both Amazon and third-party sellers are highly responsive to their own sales and lack of sales.

Subsequently, we present novel evidence of asymmetric learning contingent on competitors' sales, which has not been previously explored in the literature. Although competitors' sales data is not publicly available on the platform, Amazon theoretically has access to this information due to its role as a platform owner and, a priori, can learn from competitors' sales. To test this hypothesis, we demonstrate that when Amazon's competitors make sales, Amazon adjusts its prices accordingly. Conversely, third-party sellers do not react to their competitors' sales. This discrepancy in the ability to respond to competitors' sales highlights the potential informational advantage Amazon possesses as the platform owner when competing against third-party sellers. To strengthen the point that information, not sophistication, is the key driver, we show that, although large third-party sellers demonstrate stronger responses to their own sales when compared to smaller sellers, their reactions to competitors' sales remain similarly muted. This interpretation aligns with the findings from our industry surveys and is further supported by the later structural estimates.

To understand what could happen, we consider a simple theoretical model of linear demand in a duopoly price competition. The benchmark is where Firm 1 (the platform owner) knows the demand intercept of Firm 2 (a third-party seller), while Firm 2 is unaware of Firm 1's demand intercept. When both firms know each other's demand intercept, losing its information advantage always leads to a decline in Firm

1's profit without vertical incentives (i.e., charging no referral fee). However, if vertical incentives are present, the changes in both Firms' profits and overall consumer welfare are theoretically ambiguous. On the other hand, when neither firm has information about the other's demand intercept, Firm 2 always benefits after its competitor loses its information advantage, while Firm 1's profit and overall consumer welfare remain theoretically uncertain.

Next, we build a simple structural model of price competition, allowing firms to have heterogeneous beliefs on the state of demand. Using our long panel of prices and realized sales, we estimate demand states, which we term as true demand states h. Subsequently, we employ a flexible prediction model that uses historical true demand states (at days $< t$) to forecast the current demand state (at day t), referred to as the best predicted demand state h^{BP} (at day t). The key distinction lies in the fact that true demand states h are ex-post, determined after day t sales are realized, while best predicted demand states h^{BP} are ex-ante, determined based on past sales prior to day t sales realization. h and h^{BP} are correlated. The strength of the correlation, emerged from the data, indicates the usefulness of historical sales data in predicting current demand states. Firms' beliefs about the current demand state, depending on their access to data, are constructed by a simple reduced-form function of h^{BP}. The function is monotonic such that superior data access results in a belief closer to h^{BP}. Analogous to the reduced-form analysis, firms' heterogeneous belief functions are identified by how their prices respond to past sales (or h^{BP}).

On the demand side, we assume consumers arrive following a Poisson process, with preferences conforming to a nested logit model. The model groups sellers of a specific product together in one nest and the outside option in another. The utility model incorporates the true demand state as an intercept, allowing for preference heterogeneity through a set of fixed effects and random effects in consumer utility functions. On the supply side, each seller calculates their sales based on their own belief about demand. Using these beliefs, sellers participate in a simultaneous-move pricing game and make pricing decisions.

We estimate the model using the Method of Simulated Moments. Model parameters are selected based on their ability to replicate empirical moments, such as price, sales, and their interactions. Our estimation involves two main efforts. First, we employ a full solution approach to estimate the model, which involves jointly solving the demand and supply model to explicitly account for price endogeneity. We discretize all observed and unobserved states. For unobserved states (i.e., random effects in consumers' utility function and firms' cost function), we utilize numerical integration to compute model moments. Second, we exploit within-product listing variations in price. To achieve this, we choose a sample of local reprice events for estimation. We assume that within a 3-day time window, price variations are driven by firms' beliefs or costs, and on average, the underlying demand remains smooth. Our approach requires solving equilibria for 12,960 pricing games during each iteration.

We estimate an average own-price elasticity of -18.96, in line with existing literature (e.g., Ellison and Ellison, 2009; Dinerstein et al., 2018). The high own-price elasticity primarily results from intense competition when two sellers compete head-to-head on the same product under a shared product page. The average cross-price elasticity is 18.41, which is high as expected but still lower than a classic Bertrand model would predict. The average market price elasticity (when two sellers raise prices simultaneously) is -5.28, indicating the level of competition against the outside option, encompassing all other products on the platform as a subset. For a product priced at around \$70, consumers derive additional utility worth \$7.16 and \$6.13 when purchasing from Amazon and Fulfillment by Amazon (FBA) sellers, respectively, compared to Fulfillment by Merchant (FBM) sellers. Consumers receive an additional utility worth \$5.31 when purchasing from Amazon compared to FBA sellers. On the supply side, before accounting for referral fees, Amazon and FBA sellers demonstrate lower costs compared to FBM sellers, at \$1.81 and \$8.77. Moreover, FBA has lower costs relative to Amazon by an average of \$14.72. Overall, sellers' cost heterogeneity, as indicated by the variance of random effects on cost shocks, is substantial.

Our estimates quantitatively evaluate differences in beliefs between Amazon and third-party sellers using a rational framework of price competition. First, the estimates for own-belief parameters indicate that both Amazon and third-party sellers possess accurate beliefs about their respective demand states. Their levels of accuracy are comparable, with both aligning well with their respective h^{BP} values, and Amazon is slightly more accurate. Second, for beliefs about competitors' demand states, Amazon's belief aligns with the competitor's h^{BP}. However, third-party sellers' beliefs closely resemble random guesses. Overall, these findings are consistent with the earlier evidence, highlighting the presence of heterogeneous learning and an information advantage for Amazon over third-party sellers.

We conduct several counterfactual analyses to examine the impact of information advantage on price competition, specifically: (1) preventing Amazon from using marketplace data for its own retail business and (2) sharing marketplace data with third-party sellers.

In the first counterfactual, we examine a scenario where Amazon is prohibited from using competitors' past sales data for pricing. We model this scenario by setting Amazon's belief parameter to be identical to that of third-party sellers. Under the current 15% referral fee and restricted information access for Amazon, Amazon's first-party sales decrease by 0.64%, while profits drop by 0.30%. With the information disadvantage removed, third-party sellers experience a 2.63% increase in sales and a 2.36% profit gain. Overall, consumer welfare and social welfare improve by 0.48% and 0.33%, respectively, implied by reduced prices and increased quantity. We further emphasize the role of vertical market structure in our findings by examining the scenario where Amazon charges a 0% referral fee, eliminating its vertical incentives. In this case, when we eliminate Amazon's information advantage, we observe a larger decrease in Amazon's profits of about 4.68%, while third-party sellers' profits increase by 7.26%. It suggests that the presence of a referral fee could potentially deter Amazon from using its informational advantage to the detriment of third-party sellers.

In the second counterfactual, we investigate a situation where Amazon shares its own sales data with third-party sellers, providing equal access to information. We simulate this scenario by adjusting the belief parameter of third-party sellers to match that of Amazon. Our findings indicate a substantial increase of 14.91% in sales for third-party sellers, while Amazon's sales experience a 4.08% decline. Interestingly, due to the revenue generated from the referral fee, Amazon's profits do not decrease; instead, they increase by 0.45%. Moreover, information sharing leads to a more substantial rise in consumer welfare and social welfare on average, with estimated increases of 2.08% and 1.65%, respectively. We also explore the role of vertical incentives as well as information sharing in non-Amazon markets.

In summary, this paper examines price competition in a dominant, vertically integrated e-commerce platform where the owner possesses a data access advantage over its hosted competitors. We provide novel empirical evidence using new data and estimate a structural model that enables us to study relevant counterfactuals. We find that, compared to the current unequal access to information regarding competitors' sales data, equal data access, achieved through either eliminating Amazon's information advantage or sharing this information with third-party sellers, increases social welfare. Importantly, sharing information with third-party sellers leads to a more substantial increase in social welfare, yielding a Pareto improvement that benefits the platform owner, third-party sellers, and consumers.

1.1 Related Literature

This study builds on the theories of information sharing in oligopolies (e.g., Vives, 1984; Li, 1985; Gal-Or, 1985). A recent stream of empirical work study the impact of information disclosure in competitive markets. Rossi and Chintagunta (2016), Luco (2019), and Ater and Rigbi (2023) provide novel evidence on the competitive effects of mandatory price-disclosure policies in offline markets. These studies employ a reduced-form approach, focusing on the transparency of a strategic decision variable important to both firms and consumers.[2] In a concurrent study, Byrne et al. (2023) examine the impact of asymmetric access to competitors' price information in retail gasoline markets. Our paper introduces past sales information, which can be used as an input for predicting current demand states (e.g., Bajari et al., 2019).[3]

By incorporating demand states and heterogeneous beliefs, our paper complements recent studies that investigate the competitive effects of pricing technologies which adjust according to a rival's historical prices (e.g., Calvano et al., 2020; Klein, 2021; Clark et al., 2023; Rhodes et al., 2023). Brown and MacKay (2023) present empirical evidence and a model in which firms can differ in pricing frequency and choose pricing algorithms that are a function of rivals' prices. Asker et al. (2023) study the prices generated by AIs that use different learning protocols when there is market interaction.

Our empirical approach, which models firms' beliefs as a function of past data and identifies these beliefs from the dependence of firms' strategies on past data, is influenced by the recent literature in empirical Industrial Organization. This literature examines how firms learn and form beliefs in competitive environments (e.g., Doraszelski et al., 2018; Jeon, 2022; Huang et al., 2022).[4] Asker et al. (2020) develop empirical methodologies to examine competitive information sharing in a dynamic auction environment. Moreover, the heterogeneities in beliefs identified among Amazon and third-party sellers are related to the work of Goldfarb and Xiao (2011) and Hortaçsu et al. (2019), who employ the Cognitive Hierarchy model to empirically investigate variations in the strategic sophistication of firms.

[2]A classic empirical literature highlights how the disclosure of product quality information affects equilibrium outcome (e.g., Jin and Leslie, 2003; Tadelis and Zettelmeyer, 2015).

[3]Refer to Fudenberg and Villas-Boas (2006) for the theoretical literature on pricing with consumer-level past sales data.

[4]For a comprehensive review, see Aguirregabiria and Jeon (2020) and Aguirregabiria (2021).

This study focus on the design of access to data for sellers in the context of a vertically integrated platform which relates to the incentives and efficiency in vertically integrated markets (e.g., Hart et al., 1990; Rey and Tirole, 2007). Specifically, we add to a new and growing empirical literature on vertically integrated platforms that focus on the design of information for consumers and its implications (e.g., Chen and Tsai, 2019; Lee and Musolff, 2021; Lam, 2021; Raval, 2022; Farronato et al., 2023; Reimers and Waldfogel, 2023). We present a new, empirically unexplored perspective: the platform owner's information advantage over third-party players.[5] In a related context, Madsen and Vellodi (2022) theoretically investigate how a platform owner may leverage its informational advantage to imitate products offered by third-party sellers.

The remainder of this paper is organized as follows. Section 2 provides an overview of the empirical context. Section 3 describes the data and provides summary statistics. Section 4 presents evidence of asymmetric seller learning. Section 5 outlines the structural model. Section 6 discusses our estimation method and identification strategy. Section 7 presents our estimates and results from the counterfactual analyses. Finally, Section 8 concludes the paper. Additional technical details and robustness checks are available in the appendices.

2 Background

Amazon is a leading online retailer that has over 2.5 million third-party sellers on its platform as of 2021. These sellers are independent businesses that use Amazon's platform to sell their products to customers worldwide and make up a significant

[5]The closest empirical work to this perspective is Zhu and Liu (2018), which examines platform owners entering third-party selling markets but does not specifically identify the owner's information advantage. Hagiu et al. (2022) and Gutierrez (2021) examine the implications of vertically integrated platforms from a broader perspective. In a broader context, our study addresses small firms' access to non-personal and non-rival data (e.g., Jones and Tonetti, 2020; Bergemann et al., 2022)

portion of Amazon's overall sales. In addition to third-party sellers, Amazon also sells its own products, including both private-label and branded products, and directly competes with its third-party sellers.[6]

Amazon Market Structure There has been much discussion in academic and policy circles about Amazon's dual role as both the owner of its marketplace and as a retailer that competes with other sellers on the platform. This has led to concerns about the potential for Amazon to use data from third-party sellers, which it has access to as the platform owner, in ways that may not be in the best interests of its third-party sellers or consumers. Some have raised concerns that Amazon may misuse this data in its pricing and product launching decisions, or exert its influence in other unfair ways.

While we focus on Amazon's US store, in July 2019, the European Commission initiated an inquiry into Amazon's utilization of marketplace vendor data.[7] By November 2020, the Commission issued a Statement of Objections, highlighting that "*very large quantities of non-public seller data are available to employees of Amazon's retail business and flow directly into the automated systems of that business, which aggregate these data and use them to calibrate Amazon's retail offers and strategic business decisions to the detriment of the other marketplace sellers. For example, it allows Amazon to focus its offers in the best-selling products across product categories and to adjust its offers in view of non-public data of competing sellers.*"[8] In December 2022, Amazon pledged to refrain from using marketplace seller data, specifically including "sales data," for its retail operations, which specifically covers "Retail

[6]A third-party seller can be under either the "Fulfillment by Amazon" (FBA) or "Fulfillment by Merchant" (FBM) program. Under the FBA program, sellers can send their products to an Amazon fulfillment center, where Amazon is responsible for shipping and other related services. However, in addition to referral fees, sellers who choose FBA also pay additional FBA fulfillment fees (see https://sell.amazon.com/pricing#fulfillment-fees for more information).

[7]See https://competition-cases.ec.europa.eu/cases/AT.40462.

[8]See https://ec.europa.eu/commission/presscorner/detail/en/ip_20_2077.

Operations decisions to set the prices of ASINs (i.e., products)."[9]

Sales Information on Amazon Data plays an increasingly critical role in shaping strategic decisions, and its impact on a company's performance can be significant (e.g., Bajari et al., 2019). One way of demonstrating this importance is by using sales data to forecast demand and inform pricing decisions. Amazon and its third-party sellers are known for dynamic pricing (also known as algorithmic pricing). The prices of products on the site may vary in response to changes in customer demand and competitors' actions based on the information sellers have learned.

Our focus is on past sales information. Theoretically, the cost of providing sales data is similar to that of providing price data. While price data is transparent and relatively inexpensive to obtain through public APIs, sales data is not. To put this into perspective, we conduct a survey covering 19 prominent e-commerce platforms across the US, China, Europe, Japan, and Southeast Asia. Notably, we find that 47.4% of these platforms do disclose sales data. Among them are platforms such as eBay, Taobao, Allegro, AliExpress, and Shopee (refer to Appendix A for details).

As the platform owner, Amazon oversees all transactions and maintains extensive sales data, allowing it to access this information with virtually no marginal cost. Conversely, third-party sellers can observe their own sales but are likely to have limited knowledge beyond their own sales figures.

Inferring individual listing sales within a product category is challenging based on the publicly available information in the marketplace. However, there are a few approaches that can provide some insights into competitors' sales. Amazon's primary sales metric, the product's sales rank, consolidates sales data from all listings of a given product and ranks its total sales within its product category. Though third-party sellers can utilize their own sales data and sales rank to estimate competitors' sales to a certain extent, the information is very limited, as the sales rank only

[9]See `https://ec.europa.eu/competition/antitrust/cases1/202252/AT_40462_8825091_8265_4.pdf`.

represents a ranking and not the actual number of sales. Consequently, it can be influenced by sales of other products.

Services like Jungle Scout offer professional sales estimation for sellers, but the cost is substantial. As of April 15th, 2023, tracking 20 products costs $349 per year.[10] Another method involves manually monitoring competitors' inventory and using high-frequency inventory changes to infer sales (e.g., He et al., 2022). However, this approach is also labor-intensive and sometimes impractical.

After drafting the initial version of this paper, we came across potential A/B testing conducted by Amazon, which hinted at the possible disclosure of some type of "sales data."[11] However, it's important to note that Amazon's public disclosure of sales data is still in its fairly preliminary stages. Furthermore, Amazon has not yet shown any intentions of disclosing seller-listing level sales data, which is the primary focus of our study.

Automated Pricing on Amazon Amazon provides professional sellers with a free automated pricing tool that adjusts prices in "real time" based on either competitor prices (namely the "Featured Offer Rule" and "Competitive Price Match Rule") or their own sales, known as the "Sales Bases Rule."[12] The former allows sellers to monitor and react to competition, while the latter explicitly incorporates a pricing strategy that is contingent on own sales but *not* on competitors' sales. These algorithms function as heuristics determined by individual sellers. We also survey 11 leading automated pricing service providers on the internet. We find that while

[10]See `https://www.junglescout.com/pricing/`.

[11]This practise was observed by industry researchers including Waters (2023), who commented that "*[t]he introduction of public-facing sales data on Amazon is part of a broader trend. After years of holding tight to its customer information, Amazon is making more and more of its data available to brands. It's a huge opportunity that many are still overlooking... Ideally, we want brands to use sales data to gain a richer view of how they compare to their competitors.*" See Appendix A for details.

[12]See `https://sell.amazon.com/tools/automate-pricing`.

most algorithms are contingent on competitors' prices and own sales, none of them is contingent on competitors' sales (see Appendix B).

The reduced-form results in Section 6 indicate that sellers adopt pricing rules based on sales, but they are unable to base prices on competitors' sales, closely aligning with this institutional setup. In the structural model, we assume that, on average, these heuristics align with profit-maximizing motives. This profit-maximizing assumption is considered to be more suitable than modeling specific and exact heuristics because sellers may select varying heuristic rules based on different, unobserved circumstances that researchers cannot observe. While these heuristics are endogenous and may change in the counterfactual, the underlying profit-driven incentive is likely to remain consistent.

3 Data

We collect a daily panel dataset from Amazon.com for a period of 7 months for the Home & Kitchen category.[13] We use the historical information from Keepa.com to identify products that had an average of between 1.5 and 2.49 listings in new condition over the past 90 days.[14] From this pool of candidate products, we select the most popular products based on their average sales rank over the past 30 days, excluding those with the identical sellers offering products in different conditions or fulfillment methods. This results in a final dataset of approximately 122,000 products. We follow a daily routine for collecting data on the final set of products. We gather information on different aspects of each seller of a given product, such as its price,

[13] The Home & Kitchen category on Amazon is one of the largest, comprising a wide range of products commonly purchased by consumers, including kitchen appliances, cookware, and home décor. It is also the most popular category among sellers, according to a report by Jungle Scout (see https://www.junglescout.com/amazon-seller-report/?utm_source=twitter&utm_medium=organic&utm_content=report&utm_campaign=sots_2022).

[14] The listings usually come with various conditions, such as new, used, and refurbished.

shipping fee, seller types, and inventory. We use the current available quantity of a product on the website as our measure for inventory. When the available inventory matches the maximum quantity that can be purchased in a single transaction (e.g., 999 or a purchasing limit), we consider the inventory as unknown and do not use it for our analysis. These missing data are less common for products in the Home & Kitchen category than other categories such as Electronics.

3.1 Summary Statistics

In Table 1, we present statistics for the key variables in our data, separated into seller types whether they are third-party sellers or Amazon itself. Among the product listings in our dataset, 23% come from Amazon.

The Price variable is the amount consumers pay for a single unit of the listed product, excluding shipping fees. In our data, third-party sellers have an average price of \$49.25, while Amazon has an average price of \$66.89. Differences in price levels can be driven by both product and seller-level preferences and costs, which we will separate in our structural estimation. The Shipping variable is a fixed fee charged per order by the seller. In most cases, shipping is free for both Amazon and third-party sellers who use Fulfillment by Amazon services. However, special handling products such as oversized or frozen items may have exceptions. On average, third-party sellers charge a shipping fee of approximately \$2.71, while Amazon charges an average shipping fee of \$0.02. Shipping fees change infrequently (less than 0.3%) in the data and are absorbed by fixed effects in our analysis.

The Inventory variable indicates the maximum number of units of a product that is available for purchase. Typically, Amazon has a larger inventory than third-party sellers, with average inventory levels of approximately 92.40 compared to 79.93, respectively. We define sales based on daily changes in inventory. For instance, if the inventory decreases from 10 on day $t-1$ to 9 on day t, we assign a sale of 1 for day t. On average, daily sales for a product in our dataset amount to approximately

$0.14 - 0.15$. We exclude cases where the inventory decreases by more than 100, as significant inventory reductions is likely due to sellers manually adjusting their inventories instead of sales. We also exclude cases where stock increases as it is likely due to restock. The assumption is that sales and restock events are unlikely to happen in the same period if the length of the period is small enough. Appendix D discusses concerns related to potential bias in sales estimates due to restocking within a given day. We use data for approximately 2,400 products at a two-hour interval and show that the probability of underestimating sales is less than 1% when the stock data frequency is daily.

The Fulfilled By Amazon variable indicates whether a seller uses Amazon's Fulfilled by Amazon services. All products sold by Amazon are fulfilled by Amazon, while 14% of product listings sold by third-party sellers in our dataset are fulfilled by Amazon.

The number of seller ratings is commonly used by industry practitioners to infer seller size.[15] We use it to define large and small third-party sellers in Section 4.4. On Amazon, consumers can rate sellers that they have purchased from. This seller rating contains feedback for all products listed by a specific seller. Amazon cannot be rated by consumers, so there are no seller-level ratings available for Amazon. The third-party sellers in the data have a median of 595 ratings.

3.2 Price Variation

Prices on the Amazon marketplace are known for their frequent fluctuations. We investigate the frequency and magnitude of price adjustments, which are defined as changes in price from one day to the next. As presented in Table 2, third-party sellers have a 2.79% likelihood of adjusting their prices, while Amazon has a 5.68% likelihood of adjusting prices. Prices can be adjusted upwards or downwards, and for both third-party sellers and Amazon, prices are more likely to decrease than

[15] See https://www.marketplacepulse.com/top-amazon-usa-sellers.

Table 1: Summary Statistics of Product Listings by Third-Party Sellers and Amazon

	Third-Party Sellers			Amazon		
	Mean	Median	STD	Mean	Median	STD
Price ($)	49.25	26.58	54.52	66.89	40.59	65.67
Shipping Fee ($)	2.71	0.00	7.44	0.02	0.00	1.02
Inventory	79.93	20.00	155.38	92.40	10.00	166.28
Sales	0.15	0.00	1.89	0.14	0.00	1.97
Fulfilled By Amazon	0.14	0.00	0.35	1.00	1.00	0.00
Number of Seller Ratings	11,935.60	595.00	42,418.70			
Number of Observations		37,666,471			11,422,349	

Note: The summary statistics in Table 1 include the price, shipping fee, fulfillment by Amazon status, inventory, and sales for all listings from third-party sellers and Amazon, respectively.

increase. On average, a price adjustment results in a −$0.39 change for third-party sellers and a −$0.37 change for Amazon. When prices are decreased, the average price decrease is $4.43 for third-party sellers and $4.63 for Amazon. When prices are increased, the average price increase is $5.26 for third-party sellers and $5.53 for Amazon. Compared to third-party sellers, Amazon is more likely to adjust its prices and make larger changes.

4 Reduced-Form Analysis

In this section, we utilize the high-frequency nature of the dataset to investigate sellers' pricing behaviors. We present reduced-form evidence to support the idea that sellers adjust their prices based on past sales: (1) prices tend to decrease when there are no recent sales, (2) prices immediately increase in response to sales, and (3) Amazon and third-party sellers differ in their ability to use competitor sales information to adjust their pricing, with Amazon being more responsive to competitor sales.

Table 2: Price Variation

	Third-Party Sellers		Amazon	
	Mean	STD	Mean	STD
Price Adjustment Probability (%)	2.79	16.47	5.68	23.14
Price Decrease Probability (%)	1.63	12.66	3.29	17.85
Price Increase Probability (%)	1.16	10.72	2.39	15.26
Price Change ($)	-0.39	21.16	-0.37	23.00
Price Decrease Change ($)	-4.43	23.30	-4.63	18.01
Price Increase Change ($)	5.26	16.11	5.53	27.41

Note: Table 2 shows the probability of a price adjustment and the magnitude of the corresponding price change for third-party sellers and Amazon, respectively.

We define the following notation. Let $m \in \mathbb{M}$ be a product market. Let $j \in \mathbb{J} = \{1, 2\}$ indicate a seller. Let $t \in \mathbb{T}$ denote calendar time. We denote the sales and price in product market m by seller j on calendar time t as $q_{m,j,t} \in \mathbb{Q} \subseteq \mathbb{N}$ and $p_{m,j,t} \in \mathbb{P} \subseteq \mathbb{R}^+$, respectively. Our data consists of a panel of product listings with subscripts $(m, j, t) \in \mathbb{D} \subseteq \mathbb{M} \times \mathbb{J} \times \mathbb{T}$. We define a sample selection procedure that takes a subset of data with a subscript $\mathbb{D}c \subseteq \mathbb{D}$ satisfying condition $c \in \mathbb{C}$ and re-index the data in an ordered manner. We denote this re-indexing procedure as $\mathcal{S} : \mathbb{D}c \to \mathbb{S} \subseteq \mathbb{N}$. Throughout our analysis, we use this sample selection procedure with varying selection conditions c. We primarily use regression models to estimate conditional means, which provide interpretable estimates of seller behavior and serve as a foundation for our model and estimations.

4.1 Sellers decrease their prices when they do not make any sales

When there is no sales, sellers may become pessimistic about demand, and decrease their prices (e.g., Mason and Välimäki, 2011; Huang et al., 2019). In order to test this hypothesis, we select cases where sellers experience no sales on consecutive days and observe how they adjust their prices over this period. We define such events as the set of "no-sales" data:

$$\mathbb{S}_{\overline{r}}^{\text{no sales}} = \{\mathcal{S}\,(m,j,t) \in \mathbb{N} \mid q_{m,j,t+r} = 0\ , \quad \forall r = 0, \ldots, \overline{r}-1\}\,.$$

To analyze these events, we let $\overline{r} = 7$ and estimate the following regression specification:[16]

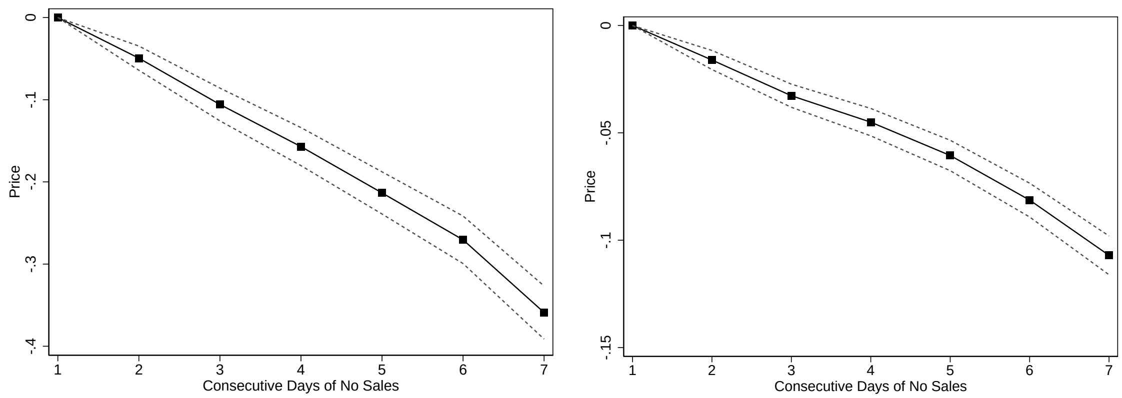
$$y_{s,\tilde{t}} = \sum_{\tau=2}^{7} \gamma_\tau \times \mathbb{1}(\tilde{t} = \tau) + \psi_s + \epsilon_{s,\tilde{t}}, \quad \forall\ s \in \mathbb{S}_7^{\text{no sales}} \wedge \tilde{t} = 1, 2, .., 7, \tag{1}$$

where $\tilde{t}$ denotes the event study time, which ranges from 1 to 7. The outcome of interest for a given event s on day $\tilde{t}$ is denoted by $y_{s,\tilde{t}}$. The indicator variable $\mathbb{1}(\tilde{t} = \tau)$ takes the value of 1 if the observation corresponds to event day τ. To allow for comparisons of the effects on other days, we normalize the coefficient of the day $\tilde{t} = 1$ dummy to zero. The fixed effects for each event are represented by ψ_s.

We investigate how price levels change by examining $p_{s,\tilde{t}}$ as the outcome variable, which provides insights into the overall pricing behavior surrounding the event. In Figure 1, we present the estimates of γ_τ from Equation 1 with $p_{s,\tilde{t}}$ as the dependent variable for third-party sellers and Amazon, respectively, with the price level of the first day normalized to zero. The results indicate that the average price level decreases during consecutive days of no sales. Specifically, after seven days of no sales, third-party sellers' prices decrease by about \$0.1, and Amazon's prices decrease by approximately \$0.3. We note that there is heterogeneity in level changes between Amazon and third-party sellers. Plausibly, Amazon and third-party sellers face different demand parameters and exercise varying degrees of market power, resulting in different responses. The structural model introduced later can explicitly address the limitations of the reduced-form evidence and incorporate the heterogeneity in demand primitives.

[16] Our results remain consistent regardless of the selection of $\overline{r}$. The choice of $\overline{r} = 7$ enhances trend visibility compared to smaller values of $\overline{r}$ and mitigates potential selection concerns associated with larger $\overline{r}$ values.

Figure 1: Price Levels during Consecutive Days of No Sales



(a) Amazon

(b) Third-Party Sellers

Note: Figure 1 shows the estimates from Equation 1 during consecutive days of no sales using the price level as the dependent variables for third-party sellers and Amazon , respectively. The robust standard errors are clustered at the product level.

4.2 Sellers increase their prices depending on their sales

We demonstrate that sellers raise their prices in response to high demand state, which is indicated by their sales. To isolate the effect of sales on prices, we filter the data to include only cases where sellers did not make any sales in the past few days leading up to the event (e.g., day -3 to day -1) and made sales on the event day (day 0). Formally, we define the set of "sales" event as follows:

$$\mathbb{S}_{\bar{r}}^{\text{sales}} = \left\{ \mathcal{S}\left(m, j, t\right) \in \mathbb{N} \mid q_{m,j,t} > 0 \ \wedge \ q_{m,j,t-r} = 0 \,, \quad \forall r = 1, \ldots, \bar{r} \right\}.$$

Our previous analysis (shown in Section 4.1) demonstrates that sellers typically decrease their prices when they do not make any sales, which would naturally result in a downward trend when the data is restricted to periods where sellers have no sales. To control for this effect, we allow for a linear trend in our event study design.

We run the following regression and set $\overline{r} = 3$:

$$y_{s,\tilde{t}} = \sum_{\tau=-3,\tau\neq-1}^{3} \gamma_\tau \times \mathbb{1}(\tilde{t} = \tau) + \eta \times \tilde{t} + \psi_s + \epsilon_{s,\tilde{t}},\ \forall\ s \in \mathbb{S}_3^{\text{sales}} \wedge \tilde{t} \in \{-4, \ldots, 3\}. \quad (2)$$

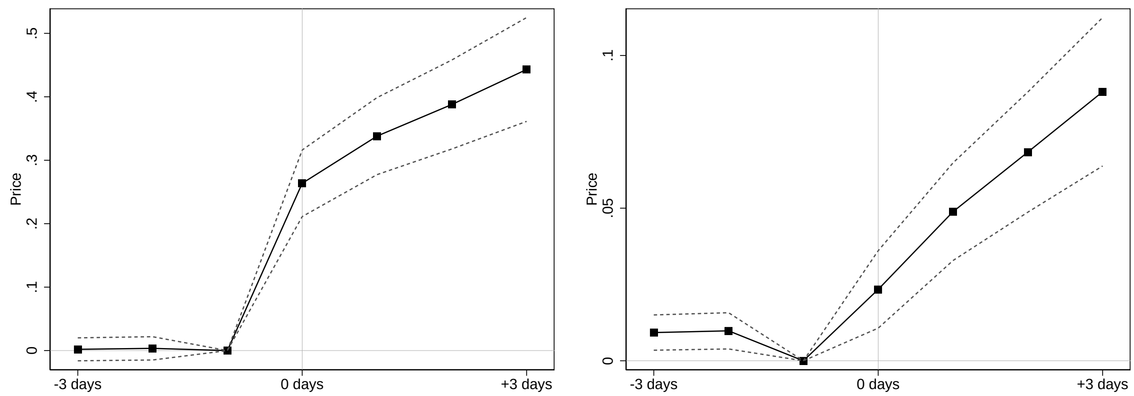
In Equation 2, $\tilde{t}$ represents the number of days relative to a sales event day, ranging from -4 to 3, and $y_{s,\tilde{t}}$ represents the outcome variable for event s on day $\tilde{t}$. The indicator variable $\mathbb{1}(\tilde{t} = \tau)$ takes a value of 1 if the observation is on event day τ. The coefficient of the day -1 dummy is normalized to zero. Note that we include data from $\tilde{t} = -4$ in this regression. This is necessary for the rank condition, as allowing for a linear trend parameter η results in 8 parameters conditional on each event, and we need at least 8 data points for each event to perform inference.

We proceed by examining how price levels change by using $p_{s,\tilde{t}}$ as the outcome variable. Figure 2 shows that the average price for both Amazon and third-party sellers increases following a sales event. Moreover, the price level increase is notably higher for Amazon compared to third-party sellers, with an average increase of around \$0.4 for Amazon and \$0.1 for third-party sellers. Overall, the results are comparable with previous results in Section 4.1. They are also consistent with our survey on Amazon's in-house repricer and leading industry repricers, as discussed in Section 2.

4.3 Heterogeneity in Learning from Competitor's Sales

Sellers can learn from not only from their own sales, but also from their competitors' sales provided that information is available. In such cases, sellers may increase their prices when their own demand is also high relative to the outside option or decrease them when their listings are less competitive than that of their competitors. As discussed in Section 2, competitors' sales is not directly visible on Amazon's website. Amazon, as a dual-role platform, has direct access to this information and can use it when setting prices. On the other hand, third-party sellers usually do not have this information. Consequently, we anticipate that Amazon is more responsive to competitors' sales compared to third-party sellers.

Figure 2: Price Levels After a Sales Event



(a) Amazon

(b) Third-Party Sellers

Note: Figure 2 shows the estimates from Equation 2 using the price level as the dependent variable for third-party sellers and Amazon, respectively, before and after sales events. The vertical line indicates the day of the sales event. The robust standard errors are clustered at the product level.

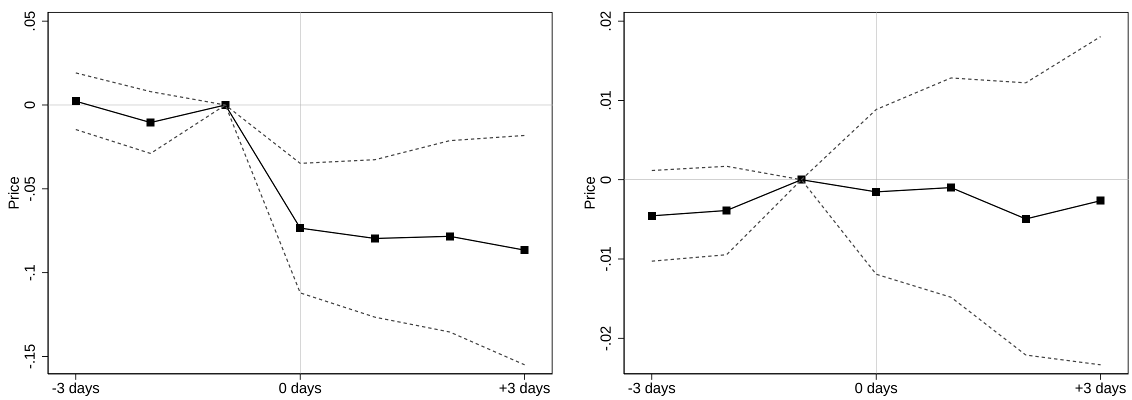
We define a set of "competitor's sales" events for a seller j in a duopoly market, where $-j$ represents the seller's competitor:

$$\mathbb{S}_{\overline{r}}^{\text{c-sales}} = \{\mathcal{S}\,(m, j, t) \in \mathbb{N} \mid \; q_{m,-j,t} > 0 \;\wedge\; q_{m,-j,t-r} = 0\,, \quad \forall r = 1, \ldots, \overline{r}\}\,.$$

Next, we examine Equation 2 using observations from competitor sales events, where $s \in \mathbb{S}_{\overline{r}}^{\text{c-sales}} \wedge \tilde{t} \in \{-4, \ldots, 3\}$. For transparency, we do not control for a competitor's price in this analysis since the competitor's price could be seen as a collider variable. Our emphasis is on the focal seller's price alone, not in relation to a competitor's price. We address price simultaneity more directly in Section 5.

In Figure 3, we present the estimates of γ_τ using the price as the dependent variable. We find that, on average, Amazon tends to lower its price level in response to competitors' sales. The price reduction aligns with the theory that Amazon learns that the competitor's offer is more appealing to consumers and decides to lower its own price.[17] On the other hand, the price levels of third-party sellers remain insignificantly different from zero following a competitor's sale.

Figure 3: Price Levels After a Competitor's Sales Event



(a) Amazon

(b) Third-Party Sellers

Note: Figure 3 displays the estimates from Equation 2 using the price level as the dependent variable for third-party sellers and Amazon, respectively, before and after a competitor's sales event. The vertical line indicates the day of a competitor's sales event. The robust standard errors are clustered at the product level.

4.4 Discussion: Interpreting the Results

We discuss the interpretation of our results and robustness checks that highlight the information mechanism.

Information Disadvantage vs. Sophistication The reduced-form results align with our survey on Amazon's in-house repricer and leading industry repricers, as discussed in Section 2. While third-party sellers can use automated repricers based

[17] This pattern differs from dynamic price competition when there are both perishability and capacity constraints (e.g., Chen and Jeziorski, 2022; Betancourt et al., 2022). In the absence of demand learning, competitors' sales result in supply scarcity and incentivize price increases.

on their own sales, to the best of our knowledge, there are no algorithms that can adjust prices based on competitor-specific sales at the listing level. However, this leads to the question of whether exceptionally large sellers might develop their proprietary systems to price based on competitors' sales. We examine this further in Appendix E. Specifically, we separately examine the response patterns of large and small third-party sellers, defined based on the number of seller ratings. It is plausible that large sellers might be more sophisticated or better equipped with technology. If our reduced-form evidence is driven by seller sophistication or technology rather than information, we would expect to observe that large sellers respond more similarly to Amazon and differ from small sellers. While large third-party sellers exhibit larger responses in price levels compared to small third-party sellers in response to their own sales events, both large and small third-party sellers are unresponsive to the competitor's sales event. This further suggests that the heterogeneity in responses to competitor's sales events is plausibly driven by information rather than the level of sophistication or technology.

Learning Competitor's Sales from Competitor's Prices Another theoretical consideration is whether third-party sellers can learn from Amazon's price changes to infer if Amazon is experiencing sales. However, in practice, this appears unlikely based on empirical data. Among all the price variations presented in Section 3.2, only 13.17% and 12.67% of the price adjustments represent instances where Amazon and third-party sellers, respectively, adjust prices within 3 days following a sales event. Many other factors, such as demand and cost shocks, can trigger price changes. As further supported by our event study results, the empirical relevance of this theoretical consideration is limited within our dataset.

Evidence on Price Adjustment and Autocorrelated Demand In Appendix F, we examine the frequency of price increases and decreases. As indicated in Appendix F.1, sellers tend to decrease prices more frequently than they increase them during consecutive days of no sales. Conversely, in Appendix F.2, we observe that sellers are more likely to increase prices and less likely to decrease them after making sales. Furthermore, in Appendix F.3, we present additional evidence that prices are affected by past sales, and demand exhibits autocorrelation. Consistent with price level responses, Appendix F.4 provides evidence that the frequency of price decreases for Amazon is approximately ten times higher than that of third-party sellers.

5 Model

To motivate the study, we first solve a simple theoretical model of linear demand in a duopoly price competition as outlined in Appendix C. The benchmark considers that Firm 1 possesses information about the demand intercept of Firm 2, while Firm 2 lacks information about Firm 1's demand intercept. Without vertical incentives, specifically referral fees, Firm 1's profit consistently decreases when both firms possess knowledge of each other's demand intercepts. However, the effect on both firms' profit and overall consumer welfare is uncertain in the presence of vertical incentives. When neither firm possesses information about the other's demand intercept, Firm 2 always benefits, and the impact on Firm 1's profit and overall consumer welfare remains theoretically ambiguous. This ambiguity is primarily driven by the tension between intensified competition in some demand states and softened competition in others, and it depends on the value of several parameters in the model. The further discussion is provided in Appendix C.

To better understand the impact of information advantage and resolve the theoretical ambiguity, we develop and estimate an empirical structural model. The structural estimation complements our reduced-form analysis in several aspects. First, it allows for an explicit treatment of price simultaneity. Second, it measures heterogeneous primitives of demand and supply and provides a quantitative assessment of the heterogeneous information and beliefs. Third, it facilitates the quantification of profits and welfare under counterfactual designs for data access.

5.1 Timing

Each duopoly market m at period t constitutes listings from seller $j \in J_{m,t} \subseteq \{1,2\}$. Seller j can have one of two types defined as $a(j) \rightarrow \mathbb{A} = \{0,1\}$ where 0 and 1 indicate third-party sellers and Amazon respectively. We use the notation $-j$ to denote seller j's competing seller in the same market m. The timing of the game is the following:

1. At the onset of time t, the sequence of past demand states is realized and represented as $\mathcal{H}_{m,j,t} = \{h_{m,j,\tau}, h_{m,-j,\tau} : \tau \leq t-1\}$. Using the information as input, the algorithm generates the best predicted demand states $h^{\text{BP}}_{m,j,t}$ and $h^{\text{BP}}_{m,-j,t}$.

2. Sellers form beliefs about their own demand states and competitor's demand states according to $h^{\text{BP}}_{m,j,t}$ and $h^{\text{BP}}_{m,-j,t}$. Sellers observe random draws for utilities and marginal costs of both products, which include additional information that deviates from $h^{\text{BP}}_{m,j,t}$ and $h^{\text{BP}}_{m,-j,t}$.

3. Based on their individual beliefs about current demand states, seller j and its competitor $-j$ engage in a simultaneous move pricing game.

4. Upon arrival, consumers make purchasing decisions according to prices and actual demand states.

5. Sales are determined, and the market proceeds to step 1 for the subsequent period $t+1$.

5.2 Consumer Demand

We outline the demand model below. Let $I_{m,t}$ be the number of consumers with unit demand arriving at period t in market m. The arriving demand $I_{m,t}$ follows a Poisson distribution according to

$$I_{m,t} \sim \text{Poisson}(\lambda_{m,t}).$$

We set $\lambda_{m,t}$ to a large value (i.e., 20), and our estimates are robust with various different values. For market m in period t, the utility of a arriving consumer i choosing alternative $j \in J_{m,t} \cup \{0\}$ is given by

$$u_{i,m,j,t} = \boldsymbol{x}'_{m,j,t}\boldsymbol{\beta} - \alpha \times p_{m,j,t} + \xi_{m,j,t} + v_{i,m,g(j),t}(\rho) + \rho \times \epsilon_{i,m,j,t}, \tag{3}$$

where $\boldsymbol{x}'_{m,j,t}$ is a vector of seller characteristics, such as whether the seller is Amazon and whether the seller uses fulfillment by Amazon services, $p_{m,j,t}$ is the price of seller j in market m at time t, $\xi_{m,j,t}$ is a seller-specific demand shock that changes over time and is important for modeling seller learning (to be discussed in Section 5.2.1), and $v_{i,m,g(j),t}$ is a nested logit random taste that is constant across sellers and differentiates the "inside" good when $g(j) = 1$ from the "outside" option of no purchase when $g(j) = 0$. The variable $\rho \in (0, 1]$ is the intraclass correlation coefficient, which measures the correlation of the unobserved factors influencing the utility of different options within the same nest. $\epsilon_{i,m,j,t}$ is an independently and identically distributed error term (across products and consumers) that follows the logit distribution.

The error structure $v_{i,m,g(j),t}(\rho) + \rho \times \epsilon_{i,m,j,t}$ is assumed to follow the distributional assumption necessary to generate the classic nested logit purchase probability, where the two nests consist of all the inside options and the outside option. When $\rho = 1$, consumers follow a simple logit decision rule on all alternatives, including the outside option. For simplicity, we define

$$\delta_{m,j,t} = \boldsymbol{x}'_{m,j,t}\boldsymbol{\beta} - \alpha \times p_{m,j,t} + \xi_{m,j,t}.$$

The mean utility of the outside option is normalized to zero $\delta_{m,0,t} = 0$. This means that the utility of the outside option, where $j = 0$, follows the error term $\epsilon_{i,m,0,t}$.

5.2.1 Demand State

One important component of the demand function is the alternative-specific intercept $\xi_{m,j,t}$, which represents the unobserved factors that affect the utility of a particular seller. This variable varies over time and is partially unknown to sellers before its realization:

$$\xi_{m,j,t} = \underbrace{\varphi \times h_{m,j,t}}_{\text{Demand state}} + \underbrace{\nu_{m,j,t}}_{\text{Other information}}, \qquad \text{where } \nu_{m,j,t} \sim N\Big(0, \sigma_\nu^2\Big). \tag{4}$$

We assume that the demand state $h_{m,j,t}$ is unobserved by sellers until after time t, meaning that sellers do not know the true demand level for their listings until they observe the actual sales and market outcomes. On the other hand, the researcher, at the time of analysis, can infer true demand states based on realized sales.

The second component in $\nu_{m,j,t}$ denotes the seller-specific error terms with zero mean and a variance of σ_ν^2. The error terms are public information to sellers but not observable by researchers. This approach, combined with our full-solution estimation, directly addresses the endogeneity in prices.

5.2.2 Sales

Based on the utility function in Equation 3 and the error structure, the probability that consumers purchase from seller j in market m at time t, conditional on that they make a purchase, can be expressed as:

$$\mathbb{P}_{m,t}\Big(\text{choosing } j \;\Big|\; \text{choosing } k \in J_{m,t}\Big) = \frac{\exp\left(\dfrac{\delta_{m,j,t}}{\rho}\right)}{\sum_{k \in J_{m,t}} \exp\left(\dfrac{\delta_{m,k,t}}{\rho}\right)}, \quad \forall j \in J_{m,t}.$$

The fraction of consumers who purchase the inside products can be written as:

$$\mathbb{P}_{m,t}\Big(\text{choosing } k \in J_{m,t}\Big) = \frac{(D_{m,t})^\rho}{1 + (D_{m,t})^\rho},$$

where $D_{m,t} = \sum_{k \in J_{m,t}} \exp\left(\frac{\delta_{m,k,t}}{\rho}\right)$. The market share of seller j in market m at period t can be represented as:

$$s_{m,j,t} = \mathbb{P}_{m,t}\Big(\text{choosing } j \;\Big|\; \text{choosing } k \in J_{m,t}\Big) \times \mathbb{P}_{m,t}\Big(\text{choosing } k \in J_{m,t}\Big), \quad \forall j \in J_{m,t}.$$

The quantity of seller j demanded by consumers in market m at time t, $\widetilde{q}_{m,j,t}$, follows a Poisson distribution:

$$\widetilde{q}_{m,j,t} \sim \text{Poisson}\big(\lambda_{m,t} \times s_{m,j,t}\big), \quad \forall j \in J_{m,t}.$$

Note that the quantity demanded may not be equal to the actual sales that were observed in the data, as the sales may be limited by the remaining inventory of the seller. We can express seller j's sales $q_{m,j,t}$ using the following equation:

$$q_{m,j,t} = \min\left\{\widetilde{q}_{m,j,t}\ , \overline{q}_{m,j,t}\right\}, \quad \forall j \in J_{m,t},$$

where $\overline{q}_{m,j,t}$ represents seller j's remaining inventory in market m at period t.

5.3 Seller Belief

We now turn to the supply side model. Based on the available data, each seller j forms estimates of the current-period demand states for the seller itself and for its competitor. If all sellers have complete access to the information and the necessary technology, the best possible estimates for the current-period demand states is denoted as $h^{\text{BP}}_{m,j,t}$, and we discuss the derivation of these estimates in detail in Section 6.1. However, due to the possibility of heterogeneous calibrated belief among sellers, some noise is introduced, and therefore, the sellers' beliefs about own demand state follows:

$$\hat{h}^{\text{Own}}_{m,j,t} = \sum_{I \in \{0,1\}} \left(I - \underbrace{\frac{\exp\left(\phi_0^{\text{Own}} + a(j) \times \phi_1^{\text{Own}}\right)}{1 + \exp\left(\phi_0^{\text{Own}} + a(j) \times \phi_1^{\text{Own}}\right)}}_{\equiv p^{\text{Own}}_{a(j)}} \right) \times \left(I - h^{\text{BP}}_{m,j,t}\right).$$

Here, $p^{\text{Own}}_{a(j)} \in (0,1)$ represents the accuracy of the seller's belief about its own demand relative to the prediction, and depends on their identity $a(j)$. When $p^{\text{Own}}_{a(j)} \to 1$, the seller's belief is perfectly calibrated, meaning their beliefs align perfectly with $h^{\text{BP}}_{m,j,t}$. When $p^{\text{Own}}_{a(j)} \to 0$, the seller's belief is perfectly miscalibrated, meaning their beliefs are completely opposite to $h^{\text{BP}}_{m,j,t}$. When $p^{\text{Own}}_{a(j)} = 0.5$, the seller perceives both states as equally likely. The parameter ϕ^{Own}_0 represents the baseline values for all sellers, while Amazon has an additional parameter value of ϕ^{Own}_1 on top of the baseline.

Similarly, the sellers' beliefs about competitor's state is defined as

$$\hat{h}^{\text{Competitor}}_{m,j,t} = \sum_{I\in\{0,1\}} \left(I - \underbrace{\frac{\exp\left(\phi^{\text{Competitor}}_0 + a(j) \times \phi^{\text{Competitor}}_1\right)}{1 + \exp\left(\phi^{\text{Competitor}}_0 + a(j) \times \phi^{\text{Competitor}}_1\right)}}_{\equiv p^{\text{Competitor}}_{a(j)}} \right) \times \left(I - h^{\text{BP}}_{m,-j,t}\right),$$

where $p^{\text{Competitor}}_{a(j)} \in (0,1)$ represents the accuracy of seller's belief about the competitor's demand state relative to its prediction and depends on its identity $a(j)$. $\phi^{\text{Competitor}}_0$ represents the baseline parameter value that is shared among all sellers, including Amazon. If Amazon has a better calibrated belief about the competitor's demand state, $\phi^{\text{Competitor}}_1$, which is specific to Amazon, is expected to be positive in addition to the baseline.

The seller makes decisions to maximize their expected profits based on their beliefs. Similar to Equation 4, the seller's belief is updated according to the following:

$$\hat{\xi}^j_{m,j,t} = \varphi \times \hat{h}^{\text{Own}}_{m,j,t} + \nu_{m,j,t}. \tag{5}$$

Sellers also form beliefs about their competitor's demand intercept, denoted as $\tilde{\xi}^j_{m,-j,t}$, based on past observations and pricing decisions. The beliefs are updated according to the following equation:

$$\hat{\xi}^j_{m,-j,t} = \varphi \times \hat{h}^{\text{Competitor}}_{m,j,t} + \nu_{m,-j,t}. \tag{6}$$

Our model of seller beliefs is parsimonious. Given our main objective is to quantify the dependence between prices and past sales data, we choose to directly model this dependence without imposing a specific learning or belief formation process.

5.4 Profit Function and Pricing Decision

The sellers in the market can be classified into two types: third-party sellers $a(j) = 0$ and Amazon $a(j) = 1$. Among third-party sellers, there are variations in their demand functions and costs depending on whether they use Amazon's fulfillment service. Amazon earns profits from its own sales and also charges fees to third-party sellers for using its platform to sell their products, including a referral fee, which is a percentage of the selling price. This referral fee is considered a cost for the third-party sellers. The profit function can be expressed as follows:

$$\Pi_{m,j,t} = \begin{cases} \big[\,(1 - r_m) \times p_{m,j,t} - mc_{m,j,t}\big] \times q_{m,j,t}, & \text{if } a(j) = 0, \\ r_m \times p_{m,-j,t} \times q_{m,-j,t} + \big(p_{m,j,t} - mc_{m,j,t}\big) \times q_{m,j,t}, & \text{if } a(j) = 1, \end{cases} \tag{7}$$

and seller j's subjective profit function depends on its beliefs:

$$\hat{\Pi}^j_{m,j,t} = \begin{cases} \big[\,(1 - r_m) \times p_{m,j,t} - mc_{m,j,t}\big] \times \hat{q}^j_{m,j,t}, & \text{if } a(j) = 0, \\ r_m \times p_{m,-j,t} \times \hat{q}^j_{m,-j,t} + \big(p_{m,j,t} - mc_{m,j,t}\big) \times \hat{q}^j_{m,j,t}, & \text{if } a(j) = 1, \end{cases}$$

where r_m is the referral fee rate charged by Amazon as a proportion of the product price, The expected quantity sold by the seller j and its competitors in market m at time t are represented by $\hat{q}_{m,j,t}$, respectively. The marginal cost of seller j is determined by a combination of its characteristics vector $\boldsymbol{w}$ and a cost shock $\varsigma_{m,j,t}$ that follows a normal distribution. The cost function is expressed as:

$$mc_{m,j,t} = \boldsymbol{w}^{'}_{m,j,t}\boldsymbol{\omega} + \varsigma_{m,j,t}, \text{ where } \varsigma_{m,j,t} \sim N\Big(0, \sigma^2_{\varsigma}\Big).$$

The supply shocks, denoted by $\boldsymbol{\varsigma}$, have a variance of σ^2_{ς} and a mean of zero. Although these shocks are observable to sellers, they are unobserved by researchers. Sellers engage in a simultaneous move pricing game to make their pricing decisions. This simultaneity aligns with the institutional setup, as nearly all leading providers of automated pricing offer instant repricing (see the repricing frequency survey in Appendix B).

Model Discussion To conclude, we discuss several assumptions embedded within the supply model. First, FBA sellers incur an additional per-unit fulfillment fee on top of the standard referral fee. The per-unit fee is constant for each product and unlike referral fee, it does not change depending on prices. Therefore, the per-unit fee is implicitly captured in the marginal cost for FBA sellers. We assume that Amazon's cost of fulfilling the FBA order is equivalent to the revenue derived from fees, and therefore, the model does not consider Amazon's profit in the FBA segment.

Second, following canonical supply models, we do not explicitly model a "decision" or cost of repricing. If a seller does not reprice, our model would rationalize this behavior using a marginal cost $\varsigma_{m,j,t}$. We assume that, on average, heuristics in pricing algorithms align with profit-maximizing incentives (see Section 2 for discussion). Moreover, since sales are stochastic and difficult to control precisely, they are unlikely to be the primary coordination tool among competitors. Nevertheless, high-frequency prices and sales measures may be valuable for studying collusive algorithmic pricing, particularly among third-party sellers (see literature cited in Section 1.1).

Lastly, Amazon has many tools for information design aimed at consumers, such as product recommendations, search ranking, and the buy box. These tools may interact with seller types (i.e., Amazon vs. non-Amazon) as shown by the current empirical literature cited in Section 1.1. We focus on price competition and information design for sellers and treat other effects as fixed in our model and in the counterfactual analysis. The study of their interactions is left for future work.

6 Estimation and Identification

We take the structural model into the data and obtain model estimates. The demand primitives includes both static preference parameters and dynamic preference parameters:

$$\boldsymbol{\theta}^D = \Big\{ \underbrace{\{\boldsymbol{\alpha}, \rho, \beta\}}_{\text{Static preference parameters}}, \underbrace{\{\varphi, \sigma_\nu\}}_{\text{Dynamic preference parameters}} \Big\}.$$

The seller primitives include the marginal cost and the parameters related to how sellers learn from past sales:

$$\boldsymbol{\theta}^S = \Big\{ \{\boldsymbol{\omega}, \sigma_\varsigma\}, \underbrace{\{\boldsymbol{\phi}^{\text{Own}}, \boldsymbol{\phi}^{\text{Competitor}}\}}_{\text{Seller learning parameters}} \Big\}.$$

6.1 Estimation

Below, we provide a detailed description of the estimation procedure. We then discuss the identification of key parameters in the estimation process.

Use Data Neighboring Repricing Events We construct a balanced panel to capture within-product price variations in a concise 3-day timeframe surrounding repricing events. These events are characterized by price adjustments taking place on day t, while no adjustments occur on the preceding or subsequent days. Formally, the set of repricing events is defined as follows:

$$\mathbb{S}^{\text{repricing}} = \Big\{ \mathcal{S}\,(m, j, t) \in \mathbb{N} \mid p_{m,j,t} \neq p_{m,j,t-1} \,\wedge\, p_{m,j,t+r} = p_{m,j,t+r-1}\,, \quad \forall r \in \{-1, 1\} \Big\}.$$

The dataset for estimation includes the sales and prices of the repricing seller and their competitor for both the day before and the day after the repricing event, denoted as $\tilde{t} = -1$ and $\tilde{t} = 1$, respectively. To focus on within-product price variation, we normalize price by the mean price of each product and set the average price across the sample to be \$70.

We outline the method used to estimate the demand and supply primitives below.

Model Specification We discuss how we model demand and supply heterogeneties. Depending on the types of sellers (i.e., Amazon [AMZ], Fulfillment by Amazon [FBA], and Fulfillment by Merchant [FBM]) present in a market, we can have five possible market structures:

$$\mathcal{M}(m) \in \{\text{AMZ-vs-FBM}, \text{AMZ-vs-FBA}, \text{FBA-vs-FBM}, \text{FBA-vs-FBA}, \text{FBM-vs-FBM}\}.$$

We allow each of these market types to have different utility intercepts, marginal costs, variances of demand shocks, and variances of cost shocks. Note that the seller types are ordered and some types of markets have asymmetric seller types. We further allow for within-market heterogeneity by permitting this asymmetry in both utility intercepts and marginal costs.

In the utility function, we include $\beta_{\text{AMZ,FBA}}$ and $\beta_{\text{AMZ,FBM}}$ to denote the utility intercept of Amazon relative to FBA or FBM sellers, respectively, and $\beta_{\text{FBA,FBM}}$ to denote the utility intercept of FBA sellers relative to FBM sellers. In the cost function, we include $\omega_{\text{AMZ,FBA}}$ and $\omega_{\text{AMZ,FBM}}$ to measure the average difference in cost between Amazon and FBA or FBM sellers, respectively, and $\omega_{\text{FBA,FBM}}$ to measure the average difference in cost between FBA and FBM sellers. We also include a cost shock, denoted as ω_{Post}, for the period after the repricing event.

True Demand States We use the following regression to estimate the demand states that explain the sales after incorporating the effect of prices and product-seller fixed effects:[18]

$$q_{m,j,t} = \eta \times p_{m,j,t} + \boldsymbol{\psi}_{m,j} + \epsilon_{m,j,t}. \tag{8}$$

We include the product-seller fixed effects, denoted by $\boldsymbol{\psi}_{m,j}$, to account for unobserved and observed heterogeneity across each product listing. By doing so, the demand state is defined by the variation in demand within a listing. To remove sales variation shifted by prices, we use the log of inventory as an instrument for the price. The inventory level acts as a cost shifter since sellers may face higher costs when they have excess inventory due to storage costs (see Appendix G).

[18] See Kalouptsidi (2014) and Jeon (2022) for a similar approach to constructing demand state when modeling demand uncertainty.

We discretize the true demand states for later structural estimation. We define the demand states $h_{m,j,t}$ as a higher state ($h_{m,j,t} = 1$) when the residuals from Equation 8 (i.e., $\hat{\epsilon}_{m,j,t}$) are greater than the median of the residuals. Alternatively, the demand states are considered to be in a lower state ($h_{m,j,t} = 0$) when the residuals are less than or equal to the median of the residuals. All medians are computed within the same market structure and seller types. Formally, we have:

$$h_{m,j,t} = \begin{cases} 1, & \text{if } \hat{\epsilon}_{m,j,t} > \text{median}\left(\hat{\epsilon}_{m',j',t} \mid \mathcal{M}(m') = \mathcal{M}(m), a(j') = a(j)\right), \\ 0, & \text{if } \hat{\epsilon}_{m,j,t} \leq \text{median}\left(\hat{\epsilon}_{m',j',t} \mid \mathcal{M}(m') = \mathcal{M}(m), a(j') = a(j)\right) \end{cases}$$

Best Predicted Demand States We aggregate all past sales information into a simple statistic called ex-ante best predicted demand state. Note that true demand states are estimated using ex-post sales, which are unavailable to sellers when setting price. To estimate the ex-ante best predicted demand states, we use past demand states of the sellers' own sales and those of their competitors:

$$h_{m,j,t} = \sum_{\tau=-T}^{-1} \kappa_{\tau,j} \times h_{m,j,\tau} + \sum_{\tau=-T}^{-1} \kappa_{\tau,-j} \times h_{m,-j,\tau} + \epsilon_{m,j,t}. \tag{9}$$

In practice, we choose $T = -7$, which corresponds to using demand states from the previous week. We find that extending the timeframe beyond a week has minimal effect on prediction accuracy. Similarly, we discretize the best predicted demand states. We define the best predicted demand state to be a higher state ($h^{\text{BP}}_{m,j,t} = 1$) when the predicted value $\hat{h}_{m,j,t}$ in Equation 9 is greater than its median, and a lower state ($h^{\text{BP}}_{m,j,t} = 0$) when the predicted value is less than or equal to its median. As before, all medians are computed within the same market structure and seller types. Formally, we have:

$$h^{\text{BP}}_{m,j,t} = \begin{cases} 1, & \text{if } \hat{h}_{m,j,t} > \text{median}\left(\hat{h}_{m',j',t} \mid \mathcal{M}(m') = \mathcal{M}(m), a(j') = a(j)\right), \\ 0, & \text{if } \hat{h}_{m,j,t} \leq \text{median}\left(\hat{h}_{m',j',t} \mid \mathcal{M}(m') = \mathcal{M}(m), a(j') = a(j)\right). \end{cases}$$

Using this approach, we find that the ex-post true demand state $h_{m,j,t}$ and the ex-ante best predicted one $h^{\text{BP}}_{m,j,t}$ coincide 79.21% of the time, indicating that past sales data plays a significant role in predicting the current demand state. It is important to note that we use the terms "true demand states" and "best predicted demand states" to refer to demand information extracted from historical sales. However, these terms should not be interpreted as literal truths or ultimate predictions. In practice, we also consider and allow additional demand information through the error term $\nu_{m,j,t}$.

Inner Loop Let the vector of demand and supply primitives be $\boldsymbol{\theta} = \left\{\boldsymbol{\theta}^D, \boldsymbol{\theta}^S\right\}$. As the demand shocks ν and cost shocks ς are unobservable in a given observation, we use a Hermite-Gauss quadrature of order 3 to simulate four normally distributed variables representing demand and cost shocks for both the seller and its competitor. This results in a total of 81 possible joint distributions of demand and cost shocks. To solve for equilibrium outcomes, including prices and sales, we simultaneously solve for demand and supply for each joint shock distribution. We approximate these outcomes using the Hermite-Gauss quadrature weights assigned to each joint distribution.

There are 16 possible states resulting from the combination of a seller's and their competitor's demand states (h_j and h_{-j}, respectively), as well as their respective best predicted demand states (h_j^{BP} and h_{-j}^{BP}). We solve for equilibrium in each of the five market structures, both before and after the day of repricing, across the 16 demand states and $3^4 = 81$ joint distributions of demand shocks and cost shocks. Thus, we solve equilibria for 12,960 pricing games during each iteration.

Outer Loop In the outer loop, we use a simulation estimator to match predicted moments from the model to observed moments in the data one day (i.e., $\tilde{t} \in \{-1, 1\}$) neighboring repricing events $\mathbb{S}^{\text{repricing}}$. We utilize the following moments.

1. Own price and sales (320 moments):

$$E\left[p_{m,j,\tilde{t}} \times \boldsymbol{z}_{m,j,\tilde{t}}\right] \text{ and } E\left[q_{m,j,\tilde{t}} \times \boldsymbol{z}_{m,j,\tilde{t}}\right],$$

2. Higher order price moments: the square of own price and the interaction with

competitor's price (320 moments):

$$E\left[p^2_{m,j,\tilde{t}} \times \boldsymbol{z}_{m,j,\tilde{t}}\right] \text{ and } E\left[p_{m,j,\tilde{t}} \times p_{d,-j,\tilde{t}} \times \boldsymbol{z}_{m,j,\tilde{t}}\right],$$

3. The interaction of own price with own sales and competitor's sales (320 moments):

$$E\left[p_{m,j,\tilde{t}} \times q_{m,j,\tilde{t}} \times \boldsymbol{z}_{m,j,\tilde{t}}\right] \text{ and}$$
$$E\left[p_{m,j,\tilde{t}} \times q_{d,-j,\tilde{t}} \times \boldsymbol{z}_{m,j,\tilde{t}}\right],$$

The vector of 160 indicators $\boldsymbol{z}_{m,j,\tilde{t}}$ is defined as:

$$\boldsymbol{z}_{m,j,\tilde{t}} = \underbrace{\boldsymbol{M}_{m,j,\tilde{t}}}_{\text{market structures } (5\times1)} \otimes \underbrace{\boldsymbol{H}_{m,j,\tilde{t}}}_{\text{predicted and true demand states } (16\times1)} \otimes \underbrace{\boldsymbol{T}_{m,j,\tilde{t}}}_{\text{before and after repricing } (2\times1)}.$$

See Appendix H for a detailed breakdown of these vectors.

We have 34 model primitives. Denote the vector of empirical moments as $\boldsymbol{\phi}$ and a vector of simulated moments as $\boldsymbol{\phi}(\boldsymbol{\theta})$. Given a weighting matrix $\boldsymbol{W}$, the outer loop minimizes the following objective function:

$$f(\boldsymbol{\theta}) = \left[\hat{\boldsymbol{\phi}}(\boldsymbol{\theta}) - \boldsymbol{\phi}\right]' \boldsymbol{W} \left[\hat{\boldsymbol{\phi}}(\boldsymbol{\theta}) - \boldsymbol{\phi}\right].$$

In practice, we use the weighting matrix $\boldsymbol{W}$ to normalize the scales of the moments. Standard errors are obtained using a sandwich formula.

6.2 Identification

We begin our discussion by examining the identification of price elasticity. To do this, we take advantage of the high-frequency data available to us. We construct a balanced panel that captures within-product price variations within a concise 3-day timeframe. Our assumption is that, during this brief period, the underlying demand remains stable. We then attribute high-frequency fluctuations in prices to two supply-side factors: a common shock in marginal cost (ω_{Post}) as well as variations in beliefs. Following the literature, we acknowledge that sellers may possess superior information compared to researchers by allowing for both demand and supply shocks that are observable to sellers but not to the researchers. Given the sparsity of sales data, it is not feasible to directly invert demand shocks. Instead, we employ a full solution approach that simultaneously solves the demand and supply models, taking into account the unobserved demand and supply shocks.

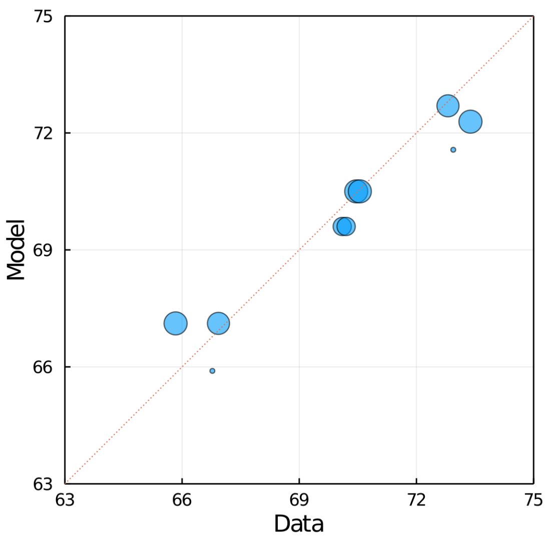
To identify the variance of demand shocks, we focus on the covariance between price and sales. Specifically, for a given price, observed sales may deviate from what is expected based on price elasticity. We attribute the distribution of these residual demands to the distribution of demand shocks. In contrast, supply shocks are independent of residual demand but still contribute to price shifts. After determining the distribution of demand shocks, we can use the variance of supply shocks to rationalize the distribution of prices, which we capture by employing the square of price as a moment.

We infer sellers' beliefs from their pricing levels (e.g., Aguirregabiria, 2021) specifically, the learning parameters can be identified through the dependence of prices on past sales. In particular, a higher value of ϕ^{Own} indicates that seller j sets higher prices when $h_j^{\text{BP}} = 1$ compared to when $h_j^{\text{BP}} = 0$. Meanwhile, a higher value of $\phi^{\text{Competitor}}$ suggests that seller j adjusts its price in response to h_{-j}^{BP}. Based on our reduced-form findings, which demonstrate that Amazon sets its prices according to competitor demand state while third-party sellers do not, we expect to observe a higher value of $\phi_1^{\text{Competitor}}$ and a lower value of $\phi_0^{\text{Competitor}}$.

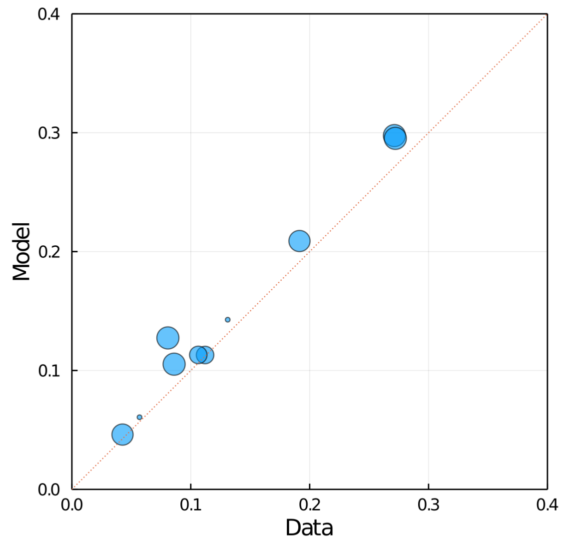
7 Results

In Figure 4, we assess the fit of our model by comparing the empirical and predicted prices and sales. Our model adequately explains the sales and prices observed in the data.

Figure 4: Moment Fit



(a) Price Moment



(b) Sales Moment

Note: Figure 4 evaluates the moment fit, with the x-axis representing the empirical moment and the y-axis representing the model prediction. The moments are aggregated at the market-type×seller level. The fit of the price moment is shown in Figure 4a, while the fit of the sales moment is presented in Figure 4b.

Columns 1 to 3 of Table 3 display the demand estimates. The average own-price elasticity is -18.96, a magnitude consistent with previous literature (e.g., Ellison and Ellison, 2009; Dinerstein et al., 2018). This high own-price elasticity primarily stems from the intense competition between two sellers offering the same product on a shared product page. The cross-price elasticity, at 18.41, is high as expected, but remains lower than what a classic Bertrand model would predict. The average market price elasticity, at -5.28, reflects the level of competition against the outside option, which includes all other products on the platform as a subset. Overall, competition across listings within the same product is notably more intense than that against the outside option.

Consumers derive greater utility from purchasing products from Amazon or FBA sellers compared to FBM sellers. The estimated premiums over FBM sellers are \$7.15 and \$6.12 respectively. In AMZ-vs-FBA markets, consumers experience an estimated gain of \$5.31 when buying from Amazon rather than from an FBA seller. The estimate of φ on the true demand state reveals that demand indeed fluctuates over time, as indicated by its positive and significant value.

Columns 4 to 6 of Table 3 display the supply estimates. In the same market, both Amazon and FBA sellers exhibit lower marginal costs than FBM sellers, with cost advantages of \$1.81 and \$8.77, respectively. In the AMZ-vs-FBA market, Amazon's marginal cost is \$14.72 higher than that of competing FBA sellers. The parameters ϕ^{own} and $\phi^{\text{competitor}}$ represent the degree to which sellers' beliefs align with the best predicted demand states h_j^{BP} and h_{-j}^{BP} based on their own and competitors' past sales. Based on the ϕ^{own} parameter estimates, both Amazon and third-party sellers have well-calibrated beliefs on their own demand states. Their accuracy levels are similar, though Amazon's is slightly better, aligning more closely with its own h^{BP} at $p_1^{\text{own}} = 0.98$, compared to third-party sellers' $p_0^{\text{own}} = 0.92$.

Regarding belief of competitors' demand states, Amazon's belief closely aligns with the competitor's h^{BP}, as indicated by $p_1^{\text{competitor}} = 0.94$. In contrast, third-party sellers' beliefs closely resemble a random guess, with $p_0^{\text{competitor}} = 0.52$.

We also find significant heterogeneity across various market structures in the variances of random effects on demand shocks, denoted by σ_ν, and on cost shocks, denoted by σ_ς.

Overall, our structural estimates are consistent with the institutional details presented in Section 2 and the reduced-form analysis in Section 4. Our structural estimations further rationalize the reduced-form patterns we observed in Section 4 with a rational framework of price competition and heterogeneous beliefs. Specifically, our estimates indicate that Amazon possesses a better-calibrated belief parameter, especially in learning from competitors' sales. This finding highlights the information advantage held by Amazon and the consequent disadvantage faced by third-party sellers.

Table 3: Parameter Estimates

Demand Parameters	Estimate	SE	Supply Parameters	Estimate	SE
α	7.97	(0.12)	$\omega_{\text{AMZ,FBM}}$	-1.81	(0.15)
$\beta_{\text{AMZ,FBM}}$	0.57	(0.02)	$\omega_{\text{AMZ,FBA}}$	14.72	(0.20)
$\beta_{\text{AMZ,FBA}}$	0.42	(0.02)	$\omega_{\text{FBA,FBM}}$	-8.77	(0.23)
$\beta_{\text{FBA,FBM}}$	0.49	(0.03)	ω_{Post}	2.27	(0.04)
Intraclass Correlation ρ	0.15	(0.01)	ϕ_0^{own}: Baseline	2.38	(0.15)
φ	3.07	(0.03)	ϕ_1^{own}: Amazon	1.67	(0.37)
Market Structure FE			$\phi_0^{\text{competitor}}$: Baseline	0.09	(0.01)
AMZ-vs-FBM	-3.53	(0.09)	$\phi_1^{\text{competitor}}$: Amazon	2.63	(0.14)
FBA-vs-FBM	-1.95	(0.08)	Market Structure FE		
FBA-vs-FBA	-0.41	(0.08)	AMZ-vs-FBM	55.34	(0.13)
FBM-vs-FBM	-5.20	(0.26)	FBA-vs-FBM	57.38	(0.16)
AMZ-vs-FBA	-1.77	(0.08)	FBA-vs-FBA	53.16	(0.14)
σ_ν			FBM-vs-FBM	51.87	(0.12)
AMZ-vs-FBM	0.56	(0.06)	AMZ-vs-FBA	47.89	(0.14)
FBA-vs-FBM	0.20	(0.18)	σ_ς		
FBA-vs-FBA	0.06	(0.28)	AMZ-vs-FBM	16.52	(0.10)
FBM-vs-FBM	2.66	(0.14)	FBA-vs-FBM	2.99	(1.25)
AMZ-vs-FBA	0.14	(0.06)	FBA-vs-FBA	0.58	(2.24)
			FBM-vs-FBM	15.01	(0.19)
			AMZ-vs-FBA	2.69	(1.02)

Note: Table 3 displays the parameter estimates. Standard errors in parentheses are clustered at the nest level. The price coefficient α is scaled by a factor of 100. We allow four parameters to differ across five market structure: (1) AMZ-vs-FBM, where Amazon competes with an FBM seller, (2) FBA-vs-FBM, where an FBA seller competes with an FBM seller, (3) FBA-vs-FBA, where two FBA sellers compete with each other, (4) FBM-vs-FBM, where two FBM sellers compete with each other, and (5) AMZ-vs-FBA, where Amazon competes with an FBA seller. We convert each repricing adjustment into a price-increasing event, so the cost shock ω_{Post} is positive.

7.1 Counterfactual Simulations

We conduct several counterfactual analyses to examine the impact of information advantage on price competition, specifically: (1) preventing Amazon from using marketplace data for its own retail business and (2) sharing marketplace data with third-party sellers.[19] To highlight the role of the vertical integrated market structure, we also vary the referral fee and examine how it mediates the effect. For each scenario, we calculate seller's profit based on Equation 7. We also calculate consumer welfare as the following:

$$\mathrm{CW} = \sum_{m,t} \lambda_{m,t} \times \frac{\log\left(\exp\left(\rho \times \log\left(D_{m,t}\right)\right) + 1\right)}{\alpha}.$$

Social welfare is the sum of surplus and profits across consumers, sellers, and Amazon. In markets where Amazon is not present as a seller, we include its revenue from referral fees in the calculation of social welfare as the following:

$$\mathrm{SW} = \mathrm{CW} + \sum_{m,t} \sum_{j \in J_{m,t}} \Pi_{m,j,t} + \underbrace{\left(1 - \max_{j \in J_{m,t}} a(j)\right) \times r_m \times p_{m,j,t}}_{\text{Referral fee in third-party-only markets}}.$$

7.1.1 Restricting Amazon's Information Advantage

In the first counterfactual scenario, Amazon loses its information advantage because it is prevented from using marketplace data for its pricing decision. This is similar to the settlement reached between Amazon and the EU, which concerns Amazon's internal structural data separation between its platform business and retail business.[20]

To simulate this, we set Amazon's learning parameter $p_1^{\text{competitor}}$ equal to that of third-party sellers $p_0^{\text{competitor}}$. The results in Columns 1 to 3 of Table 4 show that,

[19]These counterfactuals also correspond to proposals that Amazon put forward in response to EU antitrust investigations (i.e., `https://www.reuters.com/technology/amazon-offers-share-data-boost-rivals-dodge-eu-antitrust-fines-sources-2022-06-13/`).

[20]The final settlement can be found at `https://ec.europa.eu/commission/presscorner/detail/en/ip_22_7777`. See Section 2 for more discussion.

under the current 15% referral fee and restricted information access for Amazon, Amazon's sales decrease by 0.64%, while profits drop by 0.30%. With the information disadvantage removed, third-party sellers experience a 2.63% increase in sales and a 2.36% profit gain. Consumer welfare and social welfare improve by 0.48% and 0.33%, respectively, due to reduced prices and increased quantity. When Amazon has less insider information, it faces a disadvantage. Yet, the gains for consumers and third-party sellers outweigh Amazon's loss, resulting in a more competitive environment.

We further emphasize the role of vertical market structure in our findings by examining the scenario where Amazon charges a 0% referral fee, eliminating vertical incentives. This scenario addresses information design in the context of structural separation, where Amazon, as the retailer, does not collect fees from sellers (i.e., separating the platform-Amazon from the retailer-Amazon).

Columns 4 to 6 of Table 4 display the changes in equilibrium outcomes for sales, profits, and welfare when Amazon has no information advantage, compared to the case with Amazon's information advantage under a 0% referral fee. We observe that Amazon's sales and profits drop significantly by 4.61% and 4.68%, respectively. Conversely, third-party sellers experience substantial increases in sales and profits, at 7.27% and 7.26%, respectively, when Amazon loses its information advantage. The average increase in consumer welfare and social welfare across the two markets is approximately 0.48% and 0.61%, respectively. The larger effect under a 0% referral fee compared to a 15% referral fee is intriguing. It suggests that the presence of a referral fee could potentially deter Amazon from using its informational advantage to the detriment of third-party sellers.

7.1.2 Sharing Information with Third-Party Sellers

In the second counterfactual, we explore a scenario in which Amazon shares its marketplace data with third-party sellers, thereby eliminating their information disadvantage. We simulate this situation by granting third-party sellers the same access

Table 4: Counterfactual Analysis of Restricting Amazon's Information Advantage

	15% Referral Fee			0% Referral Fee		
	AMZ-vs-FBM	AMZ-vs-FBA	AMZ-Average	AMZ-vs-FBM	AMZ-vs-FBA	AMZ-Average
Panel A: Equilibrium Outcomes						
%Δ Amazon Price	-0.27	0.34	-0.22	-0.43	1.97	-0.21
%Δ Amazon Sales	0.08	-7.86	-0.64	-1.69	-33.93	-4.61
%Δ 3rd-Party Price	-0.13	-0.04	-0.12	-0.18	-0.63	-0.22
%Δ 3rd-Party Sales	2.37	5.24	2.63	6.19	18.06	7.27
Panel B: Profit and Welfare Change						
%Δ Amazon Profit	-0.30	-0.34	-0.30	-2.18	-29.73	-4.68
%Δ 3rd-Party Profit	1.94	6.52	2.36	5.99	19.99	7.26
%Δ Consumer Welfare	0.58	-0.58	0.48	0.46	0.74	0.48
%Δ Social Welfare	0.30	0.69	0.33	0.33	3.42	0.61

Note: Table 4 presents the counterfactual analysis when Amazon's information is restricted. Panel A shows the percentage changes in equilibrium prices and sales, while Panel B quantifies the changes in profits and welfare for Amazon, third-party sellers, and consumers. AMZ-vs-FBM represents the market where Amazon competes with an FBM seller. AMZ-vs-FBA represents the market where Amazon competes with an FBA seller. AMZ-Average represents the average change weighted by the number of observations across both markets, AMZ-vs-FBM and AMZ-vs-FBA. Columns 1 to 3 correspond to the current 15% referral fee, and Columns 4 to 6 correspond to the case when the referral fee is reduced to 0%.

to information and learning parameters as Amazon. Specifically, we let $p_0^{\text{competitor}} = p_1^{\text{competitor}}$. We assume that third-party sellers can use competitors' sales information to infer competitors' demand state, similar to Amazon. This assumption is supported by our model estimates, which demonstrate that third-party sellers have beliefs about their own demand state comparable to Amazon's when using their own sales information.

Columns 1 to 3 of Table 5 present the results for this counterfactual, assuming the current 15% referral fee remains in place. Our findings indicate a substantial increase of 14.91% in sales for third-party sellers, while Amazon's sales experience a 4.08% decline. However, due to the revenue generated from the referral fee, Amazon's profits increase by 0.45% instead of decreasing. Moreover, information sharing leads to a more substantial rise in consumer welfare and social welfare on average, with estimated increases of 2.08% and 1.65%, respectively. With increased information being shared, prices better reflect demand states, leading to a slight drop in average prices but a significant increase in overall output. Consumers and third-party sellers benefit greatly, and Amazon indirectly gains from market expansion through commission fees, leading to a Pareto improvement.

Similarly, Columns 4 to 6 of Table 5 present the results for the counterfactual where third-party sellers possess equal information as Amazon, compared to the scenario in which they are information disadvantaged under a 0% referral fee. In this scenario, Amazon's average profit decreases by 5.66%, driven by the loss of first-party sales without generating any referral revenue. This, in turn, incentivizes Amazon to compete more aggressively with third-party sellers. Consequently, third-party sellers experiences a smaller profit increase of 3.12%, while there is a notable increase in consumer welfare by 5.29%.

Table 5: Counterfactual Analysis of Sharing Information with Third-Party Sellers

	15% Referral Fee			0% Referral Fee		
	AMZ-vs-FBM	AMZ-vs-FBA	AMZ-Average	AMZ-vs-FBM	AMZ-vs-FBA	AMZ-Average
Panel A: Equilibrium Outcomes						
%Δ Amazon Price	0.01	-0.17	-0.01	-0.93	-0.90	-0.93
%Δ Amazon Sales	-1.80	-27.01	-4.08	-0.58	-24.53	-2.75
%Δ 3rd-Party Price	-0.12	-0.02	-0.11	-1.02	-0.62	-0.98
%Δ 3rd-Party Sales	12.51	39.00	14.91	16.76	33.43	18.27
Panel B: Profit and Welfare Change						
%Δ Amazon Profit	0.05	4.48	0.45	-3.33	-29.01	-5.66
%Δ 3rd-Party Profit	2.84	19.87	4.38	2.25	11.82	3.12
%Δ Consumer Welfare	1.31	9.82	2.08	4.38	14.45	5.29
%Δ Social Welfare	0.88	9.41	1.65	1.64	8.87	2.29

Note: Table 5 presents the counterfactual analysis when Amazon shares its information with third-party sellers. Panel A shows the percentage changes in equilibrium prices and sales, while Panel B quantifies the changes in profits and welfare for Amazon, third-party sellers, and consumers. AMZ-vs-FBM represents the market where Amazon competes with an FBM seller. AMZ-vs-FBA represents the market where Amazon competes with an FBA seller. AMZ-Average represents the average change weighted by the number of observations across both markets, AMZ-vs-FBM and AMZ-vs-FBA. Columns 1 to 3 correspond to the current 15% referral fee, and Columns 4 to 6 correspond to the case when the referral fee is reduced to 0%.

7.1.3 Information Sharing in Other Markets

Markets with Only Third-Party Sellers Information sharing among third-party sellers can considerably influence markets where these sellers vie with one another, and Amazon does not function as a seller. Given no prior information advantage among third-party sellers, we essentially compare two scenarios where competitors have equal information access, but the volume of market information varies. The counterfactual results for this scenario are explained in detail in Appendix I. Amazon, as the platform owner, earns profits from referral fees without engaging in direct market participation. In contrast to markets that involve Amazon, vertical incentives hold less significance because the referral fee primarily serves as a static modifier to sellers' marginal costs.

Our analysis indicates that in more symmetric markets, such as FBA-vs-FBA and FBM-vs-FBM, information sharing has pro-competitive effects. This leads to a redistribution of welfare from sellers to consumers, and Amazon benefits from increased fees due to higher transaction volume. However, in more asymmetric markets like the FBA-vs-FBM market, information sharing has anti-competitive effects, resulting in higher prices and reduced transactions at equilibrium. Both consumers and Amazon experience a decline in welfare, and social welfare is also reduced.

All Types of Markets In summary, when considering all market types in our dataset and using their empirical weights, the implementation of an information sharing scheme leads to a significant overall improvement in both consumer welfare and social welfare, with respective increases of 4.02% and 2.39% when the referral fee is set at the current level of 15%.

To put the numbers into context, each number represents the daily welfare change for an average product in the data, with the average consumer welfare and social welfare being \$3.19 and \$6.72 per product each day. For every one million products similar to those in our dataset that are sold over a one-year period, the resulting increase in consumer welfare and social welfare would amount to approximately \$46.81

million (\$3.19 $\times$ 365 $\times$ 4.02%) and \$58.62 million (\$6.72 $\times$ 365 $\times$ 2.39%), respectively.

8 Conclusion

This research examines the impact of asymmetric data access on price competition in a vertically integrated platform. We find that while individual sellers gain insights into demand through their own sales, the platform owner has access to competitor sales data, which enables it to better understand demand and set prices. We estimate a structural model to quantify the importance of information advantage. Counterfactual analysis shows that the design of data access has a significant impact on equilibrium outcomes. We find that granting third-party sellers with the same level of data access as Amazon results in a 1.65% increase in social welfare. Amazon's profit also rises by 0.45%, driven by the growth in referral fees that offsets for the decrease in first-party sales, leading to a Pareto improvement. On the other hand, eliminating Amazon's information advantage results in a 0.33% increase in social welfare. However, this scenario leads to a 0.30% decline in Amazon's profit due to reduced first-party sales.

There are numerous types of data with many other use in various contexts. For example, sales data can inform entry and exit decisions, whereas search data can be used to understand consumer preference. While the data is non-rivalry, the market is not. Our findings highlight the equilibrium effect of data access in a competitive setting, particularly with regards to the impact of past sales data on price competition. However, we believe that this topic has broader implications, and provides a basis for further research.

References

Aguirregabiria, V. (2021). Identification of Firms' Beliefs in Structural Models of Market Competition. *Canadian Journal of Economics/Revue canadienne d'économique*, 54(1):5–33.

Aguirregabiria, V. and Jeon, J. (2020). Firms' Beliefs and Learning: Models, Identification, and Empirical Evidence. *Review of Industrial Organization*, 56(2):203–235.

Asker, J., Fershtman, C., Jeon, J., and Pakes, A. (2020). A Computational Framework for Analyzing Dynamic Auctions: The Market Impact of Information Sharing. *The RAND Journal of Economics*, 51(3):805–839.

Asker, J., Fershtman, C., and Pakes, A. (2023). The Impact of Artificial Intelligence Design on Pricing. *Journal of Economics & Management Strategy*.

Ater, I. and Rigbi, O. (2023). Price Transparency, Media, and Informative Advertising. *American Economic Journal: Microeconomics*, 15(1):1–29.

Bajari, P., Chernozhukov, V., Hortaçsu, A., and Suzuki, J. (2019). The Impact of Big Data on Firm Performance: An Empirical Investigation. In *AEA Papers and Proceedings*, volume 109, pages 33–37.

Bergemann, D., Bonatti, A., and Gan, T. (2022). The Economics of Social Data. *The RAND Journal of Economics*, 53(2):263–296.

Betancourt, J. M., Hortaçsu, A., Oery, A., and Williams, K. R. (2022). Dynamic Price Competition: Theory and Evidence From Airline Markets. Technical report, National Bureau of Economic Research.

Brown, Z. Y. and MacKay, A. (2023). Competition in Pricing Algorithms. *American Economic Journal: Microeconomics*, 15(2):109–156.

Byrne, D. P., de Roos, N., Lewis, M. S., Marx, L. M., and Wu, X. A. (2023). Asymmetric Information Sharing in Oligopoly: A Natural Experiment in Retail Gasoline. *Available at SSRN 4513435*.

Calvano, E., Calzolari, G., Denicolo, V., and Pastorello, S. (2020). Artificial Intelligence, Algorithmic Pricing, and Collusion. *American Economic Review*, 110(10):3267–3297.

Chen, N. and Jeziorski, P. (2022). Consequences of Dynamic Pricing in Competitive Airline Markets. *Available at SSRN 4285718*.

Chen, N. and Tsai, H.-T. (2019). Steering via Algorithmic Recommendations. *Available at SSRN 3500407*.

Clark, R., Assad, S., Ershov, D., and Xu, L. (2023). Algorithmic Pricing and Competition: Empirical Evidence From the German Retail Gasoline Market.

Crémer, J., de Montjoye, Y., and Schweitzer, H. (2019). Competition Policy for the Digital Era. Technical report, European Commission.

Dinerstein, M., Einav, L., Levin, J., and Sundaresan, N. (2018). Consumer Price Search and Platform Design in Internet Commerce. *American Economic Review*, 108(7):1820–59.

Doraszelski, U., Lewis, G., and Pakes, A. (2018). Just Starting Out: Learning and Equilibrium in a New Market. *American Economic Review*, 108(3):565–615.

Ellison, G. and Ellison, S. F. (2009). Search, Obfuscation, and Price Elasticities on the Internet. *Econometrica*, 77(2):427–452.

Farronato, C., Fradkin, A., and MacKay, A. (2023). Self-Preferencing at Amazon: Evidence From Search Rankings.

Fudenberg, D. and Villas-Boas, J. M. (2006). Behavior-Based Price Discrimination and Customer Recognition. *Handbook on economics and information systems*, 1:377–436.

Gal-Or, E. (1985). Information Sharing in Oligopoly. *Econometrica: Journal of the Econometric Society*, pages 329–343.

Goldfarb, A. and Xiao, M. (2011). Who Thinks About the Competition? Managerial Ability and Strategic Entry in US Local Telephone Markets. *American Economic Review*, 101(7):3130–3161.

Gutierrez, G. (2021). The Welfare Consequences of Regulating Amazon. Technical report, mimeo, New York University.

Hagiu, A., Teh, T.-H., and Wright, J. (2022). Should Platforms Be Allowed to Sell on Their Own Marketplaces? *The RAND Journal of Economics*, 53(2):297–327.

Hart, O., Tirole, J., Carlton, D. W., and Williamson, O. E. (1990). Vertical Integration and Market Foreclosure. *Brookings Papers on Economic Activity. Microeconomics*, 1990:205–286.

He, S., Hollenbeck, B., and Proserpio, D. (2022). The Market for Fake Reviews. *Marketing Science*, 41(5):896–921.

Hortaçsu, A., Luco, F., Puller, S. L., and Zhu, D. (2019). Does Strategic Ability Affect Efficiency? Evidence from Electricity Markets. *American Economic Review*, 109(12):4302–42.

Huang, G., Luo, H., and Xia, J. (2019). Invest in Information or Wing It? A Model of Dynamic Pricing with Seller Learning. *Management Science*, 65(12):5556–5583.

Huang, Y., Ellickson, P. B., and Lovett, M. J. (2022). Learning to Set Prices. *Journal of Marketing Research*, 59(2):411–434.

Jeon, J. (2022). Learning and Investment under Demand Uncertainty in Container Shipping. *The RAND Journal of Economics*, 53(1):226–259.

Jin, G. Z. and Leslie, P. (2003). The Effect of Information on Product Quality: Evidence From Restaurant Hygiene Grade Cards. *The Quarterly Journal of Economics*, 118(2):409–451.

Jones, C. I. and Tonetti, C. (2020). Nonrivalry and the Economics of Data. *American Economic Review*, 110(9):2819–58.

Kalouptsidi, M. (2014). Time to Build and Fluctuations in Bulk Shipping. *American Economic Review*, 104(2):564–608.

Klein, T. (2021). Autonomous Algorithmic Collusion: Q-Learning Under Sequential Pricing. *The RAND Journal of Economics*, 52(3):538–558.

Kühn, K.-U. and Vives, X. (1995). *Information Exchanges Among Firms and Their Impact on Competition*. European Communities.

Lam, H. T. (2021). Platform Search Design and Market Power. *Job Market Paper, Northwestern University*.

Lee, K. H. and Musolff, L. (2021). Entry into Two-Sided Markets Shaped by Platform-Guided Search. Technical report, mimeo, Princeton University.

Li, L. (1985). Cournot Oligopoly With Information Sharing. *The RAND Journal of Economics*, pages 521–536.

Luco, F. (2019). Who Benefits From Information Disclosure? The Case of Retail Gasoline. *American Economic Journal: Microeconomics*, 11(2):277–305.

Madsen, E. and Vellodi, N. (2022). Insider Imitation. *Available at SSRN 3832712*.

Mason, R. and Välimäki, J. (2011). Learning About the Arrival of Sales. *Journal of Economic Theory*, 146(4):1699–1711.

Raval, D. (2022). Steering in one click: Platform self-preferencing in the amazon buy box. Technical report, Working Paper.

Reimers, I. and Waldfogel, J. (2023). A Framework for Detection, Measurement, and Welfare Analysis of Platform Bias. Technical report, National Bureau of Economic Research.

Rey, P. and Tirole, J. (2007). A Primer on Foreclosure. *Handbook of Industrial Organization*, 3:2145–2220.

Rhodes, A., Johnson, J. P., and Wildenbeest, M. (2023). Platform Design When Sellers Use Pricing Algorithms. *Econometrica*.

Rossi, F. and Chintagunta, P. K. (2016). Price Transparency and Retail Prices: Evidence From Fuel Price Signs in the Italian Highway System. *Journal of Marketing Research*, 53(3):407–423.

Tadelis, S. and Zettelmeyer, F. (2015). Information Disclosure as a Matching Mechanism: Theory and Evidence From a Field Experiment. *American Economic Review*, 105(2):886–905.

Varian, H. (2018). Artificial Intelligence, Economics, and Industrial Organization. In *The Economics of Artificial Intelligence: An Agenda*, pages 399–419. University of Chicago Press.

Vives, X. (1984). Duopoly Information Equilibrium: Cournot and Bertrand. *Journal of Economic Theory*, 34(1):71–94.

Waters, M. (2023). Amazon Is Displaying Sales Data In Search—Here's What To Know. Accessed: 15 Oct 2023.

Zhu, F. and Liu, Q. (2018). Competing With Complementors: An Empirical Look at Amazon. Com. *Strategic Management Journal*, 39(10):2618–2642.

Appendix

A Disclosure of Sales Information

The disclosure of data is a major design decision actively made by platforms. As for the disclosure of sales data, there is no consensus in the industry. Table A.1 presents a listing of 19 e-commerce websites. This listing is accompanied by the respective primary operational regions of these websites, along with an indicator whether these platforms disclose sales data or not.

Figure A.1a illustrates an example of sales data disclosure on an e-commerce platform. Other platforms that make sales data open typically follow a similar format. Figure A.1b, sourced from Waters (2023), is an example of Amazon's public disclosure of sales data. The incidents started around March 2023. As of October 2023, they still appear to be very selective and random. Moreover, due to differences in platform design, Amazon's disclosure is at the product level instead of the seller-listing level, given the presence of multiple competing listings for a single product on Amazon. Nevertheless, if implemented on a larger scale, this could be seen as a significant step in Amazon's sharing of sales information.[21]

B Automated Pricing Services

We conduct a survey of leading providers of automated pricing services in the market.[22] For each of these providers, we summarize the data used in their technology in Colunms 2 and 3 of Table A.2. We find that, in addition to considering competitors' prices, most automated pricing service providers offer algorithms that adjust pricing based on their own sales, often referred to as "Sales Velocity Repricing." However,

[21] Amazon offers a program called Amazon Brand Analytics, which aims to provide "the right data" to brand owners (see https://sell.amazon.com/blog/brand-analytics).

[22] See https://www.marketplacepulse.com/landscape.

Table A.1: Disclosure of Sales Data by E-commerce Websites

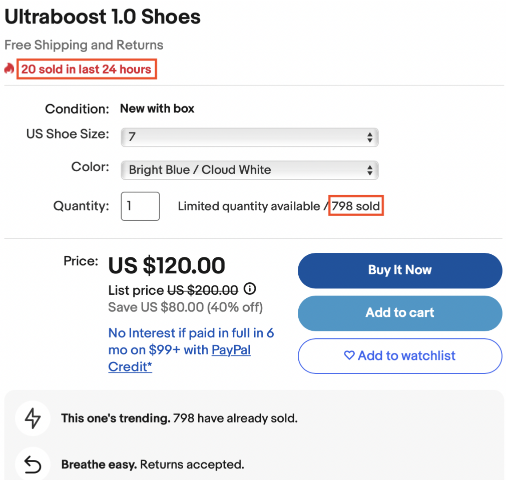
Website	**Region**	**Disclose Sales Data**
AliExpress	International	Yes
eBay	US	Yes
Walmart	US	No
Best Buy	US	No
Target	US	No
Zalando	Europe	No
Allegro	Europe	Yes
Cdiscount	Europe	No
Yahoo! Japan Shopping	Japan	No
ZOZOTOWN	Japan	No
Mercari	Japan	Yes
Lazada	Southeast Asia	Yes
Shopee	Southeast Asia	Yes
Zalora	Southeast Asia	No
Qoo10	Southeast Asia	Yes
Taobao	China	Yes
JD.com	China	No
Pinduoduo	China	Yes
Suning	China	No

Note: Table A.1 lists the 19 e-commerce websites, their primary operating regions, and whether these websites disclose sales measures or not.

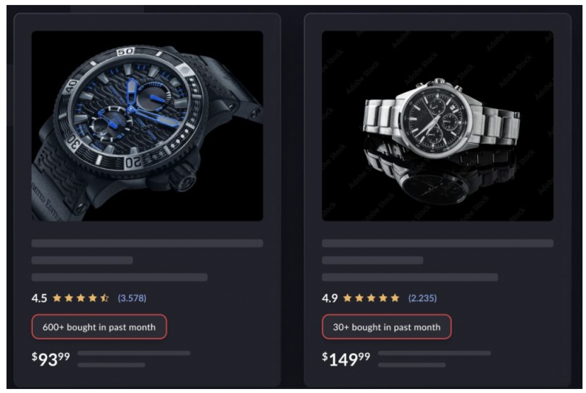
none of the companies claim to have the ability to set prices based on competitors' sales.

In Column 4 of Table A.2, we also survey the repricing frequency that each provider offers. Except for one provider whose main service is to provide A/B testing, the remaining providers claim that they reprice instantly. In practice, this means that repricing is usually done in seconds, 2 to 3 minutes, or as fast as Amazon allows.

Figure A.1: Public Sales Data on e-Commerce Platform: Examples



(a) eBay



(b) Amazon (Recent Experimentation)

Note: Figure A.1a displays open sales data on eBay, while Figure A.1b displays open sales data on Amazon, which is currently in the experimental phase.

Table A.2: Automated Pricing Services

Company	Competitor's Price	Own Sales	Instant Repricing
Amazon	✓	✓	✓
Aura	✓		✓
Bqtools	✓	✓	✓
Channelmax	✓	✓	✓
Eva	✓	✓	✓
Feedvisor360	✓	✓	✓
Flashpricer	✓	✓	✓
Informed	✓	✓	✓
Priceloop	✓	✓	
Repricer.com	✓	✓	✓
Seller Snap	✓	✓	✓

Note: Table A.2 summarizes each of the leading providers of automated pricing services, the data they claim to use in their pricing technology in Columns 2 and 3, and whether they claim to have instant repricing in Column 4.

C Theoretical Model

The demand system for two firms, 1 and 2, in a duopoly is described as follows:

$$q_1 = A_1 - k_1 A_2 - a_1 \times p_1 + b_1 \times p_2, \quad q_2 = A_2 - k_2 A_1 - a_2 \times p_2 + b_2 \times p_1.$$

Assuming that firm 1 charges a referral fee $0 < r < 1$ from firm 2, the profit functions of both firms are presented below.

$$\pi_1 = p_1 \times q_1 + r \times p_2 \times q_2, \quad \pi_2 = (1 - r) \times p_2 \times q_2,$$

The parameters $\boldsymbol{k}, \boldsymbol{a}, \boldsymbol{b}$ and r are positive deterministic numbers and it is assumed that a firm's own price has a greater impact on demand than its competitor's price ($a_1 > b_1$ and $a_2 > b_2$). Moreover, we assume a firm's price has a greater impact on its own demand than its competitors' demand ($a_2 > b_1$ and $a_1 > b_2$). We simplify the model by assuming that firms have zero marginal costs. Price and quantity must be non-negative.

The firm-specific demand shocks, A_1 and A_2, are positive and distributed according to cumulative distribution functions $F(A_1)$ and $F(A_2)$, with corresponding density functions $f(A_1)$ and $f(A_2)$, respectively. The parameters k_1 and k_2 are less than 1 and η_{A_j} denotes the mean of $A_j, \quad \forall j \in (1, 2)$. These distributions are common knowledge to both firms.

As a benchmark, we consider the current situation where Firm 1 has an information advantage over Firm 2, knowing both A_1 and A_2, while Firm 2 only knows A_2. We then explore two scenarios:

1. Perfect Information: Both firms have full information on both A_1 and A_2.

2. Limited Information: Each firm is only aware of its own A value, with Firm 1 knowing A_1 and Firm 2 knowing A_2.

C.1 Benchmark: Information Advantage

We begin by analyzing the model under the information advantage benchmark, where Firm 1 has knowledge of both A_1 and A_2, while Firm 2 is aware only of A_2. The prices and sales, p and q, respectively, are as follows:

$$p_1^A = \frac{2A_1a_2 - \eta_{A_1}k_2(b_1 + rb_2) + A_2(b_1 - 2a_2k_1 + rb_2)}{4a_1a_2 - b_2(b_1 + rb_2)} \geq 0,$$
$$p_2^A = \frac{-2a_1A_2 - A_1b_2 + A_2b_2k_1 + 2a_1\eta_{A_1}k_2}{-4a_1a_2 + b_2(b_1 + rb_2)} \geq 0.$$

$$q_1^A = \frac{b_2(-A_1 + A_2k_1)rb_2 + a_1(2A_1a_2 + A_2(b_1 - 2a_2k_1 - rb_2) + \eta_{A_1}k_2(-b_1 + rb_2))}{4a_1a_2 - b_2(b_1 + rb_2)} \geq 0,$$
$$q_2^A = \frac{a_2(2a_1A_2 + A_1b_2 - A_2b_2k_1 - 2a_1\eta_{A_1}k_2)}{4a_1a_2 - b_2(b_1 + rb_2)} - (A_1 - \eta_{A_1})k_2 \geq 0.$$

The expected profits for Firm 1 and Firm 2 can be defined as follows:

$$\mathbb{E}[\Pi_1^A] = \int_{\mathcal{A}_2}\int_{\mathcal{A}_1} p_1^A q_1^A dF(A_1)dF(A_2), \quad \mathbb{E}[\Pi_2^A] = \int_{\mathcal{A}_2}\int_{\mathcal{A}_1} p_2^A q_2^A dF(A_1)dF(A_2).$$

The consumer welfare for each firm can be expressed as follows:

$$\begin{aligned}
CW_1^A &= \int_{\mathcal{A}_2}\int_{\mathcal{A}_1} \frac{1}{2}\left(\frac{A_1 - k_1A_2 + b_1p_2^A}{a_1} - p_1^A\right) q_1^A dF(A_1)dF(A_2) \\
&= \int_{\mathcal{A}_2}\int_{\mathcal{A}_1} \frac{1}{2}\left(\frac{q_1^A + a_1p_1^A}{a_1} - p_1^A\right) q_1^A dF(A_1)dF(A_2) = \frac{1}{2a_1}\int_{\mathcal{A}_2}\int_{\mathcal{A}_1} (q_1^A)^2 dF(A_1)dF(A_2), \\
CW_2^A &= \frac{1}{2a_2}\int_{\mathcal{A}_2}\int_{\mathcal{A}_1} (q_2^A)^2 dF(A_1)dF(A_2).
\end{aligned}$$

The overall consumer welfare is defined as the sum of the consumer surplus of Firm 1 and Firm 2: $CW^A \coloneqq CW_1^A + CW_2^A$.

C.2 Perfect Information

In the perfect information scenario, both firms have knowledge of A_1 and A_2, and prices can be expressed as follows:

$$p_1^P = \frac{2A_1a_2 - A_1k_2(b_1+rb_2) + A_2(b_1 - 2a_2k_1 + rb_2)}{4a_1a_2 - b_2(b_1+rb_2)} = p_1^A + \underbrace{\frac{(\eta_{A_1} - A_1)k_2(b_1+rb_2)}{4a_1a_2 - b_2(b_1+rb_2)}}_{\text{Increased price competition}}$$

$$:= p_1^A + \Delta p_1^{A,P},$$

$$p_2^P = \frac{-2a_1A_2 - A_1b_2 + A_2b_2k_1 + 2a_1A_1k_2}{-4a_1a_2 + b_2(b_1+rb_2)} = p_2^A + \underbrace{\frac{2a_1(\eta_{A_1} - A_1)k_2}{4a_1a_2 - b_2(b_1+rb_2)}}_{\text{Improved pricing response to competitors' information}} := p_2^A + \Delta p_2^{A,P}.$$

Compared to the information advantage scenario, Firm 2 will now undercut competitors when in a higher demand state (higher A_1) and raise prices when in a lower demand state (lower A_1). This change in Firm 2's strategy will result in increased price competition for Firm 1.

Quantities are given by the following and have a simple relationship with the prices:

$$q_1^P = q_1^A + \frac{a_1(\eta_{A_1} - A_1)k_2(b_1 - rb_2)}{4a_1a_2 - b_2(b_1+rb_2)} := q_1^A + \Delta q_1^{A,P},$$
$$q_2^P = q_2^A + \frac{(A_1 - \eta_{A_1})k_2(2a_1a_2 - b_2(b_1+rb_2))}{4a_1a_2 - b_2(b_1+rb_2)} := q_2^A + \Delta q_2^{A,P}.$$

Expected Profits The expected profit for Firm 1 can be written as follows:

$$\begin{aligned}
\mathbb{E}[\Pi_1^P] &= \int_{\mathcal{A}_2}\int_{\mathcal{A}_1} (p_1^A + \Delta p_1^{A,P})(q_1^A + \Delta q_1^{A,P}) dF(A_1)dF(A_2) \\
&= \mathbb{E}[\Pi_1^A] + \int_{\mathcal{A}_1}\int_{\mathcal{A}_2} \left(p_1^A \Delta q_1^{A,P} + \Delta p_1^{A,P}\Delta q_2^{A,P}\right) dF(A_2)dF(A_1) \\
&= \mathbb{E}[\Pi_1^A] + \frac{k_2\left(\overbrace{rb_2^2(rb_2 + b_1 - ra_1k_2)}^{\text{Effects of referral fee}} - \overbrace{a_1b_1(4a_2 - b_1k_2)}^{\text{Loss of information advantage}}\right)}{(4a_1a_2 - b_2(b_1+rb_2))^2}\left(\int_{\mathcal{A}_1}(A_1 - \eta_{A_1})A_1 dF(A_1)\right).
\end{aligned}$$

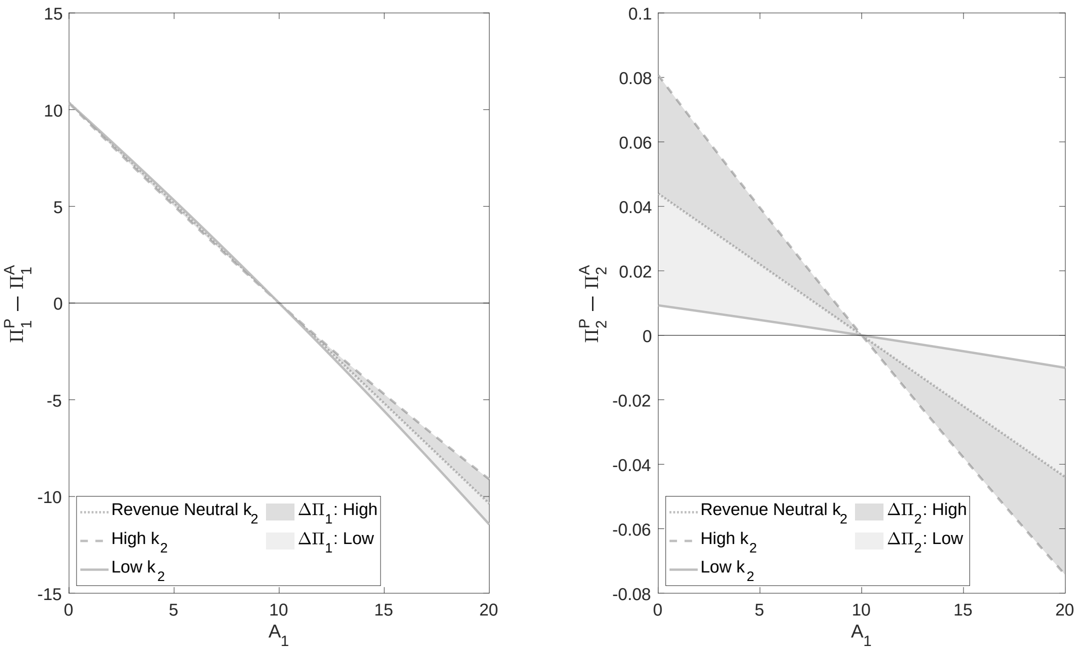
When there are no vertical incentives ($r = 0$), Firm 1's profit decreases after losing its information advantage over Firm 2, as shown in Figure A.2a. However, when vertical incentives exist ($0 < r < 1$), the outcome depends on two factors: the loss of direct sales due to the absence of the information advantage and the potential gain from referral fee revenue. If the potential gain outweighs the loss, Firm 1 may end up better off. The revenue-neutral k_2 is the value at which $\Pi_1^P = \Pi_1^A$, and the shaded region between the profit and revenue-neutral lines in Figure A.2a represents the profit change for other k_2 values. The expected profit for Firm 2 can be written as follows:

$$\begin{aligned}
\mathbb{E}[\Pi_2^P] &= \mathbb{E}[\Pi_2^A] + \int_{\mathcal{A}_1}\int_{\mathcal{A}_2}\left(p_2^A \Delta q_2^{A,P} + \Delta p_2^{A,P} a_2 p_2^P\right) dF(A_2)dF(A_1) \\
&= \mathbb{E}[\Pi_2^A] + \frac{k_2\left(\overbrace{4a_1^2 a_2 k_2}^{\text{Gain from better pricing}} - \overbrace{b_2^2(b_1 + rb_2)}^{\text{Loss from intensified competition}}\right)}{(4a_1a_2 - b_2(b_1+rb_2))^2}\left(\int_{\mathcal{A}_1}(A_1 - \eta_{A_1})A_1 dF(A_1)\right) \\
&> \mathbb{E}[\Pi_2^A] \iff k_2 > \frac{b_2^2(b_1+rb_2)}{4a_1^2a_2}.
\end{aligned}$$

Perfect information generally benefits Firm 2, particularly when k_2 values are large. However, compared to the information advantage case, its profit can still decrease. Figure A.2b illustrates a numerical example where the revenue-neutral k_2 is $\frac{b_2^2(b_1+rb_2)}{4a_1^2a_2}$, and $\Pi_2^P = \Pi_2^A$ at this point. When k_2 is low, Firm 2's profit increases if A_1 is greater than the average and decreases if A_1 is less than the average. However, the loss outweighs the gain, leading to lower profits for Firm 2.

The intuition is as follows: When Firm 2 knows A_1, it charges a lower price when A_1 is high and a higher price when A_1 is low, intensifying price competition when A_1 is high and softening it when A_1 is low. If the increased competition outweighs the benefit of the imperfect information, it can result in a decline in Firm 2's profit.

Figure A.2: Effect of Variable k_2 on Profits



(a) $\Pi_1^L - \Pi_1^A$ (b) $\Pi_2^L - \Pi_2^A$

Note: Figure A.2 displays a simulation with parameters $a_1 = 12.1, a_2 = 2.3, b_1 = 10, b_2 = 1.3, k_1 = 0.2, A_2 = 10$, and $E[A_1] = 10$, and $r = 15\%$. The concept of a "budget neutral" k refers to the value that equalizes the firm's profit to that of the information advantage case.

Consumer Welfare The consumer welfare for Firm 1 can be written as

$$\begin{aligned}
CW_1^P &= \frac{1}{2a_1}\int_{\mathcal{A}_2}\int_{\mathcal{A}_1}(q_1^P)^2 dF(A_1)dF(A_2) = \frac{1}{2a_1}\int_{\mathcal{A}_2}\int_{\mathcal{A}_1}(q_1^A + \Delta q_1^{A,P})^2 dF(A_1)dF(A_2) \\
&= CW_1^A + \frac{1}{2a_1}\int_{\mathcal{A}_1}\int_{\mathcal{A}_2}(q_1^A + q_1^P)\Delta q_1^{A,P} dF(A_2)dF(A_1) \\
&= CW_1^A + \frac{k_2(b_1 - rb_2)(2rb_2^2 - 4a_1a_2 + a_1k_2(b_1 - rb_2))}{2(4a_1a_2 - (b_1 + rb_2)b_2)^2}\int_{\mathcal{A}_1}(A_1 - \eta_{A_1})A_1 dF(A_1) \leq CW_1^A.
\end{aligned}$$

If there are no vertical incentives ($r = 0$), the loss of information advantage by Firm 1 leads to a decrease in the welfare of its consumers. However, the direction of this effect becomes ambiguous in the presence of vertical incentives ($0 < r < 1$). As for

Firm 2:

$$CW_2^P = CW_2^A + \frac{1}{2a_2}\int_{\mathcal{A}_2}\int_{\mathcal{A}_1}(q_2^A + q_2^P)\Delta q_2^{A,P} dF(A_1)dF(A_2)$$
$$= CW_2^A + \frac{k_2(2a_1a_2 - b_2(b_1 + rb_2))(2a_2(b_2 - 3a_1k_2) + b_2k_2(b_1 + rb_2))}{2a_2(4a_1a_2 - b_1b_2)^2}$$
$$\int_{\mathcal{A}_1}(A_1 - \eta_{A_1})A_1 dF(A_1).$$

The welfare of Firm 2's consumers decreases as long as k_2 is sufficiently large, aligning with the fact that a larger k_2 results in decreased profits for Firm 2. The overall consumer welfare decreases when k_2 is sufficiently large.

$$CW^P \geq CW^A \iff k_2 < \frac{2a_2(2a_1a_2(b_1 - b_2(1+r)) + b_2^2(b_1 - b_1r + b_2r(1+r)))}{a_1a_2(b_1^2 + rb_2(8b_2 + rb_2) + b_1(b_2(-8+r) + rb_2)) - 12a_1^2a_2^2 - b_2^2(b_1 + rb_2)^2}.$$

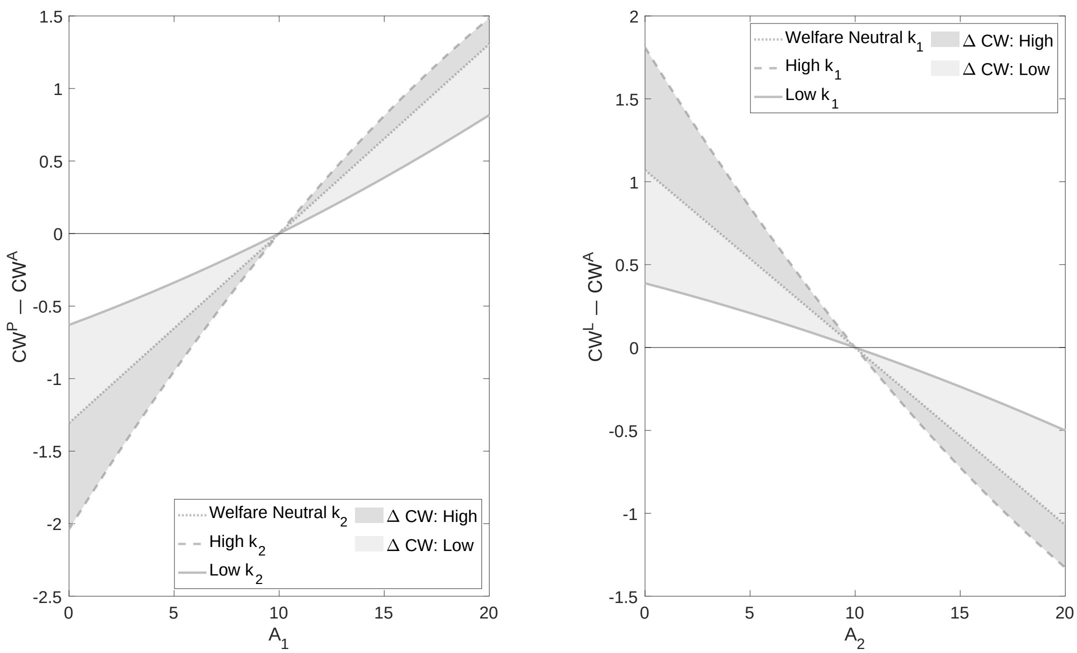
As shown in Figure A.3a, consumers benefit when A_1 is high because competition is intensified, but they suffer more when A_1 is low because competition is softened. However, consumer welfare can actually increase under perfect information compared to the information advantage scenario. When k_2 is small and Firm 2's cross-elasticity of demand b_2 is much larger than Firm 1's b_1, it becomes more likely that the gains from intensified competition when A_1 is high for Firm 2's consumers slightly outweigh the losses from softened competition when A_1 is low.

C.3 Limited Information

In the limited information scenario, each firm has knowledge only of their own A, with Firm 1 observing only A_1 and Firm 2 observing only A_2. The prices under this situation can be represented as follows:

$$p_1^L = \frac{2A_1a_2 - 2a_2\eta_{A_2}k_1 + (A_2 - \eta_{A_1}k_2)(b_1 + rb_2)}{4a_1a_2 - b_2(b_1 + rb_2)} = p_1^A + \underbrace{\frac{2a_2(A_2 - \eta_{A_2})k_1}{4a_1a_2 - b_2(b_1 + rb_2)}}_{\text{Inferior pricing without competitor information}}$$
$$:= p_1^A + \Delta p_1^{A,L},$$
$$p_2^L = \frac{2a_1A_2 + A_1b_2 - b_2\eta_{A_2}k_1 - 2a_1\eta_{A_1}k_2}{4a_1a_2 - b_2(b_1 + rb_2)} = p_2^A + \underbrace{\frac{b_2(A_2 - \eta_{A_2})k_1}{4a_1a_2 - b_2(b_1 + rb_2)}}_{\text{Soften price competition}} := p_2^A + \Delta p_2^{A,L}.$$

Figure A.3: Effect of Variables k_1 and k_2 on Consumer Welfare



(a) $CW^P - CW^A$ (b) $CW^L - CW^A$

Note: Figure A.3 displays a simulation with parameters $a_1 = a_2 = 8, E[A_1] = E[A_2] = 10$, and $r = 15\%$. Additionally, Figure A.3a has $b_1 = 1$, $b_2 = 5$, $k_1 = 0.2$, and $A_2 = 10$, while Figure A.3b has $b_1 = 5$, $b_2 = 1$, $k_2 = 0.05$, and $A_1 = 10$. The concept of a "welfare neutral" k refers to the value that equalizes the overall consumer welfare to that of the information advantage case.

Firm 1 no longer undercuts Firm 2 when its demand state is high (High A_2) but raises prices when Firm 2 has a low demand state (Low A_2). This change reduces price competition, which has a direct impact on Firm 2.

The quantity meets the following conditions:

$$q_1^L = q_1^A + \frac{(2a_1a_2 - b_1b_2)(\eta_{A_2} - A_2)k_1}{4a_1a_2 - b_2(b_1 + rb_2)} := q_1^A + \Delta q_1^{A,L},$$
$$q_2^L = q_2^A + \frac{a_2b_2(A_2 - \eta_{A_2})k_1}{4a_1a_2 - b_2(b_1 + rb_2)} := q_2^A + \Delta q_2^{A,L}.$$

Expected Profits Firm 1's expected profit in the scenario of limited information is as follows

$$\begin{aligned}\mathbb{E}[\Pi_1^L] &= \mathbb{E}[\Pi_1^A] + \int_{\mathcal{A}_2}\int_{\mathcal{A}_1}(p_1^A \Delta q_1^{A,L} + \Delta p_1^{A,L} q_1^L) dF(A_1) dF(A_2) \\ &= \mathbb{E}[\Pi_1^A] + \frac{k_1\Big(\overbrace{b_2(b_1^2 + b_1 r b_2 + 2a_2 k_1 r b_2)}^{\text{Gain from soften competition}} - \overbrace{4a_1 a_2(a_2 k_1 + r b_2)}^{\text{Loss without competitor information}}\Big)}{(4a_1 a_2 - b_2(b_1 + r b_2))^2} \\ &\quad \left(\int_{\mathcal{A}_2}(A_2 - \eta_{A_2}) A_2 dF(A_2)\right).\end{aligned}$$

For most large k_1 values, Firm 1 is generally worse off than in the information advantage case. However, there are scenarios where Firm 1 could still benefit. In the simulation shown in Figure A.4, the revenue-neutral k_1 is $\frac{b_1^2 b_2 - 4a_1 a_2 r b_2 + b_1 b_2 r b_2}{4a_1 a_2^2 - 2a_2 b_2 r b_2}$, resulting in $\Pi_1^L = \Pi_1^A$. When k_1 is high, Firm 1's profit decreases when A_2 is less than the average and increases when A_2 is larger than the average. However, the gain dominates the loss, leading to higher profits for Firm 1.

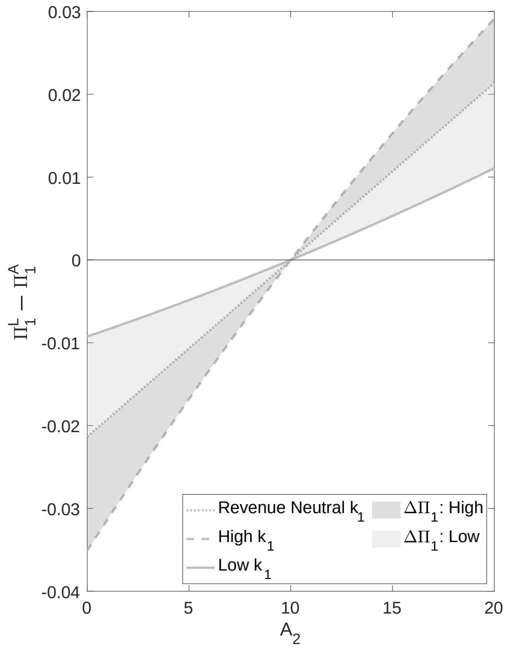
Since Firm 1 does not know A_2, it charges higher prices when A_2 is high and lower prices when A_2 is low, compared to the information advantage case. This intensifies competition when A_2 is low and softens competition when A_2 is high. When the benefit of reduced competition outweighs the cost of imperfect information, Firm 1 may benefit.

In addition, Firm 2's expected profit under the limited information scenario can be expressed as:

$$\begin{aligned}\mathbb{E}[\Pi_2^L] &= \mathbb{E}[\Pi_2^A] + \int_{\mathcal{A}_2}\int_{\mathcal{A}_1}(p_2^A \Delta q_2^{A,L} + \Delta p_2^{A,L} q_2^L) dF(A_1) dF(A_2) \\ &= \mathbb{E}[\Pi_2^A] + \underbrace{\frac{k_1 a_2 b_2(4a_1 - k_1 b_2)}{(4a_1 a_2 - b_2(b_1 + r b_2))^2}}_{\text{Gain from soften competition}} \left(\int_{\mathcal{A}_2}(A_2 - \eta_{A_2}) A_2 dF(A_2)\right) \geq \mathbb{E}[\Pi_2^A].\end{aligned}$$

Firm 2's profits increase in comparison to the scenario where Firm 1 has an information advantage.

Figure A.4: Effect of Variable k_1 on $\Pi_1^L - \Pi_1^A$



Note: Figure A.4 displays a simulation with parameters $a_1 = 12, a_2 = 10$, $b_1 10, b_2 = 5$, $k_1 = 0.2$, and $r = 15\%$. Additionally, Figure A.2b has $A_1 = 10$, and $E[A_2] = 10$, while Figure A.4 has $A_2 = 10$, and $E[A_1] = 10$. The concept of a "budget neutral" k refers to the value that equalizes the firm's profit to that of the information advantage case.

Consumer Welfare The consumer welfare for Firm 1 can be written as

$$
\begin{aligned}
CW_1^L &= CW_1^A + \frac{1}{a_1}\int_{\mathcal{A}_2}\int_{\mathcal{A}_1}(q_1^A + q_1^L)\Delta q_1^{A,L} dF(A_1)dF(A_2) \\
&= CW_1^A + \frac{k_1(2a_1a_2 - b_1b_2)\left(2a_1(-b_1 + 3a_2k_1 + rb_2) - b_2k_1(b_1 + 2rb_2)\right)}{2a_1(4a_1a_2 - (b_1 + rb_2)b_2)^2} \\
&\int_{\mathcal{A}_2}(A_2 - \eta_{A_2})A_2 dF(A_2) \geq CW_1^A \iff k_1 > \frac{2(a_1b_1 - a_1rb_2)}{6a_1a_2 - b_1b_2 - 2b_2rb_2}.
\end{aligned}
$$

Firm 1's consumer welfare increases with sufficiently large k_1, consistent with Firm 1's decreasing profit. On the other hand, the consumer welfare for Firm 2 is as follows:

$$\begin{aligned} CW_2^L &= CW_2^A + \frac{1}{a_2}\int_{\mathcal{A}_2}\int_{\mathcal{A}_1}(q_2^A + q_2^L)\Delta q_2^{A,L} dF(A_1)dF(A_2) \\ &= CW_2^A + \frac{a_2 b_2 k_1(4a_1 - b_2 k_1)}{2(4a_1 a_2 - (b_1 + rb_2)b_2)^2}\int_{\mathcal{A}_2}(A_2 - \eta_{A_2})A_2 dF(A_2) \geq CW_2^A. \end{aligned}$$

The welfare of Firm 2's consumers increases as Firm 1 is unable to factor in information from A_2 when setting their prices. An increase in overall consumer welfare compared to the information advantage case is indicated when the following inequality is satisfied:

$$CW^L > CW^A \iff k_1 > \frac{2a_2(2a_1a_2(b_1 - b_2 - rb_2) + b_1b_2(b_2 - rb_2) + b_2rb_2(b_2 + rb_2))}{12a_1^2a_2^2 + b_2^2(b_1 + rb_2)^2 - a_1a_2(b_1^2 + 8b_1b_2 - 2b_1rb_2 + 8b_2rb_2 + rb_2^2)}.$$

As depicted in Figure A.3b, when k_1 is sufficiently large, the overall consumer welfare generally increases. The intensified competition in Firm 2's market when A_2 is low leads to welfare gains, while the softened competition in the same market when A_2 is high results in welfare losses. However, under limited information relative to the information advantage case, it is still possible for consumer welfare to decrease. This occurs primarily when k_1 is small and Firm 1's cross-elasticity of demand b_1 is much larger than that of Firm 2's b_2. In such a scenario, the intensified competition when A_2 is low benefits Firm 1's consumers, but the softened competition when A_2 is high leads to losses that slightly outweigh the gains.

D Sales Error Probability from Restocking

One concern for inaccuracies in our sales measures is the restocking of products. We track the stock levels of approximately 2,400 products on an bi-hourly basis. Using the high frequency data, we can more effectively distinguish when sales occur versus when products are restocked. We identify these restocking events and compare how likely it is to underestimate sales in scenarios where the stock data is on a daily or weekly basis.

In Table A.3, we calculate the probability of underestimating sales as a result of restocking events. When the data is at a daily level, the likelihood of having measurement errors due to restocking is less than 1%. As expected, this probability increases as the data frequency becomes less frequent, such as when using weekly data, where there is a 7.12% chance of having measurement errors attributable to restocking.

Table A.3: Sales Error Probability from Restocking

Stock Data Frequency	Probability (%)
1 Day	0.98
2 Days	2.15
3 Days	3.12
4 Days	4.14
5 Days	5.22
6 Days	6.06
7 Days	7.12

Note: Table A.3 presents probabilities of sales underestimation due to restocking events at different stock data frequency levels.

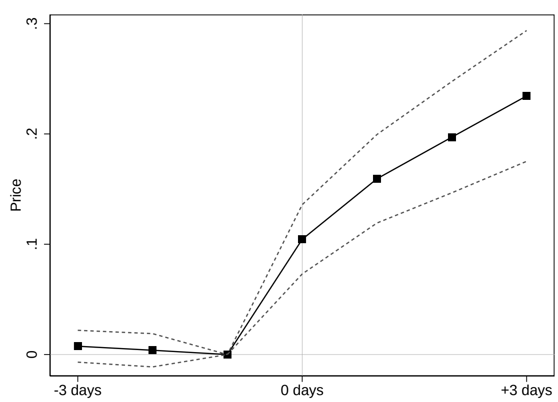
E Response Patterns of Large and Small Sellers

We examine whether the responses to one's own sales events and a competitor's sales events may depend on whether a third-party seller is large or small. Plausibly, a large seller might be more sophisticated or equipped with better technology. So, if our reduced-form evidence is driven by seller sophistication or technology, rather than information, we would expect to observe that large sellers respond more similarly to Amazon and differ from small sellers.

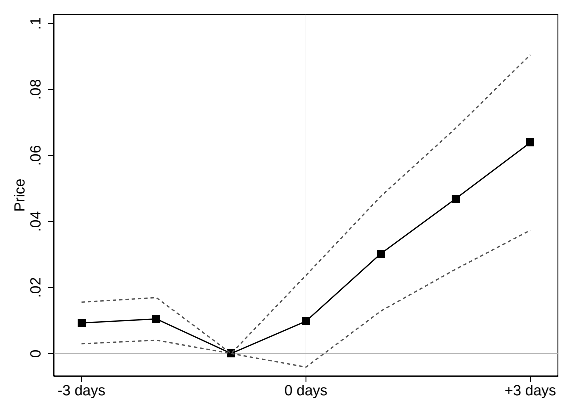
We define large sellers as those with the top 10% number of seller ratings, meaning a number of seller rating greater than approximately $17,500$. Amazon aggregates seller ratings for different products at the seller level, so a greater number of seller rating is associated with sellers who have more aggregate sales and are larger in size.

Figure A.5 displays the responses to a sales event for both large and small third-party sellers. Similar to the patterns observed between Amazon and third-party sellers, large third-party sellers also exhibit faster and larger responses in the price levels compared to small third-party sellers. However, in Figure A.6, both large and small third-party sellers are unresponsive to the competitor's sales event. Therefore, it is more plausible to explain the heterogeneity in response to competitor's sales events as being driven by information.

Figure A.5: Response of Large and Small Third-Party Sellers to a Sales Event



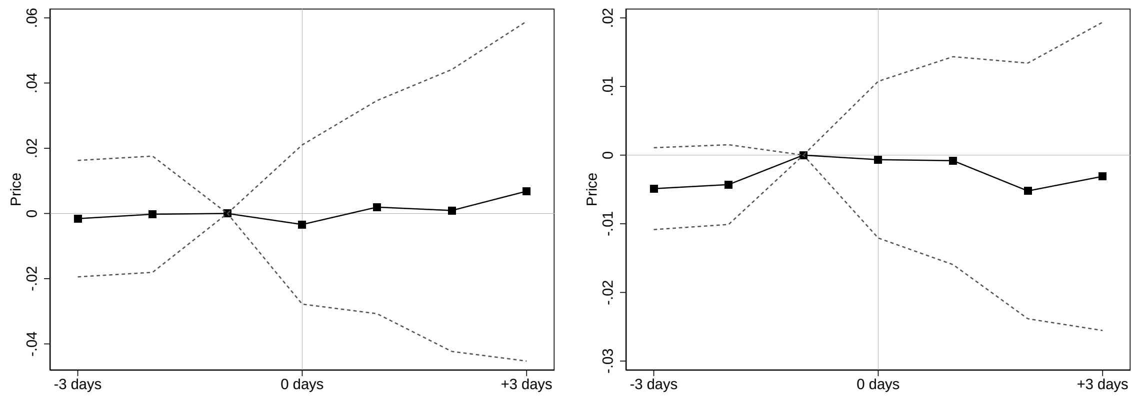
(a) Large Third-Party Sellers



(b) Small Third-Party Sellers

Note: Figure A.5 displays the estimates from Equation 2 using the price level as the dependent variable for large and small third-party sellers, respectively, before and after a sales event. The vertical line indicates the day of a sales event. We define large and small sellers using the median number of seller ratings. The robust standard errors are clustered at the product level.

Figure A.6: Response of Large and Small Third-Party Sellers to Competitor's Sales Event



(a) Large Third-Party Sellers

(b) Small Third-Party Sellers

Note: Figure A.6 displays the estimates from Equation 2 using the price level as the dependent variable for large and small third-party sellers, respectively, before and after a competitor's sales event. The vertical line indicates the day of a competitor's sales event. We define large and small sellers using the median number of seller ratings. The robust standard errors are clustered at the product level.

F Price Adjustments

We consider two types of outcomes related to pricing. First, we differentiate the direction of price adjustments and define the cumulative number of price adjustments in each direction for each event s on its event day $\tilde{t}$. For instance, to count price increases for an event s, we sum up the instances when the price on a particular day, $p_{s,\tilde{t}}$, is greater than the previous day's price, $p_{s,\tilde{t}-1}$. More formally, for all $s \in \mathbb{S}_7^{\text{no sales}}$, we define the cumulative number of price increases $\zeta^{\uparrow}_{s,\tilde{t}}$ and the cumulative number of price decreases $\zeta^{\downarrow}_{s,\tilde{t}}$ as

$$\zeta^{\uparrow}_{s,\tilde{t}} = \begin{cases} \sum_{\tau=2}^{\tilde{t}} \mathbb{1}(p_{s,\tau-1} < p_{s,\tau}) & \text{if } \tilde{t} \in \{2, \ldots, 7\}, \\ 0 & \text{if } \tilde{t} = 1; \end{cases}$$
$$\zeta^{\downarrow}_{s,\tilde{t}} = \begin{cases} \sum_{\tau=2}^{\tilde{t}} \mathbb{1}(p_{s,\tau-1} > p_{s,\tau}) & \text{if } \tilde{t} \in \{2, \ldots, 7\}, \\ 0 & \text{if } \tilde{t} = 1. \end{cases} \tag{A.1}$$

F.1 No-Sales Events

In Figure A.7, we present the estimates of γ_τ obtained from Equation A.1 using the cumulative number of price increases $\zeta^{\uparrow}_{s,\tilde{t}}$ and the cumulative number of price decreases $\zeta^{\downarrow}_{s,\tilde{t}}$ as dependent variables for third-party sellers and Amazon, respectively. The findings suggest that, during consecutive days of no sales, prices tend to decrease more often than they increase. This pattern is consistent for both Amazon and third-party sellers.

F.2 Sales Events

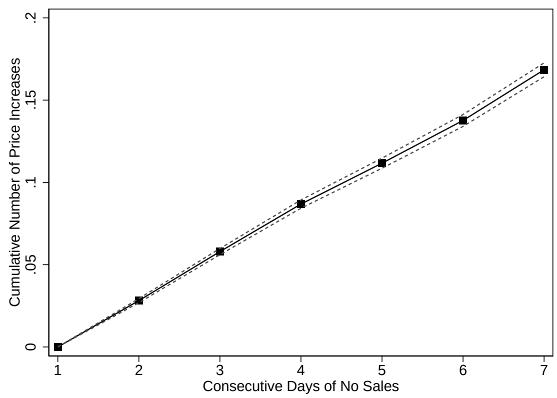
We examine the frequency and direction of price adjustments using the cumulative number of price increases $\zeta^{\uparrow}_{s,\tilde{t}}$ and decreases $\zeta^{\downarrow}_{s,\tilde{t}}$ for Amazon and third-party sellers, respectively. We define these variables similarly to Section 4.1, but restrict s to be in the set $\mathbb{S}_3^{\text{sales}}$ and $\tilde{t}$ to take values from -4 to 3. As shown in Figure A.8, both third-party sellers and Amazon are more likely to increase their prices following a sales event. Additionally, sellers are less likely to lower their prices after a sales event.[23]

F.3 Sellers prices depending on their past sales

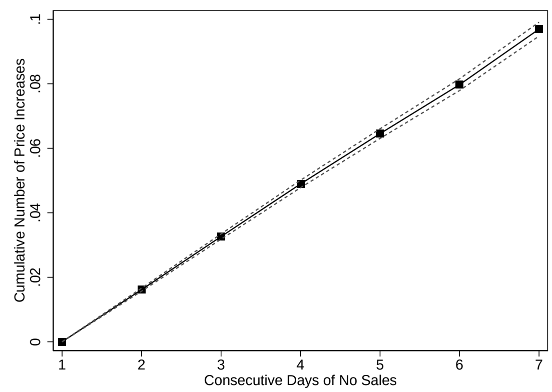
We investigate the relationship between a seller's current price and their past sales by measuring the number of days in which the seller has made any sales in the past.

[23]The outcomes in Figure A.9c and Figure A.9d are the cumulative number of price decreases, which is expected to increase over time only. However, we control for a linear trend, so the estimate can decrease if the increase in the cumulative number is less than the linear trend.

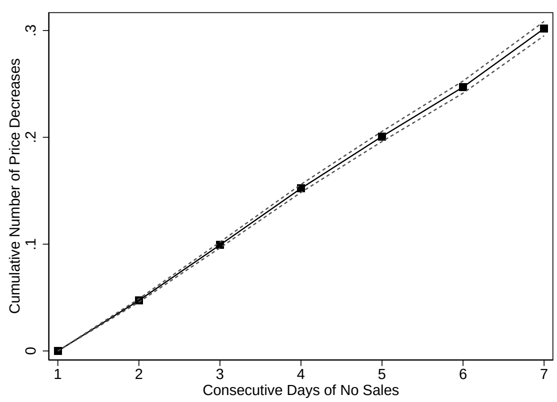
Figure A.7: Cumulative Number of Price Adjustments during Consecutive Days of No Sales



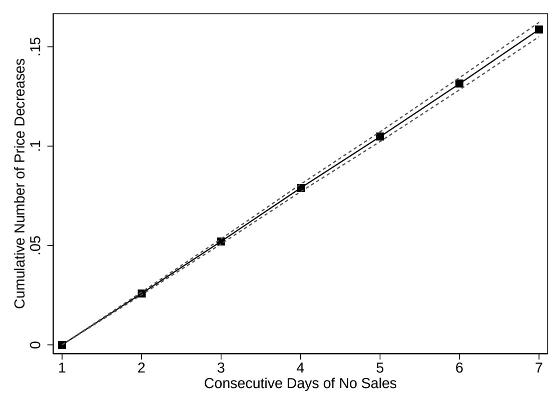
(a) Price Increase: Amazon



(b) Price Increase: Third-Party Sellers



(c) Price Decrease: Amazon



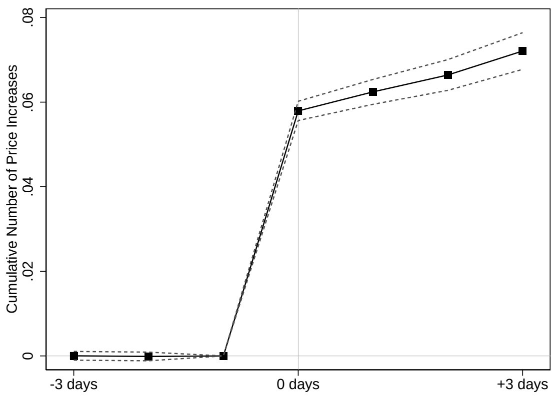
(d) Price Decrease: Third-Party Sellers

Note: Figure A.7 shows the estimates from Equation 1 during consecutive days of no sales using the cumulative number of price adjustments as the dependent variables for third-party sellers and Amazon, respectively. Figure A.7a and Figure A.7b present the cumulative number of price increases and Figure A.7c and Figure A.7d present the cumulative number of price decreases. The robust standard errors are clustered at the product level.

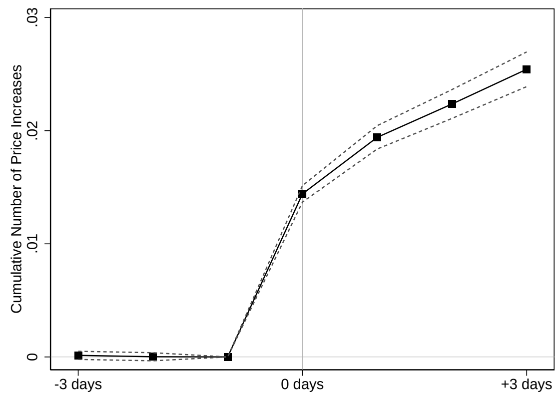
Specifically, we define the number of days with sales in the past as the following:

$$\mathbb{F}^{\text{sales}}_{\overline{\tau}m,j,t} = \sum_{r=1}^{\overline{\tau}} \mathbb{1}(q_{m,j,t-r} > 0).$$

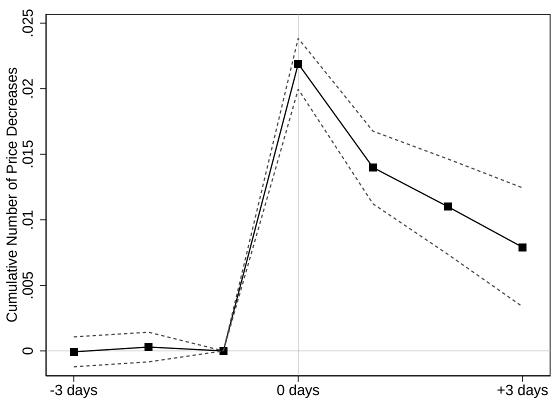
Figure A.8: Price Adjustments After a Sales Event



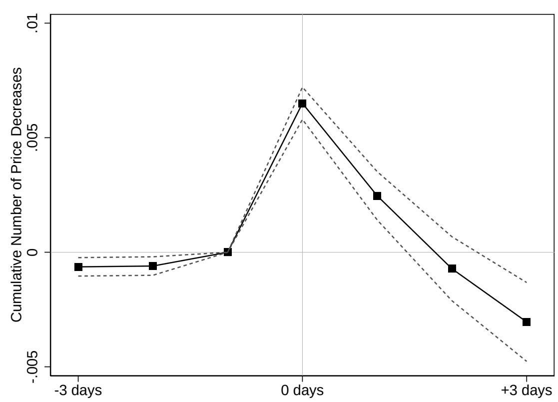
(a) Price Increase: Amazon



(b) Price Increase: Third-Party Sellers



(c) Price Decrease: Amazon



(d) Price Decrease: Third-Party Sellers

Note: Figure A.8 shows the estimates from Equation 2 using the cumulative number of price adjustments as the dependent variable for third-party sellers and Amazon, respectively, before and after sales events. Figure A.8a and Figure A.8b present the cumulative number of price increases and Figure A.8c and Figure A.8d present the cumulative number of price decreases. The vertical line indicates the day of the sales event. The robust standard errors are clustered at the product level.

We choose $\overline{\tau} = 7$, meaning that $\mathbb{F}^{\text{sales}}_{m,j,t}$ captures the number of days with sales in the past week. We then use the following specification to understand how the current price may vary based on the number of days with sales in the past week:

$$y_{m,j,t} = \sum_{\tau=0}^{4} \gamma \times \mathbb{1}(\mathbb{F}^{\text{sales}}_{m,j,t} = \tau) + \vartheta_{m,j} + \lambda_t + \iota_{m,j,t} + \epsilon_{m,j,t}. \tag{A.2}$$

In Equation A.2, the variable $y_{m,j,t}$ represents the outcome of seller j in market m on day t. The indicator variable $\mathbb{1}(\mathbb{F}^{\text{sales}}_{m,j,t} = \tau)$ takes a value of 1 if the observation has τ days of sales in the past week. The term $\vartheta_{m,j}$ represents seller-market fixed effects, which allow us to compare prices of the same seller in a given market based on the number of days during the past week when the seller made any sales. The term λ_t represents daily fixed effects, which capture the variation in prices across calendar dates. We also control for inventory fixed effects ($\iota_{m,j,t}$) in order to isolate the effects of inventory on prices.

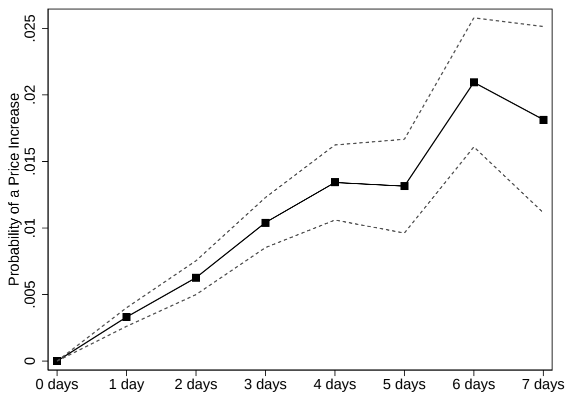
We present the estimates of γ from Equation A.2, using the cumulative number of price increases $\zeta^{\uparrow}_{m,j,t}$ and the cumulative number of price decreases $\zeta^{\downarrow}_{m,j,t}$, as defined in Equation A.1, in Figure A.9. Both third-party sellers and Amazon tend to increase their prices more and decrease their prices less frequently when they have had more days with sales in the past week.

Furthermore, we present the estimates of γ_τ in Figure A.10, using the price $p_{m,j,t}$ as the dependent variable in Equation A.2. The average price increase is larger when a seller has had more days with sales in the past, consistent with our findings for the probability of price adjustments.

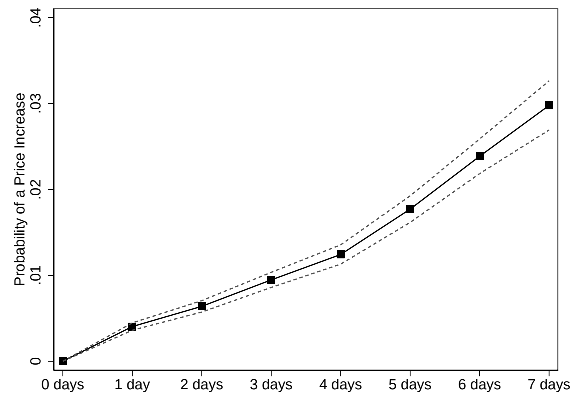
F.3.1 Correlated Demand

The estimates of γ_τ using sales $q_{m,j,t}$ as the dependent variable are displayed in Figure A.11, indicating a positive correlation between past sales and current sales. This implies that sales are autocorrelated and that a higher number of past sales is associated with a greater likelihood of higher sales in the present.

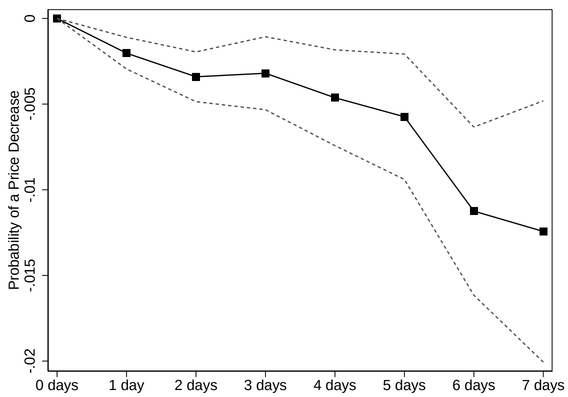
Figure A.9: Price Adjustments Over Number of Days With Sales During The Past Week



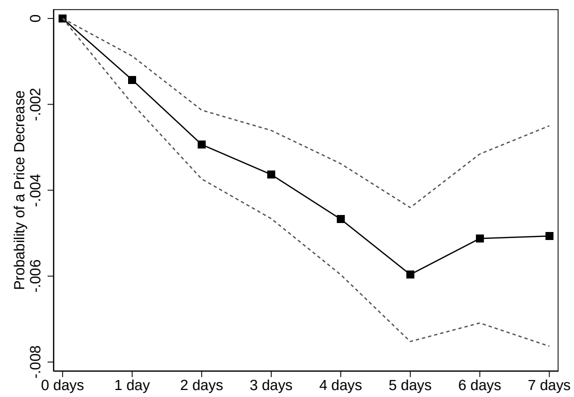
(a) Price Increase: Amazon



(b) Price Increase: Third-Party Sellers



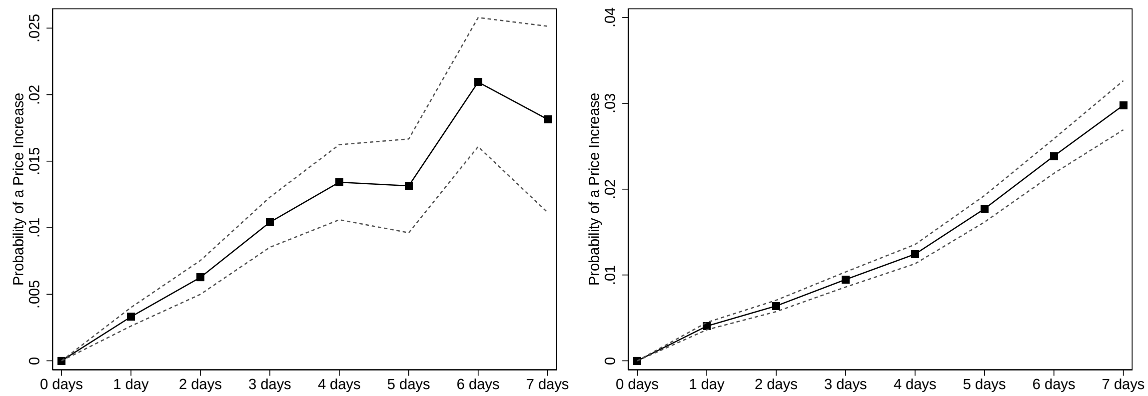
(c) Price Decrease: Amazon



(d) Price Decrease: Third-Party Sellers

Note: Figure A.9 plots the estimates from Equation A.2 over the number of days with sales during the past week using the probability of a price adjustment as the dependent variable for third-party sellers and Amazon, respectively. Figure A.9a and Figure A.9b present the probability of a price increase, and Figure A.9c and Figure A.9d present the probability of a price decrease. The robust standard errors are clustered at the product level.

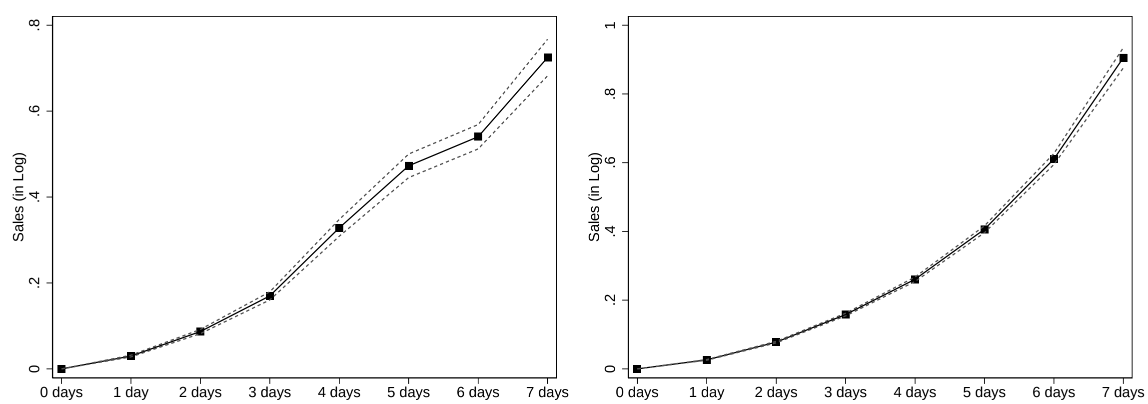
Figure A.10: Price Levels Over Number of Days With Sales During The Past Week



(a) Amazon (b) Third-Party Sellers

Note: Figure A.10 plots the estimates from Equation A.2 over the number of days with sales during the past week using the price level as the dependent variable for third-party sellers and Amazon, respectively. The robust standard errors are clustered at the product level.

Figure A.11: Sales Over Number of Days With Sales During The Past Week



(a) Amazon (b) Third-Party Sellers

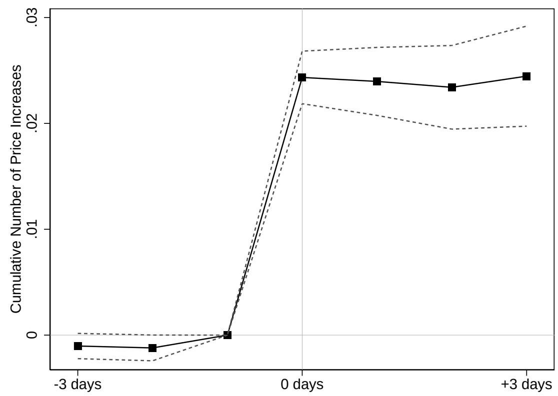
Note: Figure A.11 plots the estimates from Equation A.2 over the number of days with sales during the past week using the log of sales as the dependent variable for third-party sellers and Amazon, respectively. The robust standard errors are clustered at the product level.

F.4 Competitor's Sales Events

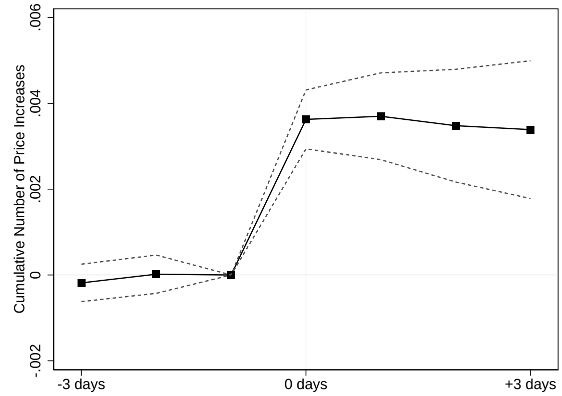
In Figure A.12, we plot the estimates of γ_τ using the cumulative number of price adjustments as the dependent variable for Amazon and third-party sellers, respectively. According to our findings in Section 4.2, after a competitor's sales event, the competitor increases its price. In equilibrium, Amazon and third-party sellers may also increase their prices in response. We introduce a structural model of price competition that deals with this equilibrium effect in Section 5.

Interestingly, we find that Amazon may reduce its prices following a competitor's sales event. The frequency of price decreases for Amazon is approximately ten times greater than that of third-party sellers. This effect is highly significant for Amazon, while it is insignificant for third-party sellers when the pretrend is taken into account.

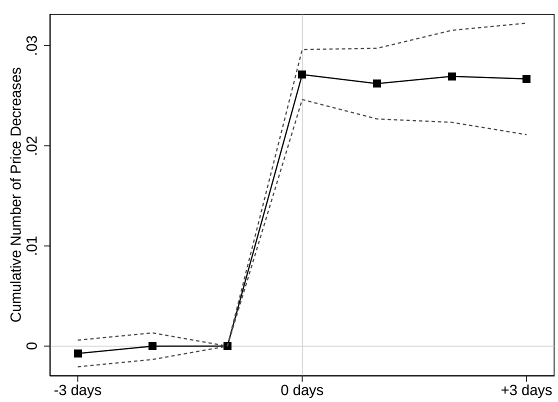
Figure A.12: Price Adjustments After a Competitor's Sales Event



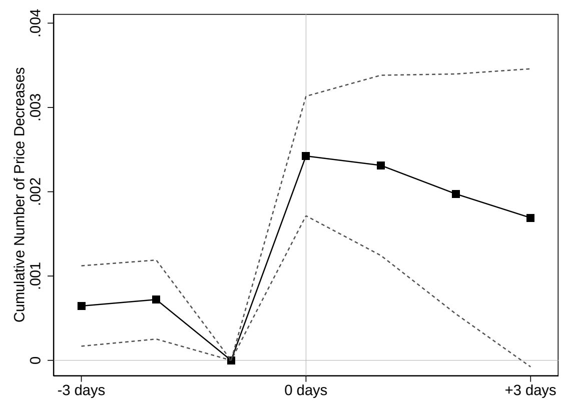
(a) Price Increase: Amazon



(b) Price Increase: Third-Party Sellers



(c) Price Decrease: Amazon



(d) Price Decrease: Third-Party Sellers

Note: Figure A.12 displays the estimates from Equation 2 using the cumulative number of price adjustments as the dependent variable for third-party sellers and Amazon, respectively, before and after a competitor's sales events. Figure A.12a and Figure A.12b present the cumulative number of price increases and Figure A.12c and Figure A.12d present the cumulative number of price decreases. The vertical line indicates the day of a competitor's sales event. The robust standard errors are clustered at the product level.

G Estimation of Demand States

To address the potential bias in the coefficient of η caused by the endogeneity of prices when estimating Equation 8, we use the log of inventory as an instrument for price. The inventory level serves as a cost shifter since sellers may face higher costs when they have excess inventory due to storage costs.

In Column 1 of Table A.4, we present the results of the first stage regression, which show that the price tends to be higher when the log inventory is higher. In Column 2 of Table A.4, we provide the estimates from the IV regression. As expected, the coefficient of the price is negative.

Table A.4: Estimation of Demand States

	(1) First Stage	(2) 2SLS
$\log(I_{m,j,t})$	0.267***	
	(0.021)	
$p_{m,j,t}$		-0.044***
		(0.005)
Product-Seller Fixed Effects	Yes	Yes
No. of Observations	31,097,387	31,097,387

Note: Table A.4 shows the results from Equation 8. The first stage estimates are presented in Column 1 and the two-stage least squares (2SLS) estimates are displayed in Column 2. The robust standard errors are cluster at the product level. The significance levels are indicated as follows: *(p < 0.10), **(p < 0.05), ***(p < 0.01).

H Moment Conditions

The vector $\boldsymbol{z}_{m,j,\tilde{t}} = \boldsymbol{M}_{m,j,\tilde{t}} \otimes \boldsymbol{H}_{m,j,\tilde{t}} \otimes \boldsymbol{T}_{m,j,\tilde{t}}$ represents a vector of 160 indicators, where $\boldsymbol{M}$ represents the 5 market structures (5×1), $\boldsymbol{H}$ represents the combinations of best predicted and true demand states of focal seller and the competitor (16×1), and $\boldsymbol{T}$ represents the period before and after repricing (2×1) as the following:

$$\boldsymbol{M}_{m,j,\tilde{t}} = \begin{bmatrix} \mathbf{1}(\mathcal{M}(m) = \text{AMZ-vs-FBM}) \\ \mathbf{1}(\mathcal{M}(m) = \text{AMZ-vs-FBA}) \\ \mathbf{1}(\mathcal{M}(m) = \text{FBA-vs-FBM}) \\ \mathbf{1}(\mathcal{M}(m) = \text{FBA-vs-FBA}) \\ \mathbf{1}(\mathcal{M}(m) = \text{FBMvs-FBM}) \end{bmatrix}, \boldsymbol{H}_{m,j,\tilde{t}} = \begin{bmatrix} \mathbf{1}(\hat{h}_{m,j,t} = 1) \\ \mathbf{1}(\hat{h}_{m,j,t} = 0) \\ \mathbf{1}(\hat{h}_{m,-j,t} = 1) \\ \mathbf{1}(\hat{h}_{m,-j,t} = 0) \end{bmatrix} \otimes \begin{bmatrix} \mathbf{1}(h^{\text{BP}}_{m,j,t} = 1) \\ \mathbf{1}(h^{\text{BP}}_{m,j,t} = 0) \\ \mathbf{1}(h^{\text{BP}}_{m,-j,t} = 1) \\ \mathbf{1}(h^{\text{BP}}_{m,-j,t} = 0) \end{bmatrix},$$

and $\boldsymbol{T}_{m,j,\tilde{t}} = \begin{bmatrix} \mathbf{1}(\tilde{t} = 1) \\ \mathbf{1}(\tilde{t} = -1) \end{bmatrix}$.

I Information Sharing in Third-Party-Only Markets

Information sharing among third-party sellers can significantly impact markets where these sellers compete with one another, and Amazon does not operate as a seller. The counterfactual outcomes for this scenario are delineated in Table A.5. Amazon, as the platform owner, earns profits from referral fees without engaging in direct market participation. In contrast to markets that involve Amazon, vertical incentives hold less significance because the referral fee primarily serves as a a static modifier to sellers' marginal costs.

In markets characterized by the presence of one FBA seller and one FBM seller, consumer welfare and social welfare experience a decrease of 1.22% as competition weakens and both sellers elevate their prices. Concurrently, the total profit for sellers witnesses a 2.21% increase. Amazon's referral fee earnings decrease as a result of the overall decrease in transactions. This decrease in competition could be due to the the asymmetry among sellers.

Regarding FBA-vs-FBA and FBM-vs-FBM markets, competition demonstrates greater intensity, leading to average price reductions of 3.34% and 0.85%, respectively. Consequently, a larger share of welfare is allocated to consumers, resulting in a significant increase in consumer welfare by 35.84% and 2.5%, respectively. In addition, social welfare increases by 18.82% and 0.89% for each market type, respectively.

Table A.5: Information Sharing with Third-Party Sellers, Third-Party-Only Markets

	15% Referral Fee			0% Referral Fee		
	FBA-vs-FBM	FBA-vs-FBA	FBM-FBM	FBA-vs-FBM	FBA-vs-FBA	FBM-FBM
Panel A: Equilibrium Outcomes						
%Δ FBA Price	1.10		-0.85	1.27		-0.88
%Δ FBA Sales	0.41		2.22	1.08		2.07
%Δ FBM Price	0.04	-3.34		-0.23	-3.98	
%Δ FBM Sales	-15.94	34.68		-14.02	33.46	
Panel B: Profit and Welfare Change						
%Δ Amazon Profit	-2.26	27.71	1.57	0.00	0.00	0.00
%Δ 3rd-Party Profit	2.21	-24.13	-1.67	2.35	-25.20	-1.53
%Δ Consumer Welfare	-2.49	35.84	2.50	-1.84	35.80	2.26
%Δ Social Welfare	-1.22	18.82	0.89	-0.11	11.30	0.45

Note: Table A.5 presents the counterfactual analysis when Amazon shares its information with third-party sellers. Panel A shows the percentage changes in equilibrium prices and sales, while Panel B quantifies the changes in profits and welfare for Amazon, third-party sellers, and consumers. FBA-vs-FBM represents the market where an FBA seller competes with an FBM seller. FBA-vs-FBA represents the market where two FBA sellers competes with each other. FBM-vs-FBM represents the market where two FBM sellers competes with each other. Columns 1 to 3 correspond to the current 15% referral fee, and Columns 4 to 6 correspond to the case when the referral fee is reduced to 0%.