Price Competition Under Information (Dis)Advantage

Nan Chen Hsin-Tien Tsai*

September, 2024

Abstract

We examine the impact of asymmetric access to competitively sensitive information on a vertically integrated e-commerce platform. High-frequency data variation facilitates the identification of supply-side information asymmetry: while both the owner and third-party sellers base pricing on past sales, only the owner leverages competitors' sales. We estimate a price competition model with demand uncertainty and heterogeneous beliefs. Counterfactual analyses of symmetric access, whether by restricting or sharing superior information, suggest that information advantage can be anti-competitive, reducing consumer and social welfare, and that this impact interacts with vertical incentives. Furthermore, sharing information leads to greater welfare by enhancing market efficiencies.

Keywords: Asymmetric information, price competition, e-commerce, digital platform, vertical integration, information design.

*Chen: Department of Information Systems and Analytics, 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 Zach Brown, David Byrne, Tat Chan, Soheil Ghili, Jasmine Hao, Julian Hidalgo, Brett Hollenbeck, Chuqing Jin, Ginger Zhe Jin, Martin Peitz, Andrew Rhodes, Joel Waldfogel, Kevin Williams, Mo Xiao, Dongni Zhu, and participants at APIOC, Berkeley SICS, BU Platform Strategy Symposium, CEPR-JIE, CRESSE, IIOC, NBER-SI, TSE Digital Economics Conference, and other conferences and seminars. 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 information sharing and its role in shaping competitive outcomes (Kühn and Vives, 1995) have become increasingly relevant in a data-driven world. Data, often characterized as the "non-rival" new oil (Varian, 2018), exhibits potential asymmetric access for firms. One such asymmetry may arise from vertical integration (Arrow, 1975), which could lead to an integrated firm's access to competitively sensitive information, as regulators emphasize in the vertical merger guideline.[1] However, limited empirical evidence points to the anti-competitive effects related to information. This paper attempts to make progress on the competitive and welfare implications of information asymmetry and its interaction with vertical incentives.

Our empirical context is a dominant, vertically integrated e-commerce platform where the platform owner, Amazon, competes with hosted third-party sellers on prices. The platform owner manages all transactions and maintains extensive information, while third-party sellers likely possess limited information beyond their own business. Specifically, we aims to empirically identify supply-side information asymmetry. In addition, with concerns over platforms monopolizing data-driven competitive advantages, the information design within these vertically integrated platforms holds important implications for competition policy (Crémer et al., 2019; Cabral et al., 2021).

We compile a daily panel dataset for products on Amazon.com spanning a 7-month period. The dataset comprises over 122,000 products, each with two sellers offering the same product in new condition, allowing us to study price competition in a duopoly market, the most prevalent oligopoly structure on the platform. We collect daily information on price and inventory for each product listing, where each listing represents a seller offering a given product. We use inventory data to approximate daily sales at the listing level.

We begin by presenting reduced-form evidence on how sellers set prices based on past sales information.[2] The high-frequency pricing and sales data enable us to plausibly infer sellers' causal responses to past sales. We demonstrate that sellers tend to gradually decrease their prices when experiencing consecutive days without sales, aligning with pessimistic beliefs about demand (Mason

[1] "[T]he downstream component of the merged firm could now have access to its rival's sensitive business information. In some circumstances, the merged firm can use access to a rival's competitively sensitive information to moderate its competitive response to the rival's actions." See U.S. Department of Justice and The Federal Trade Commission (2020).

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

and Välimäki, 2011). Subsequently, we find that sellers are likely to increase their prices immediately following a sale event, suggesting that they adjust 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.

Next, we present novel evidence supporting supply-side asymmetric information regarding competitors' sales. Although competitors' sales information is non-public, the platform owner has access to this information, while third-party sellers do not, potentially leading to information asymmetry. We find that when Amazon's competitors make sales, Amazon adjusts its prices accordingly. On the other hand, third-party sellers show limited response to their competitors' sales. We also show that although large third-party sellers exhibit stronger responses to their own sales compared to smaller sellers, their reactions to competitors' sales remain similarly muted, highlighting that it is the information, not the sophistication, that drives the response. The findings align closely with the results of our industry surveys. This discrepancy in the ability to respond to competitors' sales highlights the informational asymmetry and potentially competitive advantage Amazon possesses as the platform owner when competing against third-party sellers.

Moving forward, we build a simple structural model of price competition.[3] On the demand side, consumer preferences conform to a nested logit model, with sellers of a product in one nest and the outside option in another. The utility model incorporates a time-varying demand intercept; the autocorrelation of the demand intercept allows past demand information (prior to day t) to forecast the current demand (on day t). The demand prediction, which uses all available past information, is denoted as the best predicted demand states h^{BP}. On the supply side, each seller forms heterogeneous beliefs about the intercepts for both their own and their competitors' demand, which may differ from h^{BP}, and participates in a simultaneous-move pricing game.

Using the realized sales, we estimate the demand intercept, discretize it into true demand states h, and use this along with a prediction model to estimate the discretized h^{BP}.[4] Firms' beliefs are specified as a reduced-form function of h^{BP}. This function is monotonic, indicating better information access when firms' prices approximate the price they would set with accurate beliefs h^{BP} and worse otherwise. Firms' heterogeneous belief functions are identified by how their

[3] Online Appendix A considers a toy model of price competition under information advantage and finds that, theoretically, the welfare changes are largely uncertain.

[4] The key distinction is that true demand states h are determined ex-post, after the realization of day t sales, while the best predicted demand states h^{BP} are determined ex-ante, based on past sales data preceding the realization of day t sales.

prices respond to past sales (or h^{BP}).

We estimate the model using the Method of Simulated Moments. We employ a full solution approach, which jointly solves the demand and supply model to explicitly account for price endogeneity and the large fraction of zero sales in the data. We numerically integrate unobserved states (i.e., random effects in consumers' utility functions and firms' cost functions) to compute moments. Model parameters are selected based on their ability to replicate empirical moments, such as prices, sales, and their interactions.[5] For each iteration of candidate parameters, we solve equilibria for 12,960 pricing games.

We estimate an average own-price elasticity of -18.96, in line with existing literature (Ellison and Ellison, 2009; Dinerstein et al., 2018). Under a price competition framework where prices are endogenous and set simultaneously, our structural estimates provide an additional test for belief differences between Amazon and third-party sellers. First, the estimates for own-belief parameters indicate that both Amazon and third-party sellers have accurate beliefs about their own demand states. Second, regarding beliefs about competitors' demand states, Amazon's beliefs align with the competitor's h^{BP}. However, third-party sellers' beliefs closely resemble random guesses. These estimates are consistent with the reduced-form evidence, which suggests that Amazon has an information advantage regarding competitors' sales compared to third-party sellers.

We conduct counterfactual analyses to examine the equilibrium outcomes of a symmetric information environment, specifically: (1) preventing the platform owner from using marketplace information for its retail business and (2) sharing marketplace information with third-party sellers. These counterfactuals correspond to proposals that Amazon put forward in response to EU antitrust investigations.[6]

In the first counterfactual, we examine a scenario where Amazon is prohibited from using competitors' past sales for pricing by setting Amazon's belief parameter to be identical to that of third-party sellers. Under the current 15% commission rate, Amazon's first-party sales and profits decrease by 0.48% and 0.30%, respectively, while third-party sellers experience a 3.10% increase in sales and a 3.04% profit gain. Although more information can increase efficiency, superior information also grants market power, enabling the advantaged party to know when to

[5] We exploit the moments generated from within-product listing variations in price, using a sample of local repricing events. We assume that within a 3-day time window, price variations are driven by firms' beliefs or costs, while the underlying demand remains relatively stable.

[6] See https://www.reuters.com/technology/amazon-offers-share-data-boost-rivals-dodge-eu-antitrust-fines-sources-2022-06-13/.

raise prices (e.g., Acemoglu, 2021). This results in an anti-competitive effect that outweighs the efficiency gains. As a result of removing the information advantage and, consequently, market power, consumer welfare and social welfare improve by 0.45% and 0.35%, respectively.

We find that the information advantage interacts with vertical incentives. When Amazon does not charge a commission fee, eliminating its vertical incentives, restricting Amazon's information advantage results in a larger decrease in Amazon's profits of about 4.22%, while third-party sellers' profits increase by 10.11%. Superior information harms third-party sellers to a greater extent when Amazon does not charge a commission fee. As vertical incentives soften the competition between Amazon and third-party sellers, the absence of these incentives makes the market power stemming from informational advantage a more prominent concern, reducing social welfare by 0.78%.

In the second counterfactual, we investigate another symmetric information scenario where Amazon shares its superior information with third-party sellers by resetting the belief parameter of third-party sellers to match that of Amazon. Third-party sellers' sales increase by 19.78%, while Amazon's sales decrease by 3.58%. Moreover, information sharing leads to a more substantial increase in consumer welfare and social welfare on average, with estimated increases of 2.41% and 2.01%, respectively. Sharing information yields a lesser average price reduction but a greater social welfare gain than restricting information. The gain mostly comes from the increased information enhancing the matching quality of supply and demand. Despite reduced first-party profits, Amazon may still benefit from sharing information, as the increase in commission fees outweighs the losses.

We further highlight that vertical incentives, shaped by commission rates, mediate the competition and welfare implications of different information structures. We also find that symmetric information would lead to a lower commission rate, if reoptimized. The positive welfare impact of sharing information or restricting Amazon's information advantage would be amplified with a reoptimized commission fee.

In summary, this paper examines the implications of information asymmetry on competition between a vertically integrated platform owner and its hosted competitors. We empirically identify the presence of information advantage for the owner and show that this advantage can serve as a source of market power, enabling the advantaged party to extract greater economic rents. Compared to the current asymmetric information access, symmetric information reduces market power and increases social welfare. Furthermore, sharing superior information leads to a greater increase in market efficiency than restricting it.

1.1 Related Literature

This study builds on theories of information sharing in oligopolies (Vives, 1984; Li, 1985; Gal-Or, 1985). Recent empirical work examines 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.[7] In a concurrent study, Byrne et al. (2023) examine the impact of asymmetric access to competitors' price information in retail gasoline markets. They show that price competition becomes less intense when information shifts from symmetric to asymmetric, as this allows a less informed firm to use price commitment. We study information asymmetry in a vertical market structure without collusion and find that asymmetric information access is anti-competitive and inefficient compared to symmetric information. In addition, we focus on past sales information, which can be used as an input for predicting current demand states (Bajari et al., 2019).[8]

We model firms' beliefs as a function of past data and identifies these beliefs from the dependence of firms' strategies on past data. This approach relates to recent empirical studies that examine how firms learn and form beliefs in competitive environments (Doraszelski et al., 2018; Jeon, 2022; Huang et al., 2022).[9] 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. We identify the presence of supply-side information friction in competitive environment and develop a tractable model to assess the implications of asymmetric information.

This study focuses on the design of access to data for sellers within the context of a vertically

[7] Classic empirical literature highlights how the disclosure of product quality information affects equilibrium outcomes (Jin and Leslie, 2003; Tadelis and Zettelmeyer, 2015).

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

[9] For a comprehensive review, see Aguirregabiria and Jeon (2020) and Aguirregabiria (2021). Asker et al. (2020) examine competitive information sharing in a dynamic auction environment. Our paper is broadly related to recent work investigating the competitive effects of pricing technologies that adjust according to a rival's historical prices (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 using different learning protocols when there is market interaction.

integrated platform, relating to the incentives and efficiency in vertically integrated markets (Hart et al., 1990; Rey and Tirole, 2007). Specifically, we contribute to a new and growing empirical literature on vertically integrated platforms that focuses on the design of information for consumers and its implications (Chen and Tsai, 2019; Lee and Musolff, 2021; Lam, 2021; Raval, 2022; Farronato et al., 2023; Reimers and Waldfogel, 2023). To the best of our understanding, we present a perspective that has yet to be empirically explored: the information advantage of integrated firms over non-integrated firms (i.e., the platform owner over third-party sellers in this context), which can be termed "data-driven self-preferencing" (Cabral et al., 2021). 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 (see also Crawford et al., 2022). High-frequency variations in pricing and sales allow us to causally identify information advantages, which are typically difficult to observe due to the unobservable nature of information structures. In a theoretical work, Madsen and Vellodi (2024) investigate how a platform owner may leverage its informational advantage to imitate products offered by third-party sellers.[10]

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 information asymmetry that affects pricing patterns. 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

There has been much discussion in academic and policy circles about platforms' dual role. One specific concern is monopolistic market power resulting from platforms' data collection and exclusive control over this data, as they may exploit this information asymmetry to compete unfairly with their users (e.g., Cabral et al., 2021).

The empirical context of this paper is Amazon, which operates in a dual role, which operates in a dual role. 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

[10] 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 (Jones and Tonetti, 2020; Acemoglu, 2021; Bergemann et al., 2022).

their products to customers worldwide and make up a significant 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, directly competing with these third-party sellers.[11]

Information Asymmetry. Data collected from third-party sellers, which Amazon can access due to its role as the platform owner, may be used in ways that do not align with the best interests of these sellers or consumers. Concerns have been raised about Amazon potentially using marketplace information in its product launch decisions; however, information about best-selling products is publicly available. Zhu and Liu (2018) show that launch decisions can be based on sales rank, which is public information.[12] We focus on the aspect of information asymmetry, specifically *non-public data*, that could lead to "data-driven self-preferencing."

In July 2019, the European Commission initiated an inquiry into Amazon's utilization of marketplace vendor data.[13] In December 2022, Amazon pledged to refrain from using marketplace seller data for retail operations decisions. Although we began collecting data from Amazon's US store before this pledge, we do not observe any noticeable changes in the data afterward.

Sales Information on Amazon. Data-driven strategic decisions increasingly play a critical role in a company's performance (e.g., Bajari et al., 2019). One way to demonstrate 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

[11] Third-party sellers can operate 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 handles shipping and related services. However, in addition to commission fees, FBA sellers also incur extra FBA fulfillment fees. Refer to https://sell.amazon.com/pricing#fulfillment-fees for further details.

[12] See https://www.wsj.com/articles/amazon-scooped-up-data-from-its-own-sellers-to-launch-competing-products-11587650015. Amazon replies that "Like other retailers, we look at sales and store data to provide our customers with the best possible experience."

[13] 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.", see https://competition-cases.ec.europa.eu/cases/AT.40462 for more details. A potential violation could result in a fine amounting to 10% of its annual turnover, exceeding $60 billion USD in 2022.

on new information.

Our focus is on past sales information. Theoretically, the cost of providing sales information is similar to that of providing price information. While price information is transparent and relatively inexpensive to obtain through public APIs, sales information 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. We find that 47.4% of these platforms disclose sales information (refer to Appendix A for details).

Inferring individual listing sales within a product based on publicly available information on Amazon is difficult. Amazon's primary sales metric, the product's sales rank, is an aggregate of sales information across all listings of a given product and ranks its total sales within its product category. The platform owner may access sales information with minimal marginal cost, as it oversees all transactions. Conversely, third-party sellers can observe only their own sales and may have limited information beyond their own sales figures.[14] Although they can utilize their own sales information and sales rank to estimate competitors' sales to some extent, the information is noisy, as the sales rank only represents a relative ranking and not the actual number of sales; it can change even if the sales of the focal product remain constant, due to changes in the sales of other products.

Third-party sellers may still use external resources to gain insights into competitors' sales. Services like Jungle Scout offer professional sales estimation for sellers, but the estimates are not high-frequency and can be costly.[15] 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- or technology-intensive and requires significant costs to integrate this information into pricing.

[14] There were practices of potential A/B testing conducted by Amazon, indicating the possible disclosure of some form of "sales data." Industry researcher Waters (2023) notes 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... 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. The public disclosure of sales data is still in its early stages, and there is no indication of disclosing seller-listing level sales data, which is our focus.

[15] As of September 2024, tracking 50 products costs $49 per month; see https://www.junglescout.com/pricing/.

3 Data

We collect a daily panel dataset from Amazon.com for a period of 7 months, covering from October 2022 to May 2023, for the Home & Kitchen category.[16] We follow a daily routine for collecting data on the final set of approximately 122,000 products. We use the historical information from Keepa.com to identify products that had an average of between 1.5 to 2.49 listings in new condition over the past 90 days. 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. We also collect data for a small sample of 2,400 products in *two-hour interval* to complement the main daily dataset for some robustness checks.

We gather information on different aspects of each seller of a given product, such as its price, shipping fee, seller types, and inventory. We use the current available quantity of a product for purchase on the website as our measure of 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 in 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. See Appendix B for the discussion of price variation. The Shipping variable is a fixed fee charged per order by the seller.[17]

[16] 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).

[17] On average, third-party sellers charge a shipping fee of approximately $2.71, while Amazon charges an average

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.[18] On average, daily sales for a product in our dataset amount to approximately $0.14 - 0.15$. Appendix C discusses concerns related to potential bias in sales estimates due to restocking within a given day. We use the smaller dataset collected at two-hour intervals and show that the probability of underestimating sales is less than 1% when the inventory 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. We also collect the "buy box" status, which indicates whether a seller is the featured seller. In the data, the average third-party seller has a 41% probability of being the featured seller, while Amazon has an 81% probability of being the featured seller.

The number of seller ratings is commonly used by industry practitioners to infer seller size.[19] We use it to define large and small third-party sellers in Section 4.4. The third-party sellers in the data have a median of 595 ratings.[20]

4 Reduced-Form Analysis

In this section, we use the high-frequency dataset to investigate sellers' pricing behaviors. We present reduced-form evidence to support the idea that sellers adjust their prices based on past sales:

shipping fee of $0.02. 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.

[18] 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.

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

[20] 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.

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
Buy Box	0.41	0.00	0.49	0.81	1.00	0.39
No. of Seller Ratings	11,935.60	595.00	42,418.70			
No. 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.

(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 adjust pricing based on competitor sales information, with Amazon being more responsive to changes in competitor sales. Additionally, in Online Appendix B.3, we present additional evidence that prices are affected by past sales, and demand exhibits autocorrelation in Appendix D.

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 $\mathcal{S} : \mathbb{D}c \to \mathbb{S} \subseteq \mathbb{N}$. We use this sample selection procedure with varying selection conditions c in the reduced-form analysis.

4.1 Seller Prices Based on No-Sales Information

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" events:

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

To analyze these events, we let $\overline{r} = 7$ and estimate the following regression specification:[21]

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

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.[22]

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 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.[23]

4.2 Seller Prices Based on Sales Information

We demonstrate that sellers raise their prices in response to higher demand, as 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 few days prior to the event (e.g., days -3 to -1) but made sales on the event day (day 0). Formally, we define the set of "sales" event as follows:

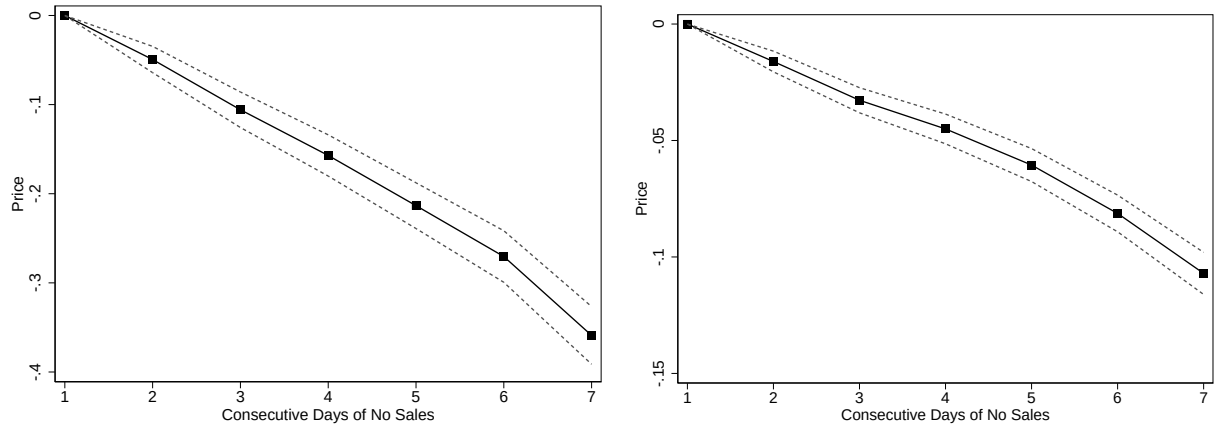
$$\mathbb{S}_{\overline{r}}^{\text{sales}} = \{\mathcal{S}(m,j,t) \in \mathbb{N} \mid q_{m,j,t} > 0 \ \wedge \ q_{m,j,t-r} = 0\,, \quad \forall r = 1,\ldots,\overline{r}\}\,.$$

[21] 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.

[22] The fixed effects ψ_s are more granular than product fixed effects and enable us to compare prices within a given product and across neighboring days.

[23] 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 address the limitations of the reduced-form evidence and incorporate the heterogeneity in demand primitives.

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.

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 $\overline{r} = 3$:

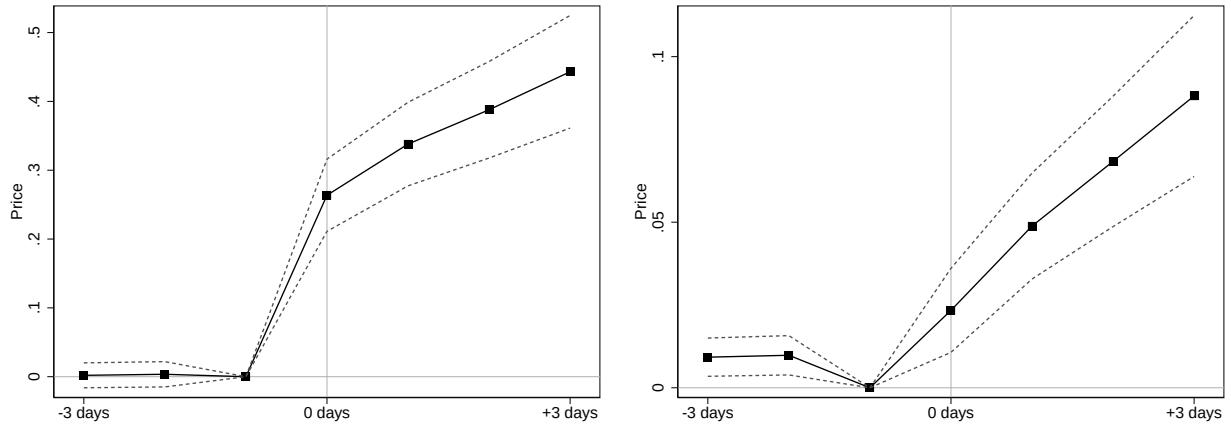
$$y_{s,\tilde{t}} = \sum_{\tau=-3,\tau\neq-1}^{3} \gamma_\tau \times \mathbb{1}(\tilde{t} = \tau) + \eta \times \tilde{t} + \psi_s + \epsilon_{s,\tilde{t}}, \ \forall \ s \in \mathbb{S}_3^{\text{sales}} \wedge \tilde{t} \in \{-4, \ldots, 3\}. \tag{2}$$

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.[24]

We use price levels $p_{s,\tilde{t}}$ as the outcome variable. Figure 2 shows that the average price for both Amazon and third-party sellers increases following a sales event, with an average increase of around \$0.40 for Amazon and \$0.10 for third-party sellers.

[24] 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.

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 pre-trend is controlled by the linear trend η. The robust standard errors are clustered at the product level.

4.3 Information Asymmetry Regarding Competitor's Sales

Sellers can also set prices based on 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 are not directly accessible to third-party sellers on the website, while Amazon may access this information at minimal cost. Consequently, we expect that Amazon is more responsive to competitors' sales compared to third-party sellers.

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

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

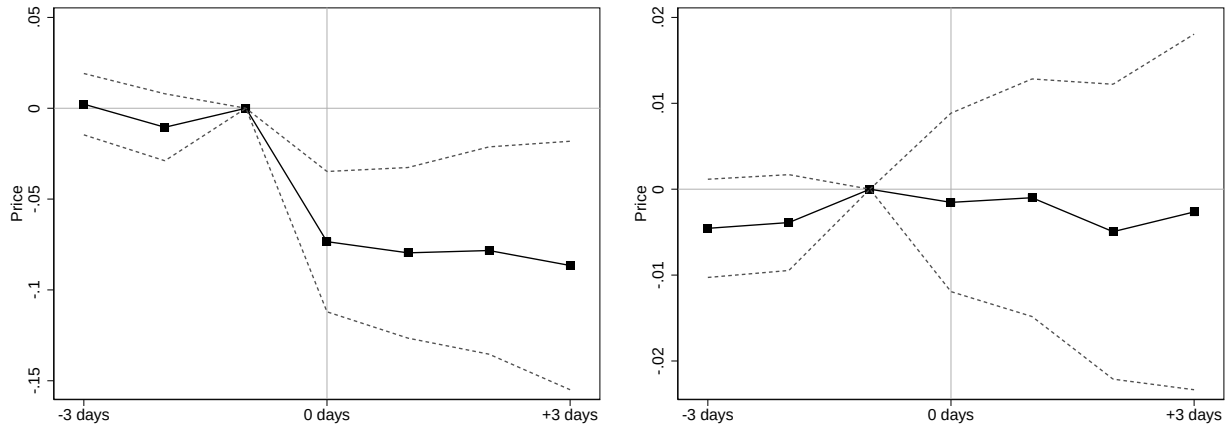
To isolate the effect of own sales on prices, as discussed in Section 4.2, we restrict our attention to cases where only competitors make sales and the focal sellers do not.

Next, we examine Equation 2 using observations from competitor sales events, where $s \in \mathbb{S}_{\bar{r}}^{\text{c-sales}} \wedge \tilde{t} \in \{-4,\ldots,3\}$. 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.[25] The price reduction supports the theory that Amazon reduces its own prices in response to a competitor's more appealing offer to consumers.[26] On the other hand, the price levels of third-party sellers have no significant response following a competitor's sale.[27]

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 pre-trend is controlled by the linear trend η. The robust standard errors are clustered at the product level.

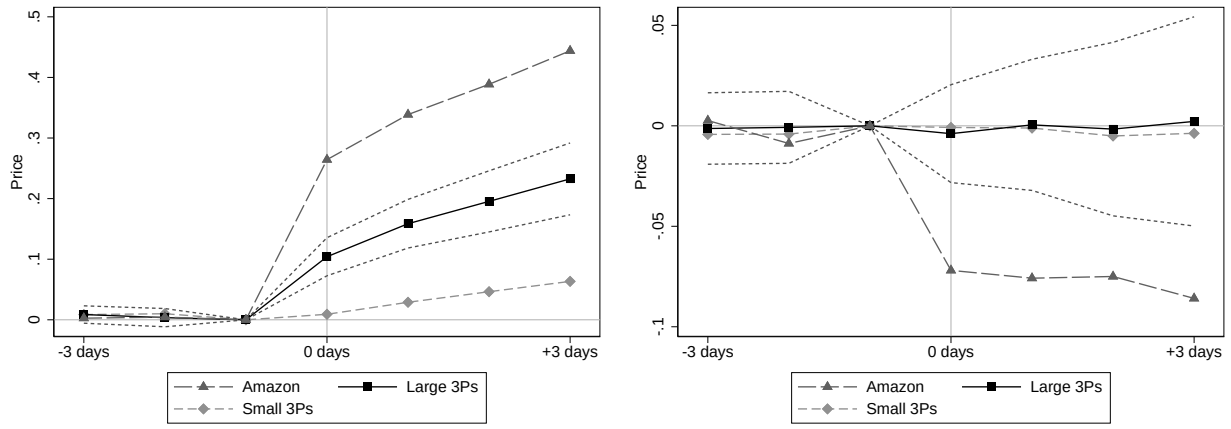
[25] This result does not imply Amazon's *equilibrium average* price is lower under information advantage.

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

[27] While competitors are more likely to raise their prices after their own sales, if there is price matching, we might expect third-party sellers to raise their prices after a competitor's sales. We examine the frequency of price increases and decreases in Online Appendix B.4. A small fraction of third-party sellers also raise their prices, but this indirect price response is minimal. Additionally, a tiny fraction of third-party sellers undercut, potentially based on imprecise sales rank information; thus, on average, we do not observe a change in the average price. This price simultaneity is explicitly modeled in our structural model.

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 Appendix E. While most repricers are a function of competitors' prices and own sales, none of them is contingent on competitors' sales, likely due to the lack of information.

Figure 4: Responses of Large Third-Party Sellers



(a) Own Sales Events

(b) Competitor's Sales Events

Note: Figure 4 presents the estimates from Equation 2, using the price level as the dependent variable for large third-party sellers before and after a sales event (Figure 4a) and a competitor's sales event (Figure 4b), alongside estimates for Amazon and small third-party sellers for comparison. The vertical line indicates the day of the event. Large sellers are in the top 10% of seller ratings, while small sellers are the remaining. The robust standard errors are clustered at the product 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 Figure 4. Specifically, we separately examine the response patterns of large and small third-party sellers, defined based on the number of seller ratings.[28] 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

[28] 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.

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 driven by information rather than the level of sophistication or technology.

4.4 Discussion

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

Buy Box. We compare buy box status changes three days before and after a competitor's sales event in Appendix F. If Amazon's recommendation algorithm steers demand in response to third-party sales, the information advantage could be amplified. However, we find that buy box changes are small for both Amazon and third-party sellers after competitors' sales events. Most importantly, all these variations are explained by prices, indicating that the buy box does not directly interact with the information advantage. This suggests that pricing and buy box algorithms are likely managed independently, as often occurs in practice (Hortaçsu et al., 2024). Based on this, we do not jointly model the buy box algorithm in this paper.

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 Appendix B, 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.

Tacit Collusion. Tacit collusion based on sales information is not a main concern to our study, as competitors can easily react to competitors' prices instantly (see Appendix E). Sales are stochastic and difficult to control precisely, and they are unlikely to be a primary coordination tool among competitors, especially given the transparency of prices. Furthermore, we examine our bi-hourly data, as used in Appendix C, and similar to the daily dataset, we did not find any pre-existing trend in prices before sales events. This suggests that the sales event is not confounded by pre-existing price events. Lastly, the comparison provided in Figure 4 between Amazon and large third-party

sellers does not support the notion that asymmetric price patterns are driven by the asymmetric incentives between dominant and small sellers.

5 Model

To motivate the study, we first solve a simple theoretical model of linear demand in a duopoly price competition. We find that the welfare results are mostly 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. Further discussion is provided in the Online Appendix A. To better understand the impact of information advantage and resolve the theoretical ambiguity, we develop and estimate an empirical structural model.[29]

5.1 Timing

Similar to Lee and Musolff (2021), we define a market m as a single product in period t, with listings from sellers $j \in J_{m,t} \subseteq \{1, 2\}$. Consumers have nested logit preferences and decide whether to buy from one of the two sellers or choose not to buy (i.e., the outside option). Seller j can have one of two types, defined as $a(j) \to \mathbb{A} = \{0, 1\}$, 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 as follows:

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.

[29] 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.

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

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

For market m in period t, the utility of an 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.[30] The seller-specific demand shock, $\xi_{m,j,t}$, changes over time (see Section 5.2.1 for a detailed discussion), 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 $g(j) = 0$.[31] 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

[30] Our choice model implicitly combines consumer choices with algorimic choices such as the buy box. As discussed in Section 4.4, we do not find that the buy box interacts with the information advantage, so we assume that, similar to consumer choice, buy box algorithms are fixed in the counterfactual. See Lee and Musolff (2021) and Raval (2022) for recent efforts to separate the buy box algorithm from consumer preferences.

[31] The seller-specific error structure in consumer utility reflects the sparse and dynamic nature of e-commerce sales. It comprises two components: (1) preferences for inside versus outside goods, influenced by product-level demand shocks, and (2) relative preferences between sellers, which can shift as consumers directed from different websites (e.g., TikTok vs. Facebook) bring varying shares of Prime and non-Prime users, along with differences in age, income, and shopping habits.

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. 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:

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

In the presence of demand uncertainty, the demand state $h_{m,j,t}$ is not perfectly observed by sellers until after time t, when the actual sales and market outcomes are realized. This means that sellers do not perfectly observe but instead hold a belief about the true demand level (based on historical information). 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 observed by sellers but not by researchers. This approach, combined with our full-solution estimation, directly addresses the endogeneity in prices and the issue of zero sales in high-frequency data.

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 them making a purchase, can be expressed as:

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

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 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 first-party sales and also collects commission fees from third-party sellers when their products sell. This commission fee is a percentage of the selling price and is considered a marginal cost for the third-party sellers. The profit function can be expressed as follows:

$$\Pi_{m,j,t} = \begin{cases} \big[\,(1 - r_m) \times p_{m,j,t} - mc_{m,j,t}\big] \times q_{m,j,t}, & \text{if } a(j) = 0, \\ r_m \times p_{m,-j,t} \times q_{m,-j,t} + \big(p_{m,j,t} - mc_{m,j,t}\big) \times q_{m,j,t}, & \text{if } a(j) = 1, \end{cases} \tag{5}$$

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 commission 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:

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

The supply shocks, denoted by $\boldsymbol{\varsigma}$, have a variance of σ_{ς}^2 and a mean of zero. These shocks are observable to sellers but are unobserved by researchers. Sellers engage in a simultaneous move pricing game to make their pricing decisions.[32] We consider the equilibrium Nash pricing in each period, allowing for a flexible belief regarding own and competitors' demand states.

5.3.1 Seller Belief

Based on the available information, each seller j forms estimates of the current-period demand states for the seller itself and for its competitor. The seller makes decisions to maximize their expected profits based on their beliefs. Similar to Equation 4, the seller's belief is written as follows:

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

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:

$$\hat{\xi}^{j}_{m,-j,t} = \varphi \times \hat{h}^{\text{Competitor}}_{m,j,t} + \nu_{m,-j,t}. \tag{7}$$

If all sellers had complete access to the information and the necessary technology, the best possible estimates for the current-period demand states would be denoted as $h^{\text{BP}}_{m,j,t}$ (see Section 6.1 for the detailed derivation). However, due to potential heterogeneity in sellers' information and calibration, some noise is introduced, and the estimates $\hat{h}$ may differ from h^{BP}. The model identifies better information access when firms' prices approximate the price they would set if they had accurate beliefs h^{BP}, and worse otherwise.[33] We later parametrize $\hat{h}$ and h^{BP} in the estimation (See Section 6.1 for the procedure and their parameter support).

[32] 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 E).

[33] One possible specification is to model $\hat{h}^{\text{Competitor}}_{m,j,t}$ as a function of $\hat{h}^{\text{Own}}_{m,j,t}$, allowing seller j's belief in the competitor's demand state to depend on their belief about their own state. However, since our primary interest is in measuring the deviation of belief relative to the best prediction, we choose a more direct approach.

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.

Model Discussion. We discuss several assumptions embedded within the supply model. First, FBA sellers incur an additional per-unit fulfillment fee on top of the standard commission fee. The per-unit fee is constant for each product and unlike commission 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.[34]

6 Estimation and Identification

We apply the structural model to the data and obtain model estimates. Given that sales at the listing level with high frequency can have a large fraction of zeros, it is not feasible to directly invert demand shocks. We employ a full-solution approach that simultaneously solves the demand and supply models, taking into account the unobserved demand and supply shocks. 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:

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

[34] Sellers may adopt pricing heuristics, assumed to align with profit-maximizing motives. This profit-maximizing assumption might be considered 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.

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.

True Demand States. As sales realizations are observed ex post, we use this information to construct the residual demand, $\hat{\epsilon}_{m,j,t}$, which represents the residuals that rationalize the sales data beyond the effects of prices and product-seller fixed effects (see Appendix G for details). We discretize the demand states $h_{m,j,t}$ as a higher state ($h_{m,j,t} = 1$) when the residual demands are greater than the median of the residuals, and a lower state ($h_{m,j,t} = 0$) otherwise. All medians are computed within the same market structure and seller types.

Best Predicted Demand States. Sellers can use past states to predict the current state, since demand exhibits autocorrelation (see Appendix D). Note that true demand states are estimated using ex-post sales, which are unavailable to sellers when setting price.

We use all past sales information to predict the demand sates (see Appendix H for details) and obtain discrete best-predicted demand states $h^{\text{BP}}_{m,j,t}$.[35] We define the best-predicted demand state as a higher state ($h^{\text{BP}}_{m,j,t} = 1$) when the predicted probability of a high demand state is greater than its median, and a lower state ($h^{\text{BP}}_{m,j,t} = 0$) otherwise. As before, all medians are computed within the same market structure and seller types.[36]

We find that the ex-post true demand state hm, 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 information is important in predicting the current demand state.

[35] Consistent with the empirical sales patterns, the estimated best predicted demand states for own $h^{\text{BP}}_{m,j,t}$ and competitors $h^{\text{BP}}_{m,-j,t}$ are correlated, as we use of both own and competitors' demand states to obtain the prediction.

[36] We use the term "best predicted demand states" to refer to demand information extracted from historical sales. It serves as a benchmark based on past sales against which we evaluate sellers' beliefs, and is not intended to be viewed as literal truths or ultimate predictions. In practice, we also consider and allow additional demand information through the error term $\nu_{m,j,t}$.

Seller Belief. Seller's beliefs are a continuous parameter between 0 and 1 and are estimated as a simple reduced-form function of h^{BP}, 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 $p^{\text{Own}}_{a(j)} = 0.5$, the seller perceives both states as equally likely, similar to a random guess. The parameter ϕ_0^{Own} represents the baseline values for all sellers, while Amazon has an additional parameter value of ϕ_1^{Own} on top of the baseline.

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

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

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. This belief also depends on the seller's identity $a(j)$. $\phi_0^{\text{Competitor}}$ 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, $\phi_1^{\text{Competitor}}$, which is specific to Amazon, is expected to be positive in addition to the baseline.

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:

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

The dataset for estimation includes the sales and prices of the repricing seller and its 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 also include a cost shock, denoted as ω_{Post}, for the period after the repricing event to account for cost-side explanations of price changes.

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:

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

We allow each of these market types to have different utility intercepts, marginal costs, variances of demand shocks, and variances of cost shocks.[37] 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.

Inner Loop. Let the vector of demand and supply primitives be $\boldsymbol{\theta} = \big\{\boldsymbol{\theta}^D, \boldsymbol{\theta}^S\big\}$. 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. We also set $\lambda_{m,t}$ to a large value (i.e., 20), and our estimates are robust with various different values.

There are 16 possible states resulting from the combination of a seller's and its 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. We solve equilibria for 12,960 pricing games during each iteration.

[37] 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.

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 (see Appendix I for a detailed breakdown of these vectors):

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

In total, 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 the discussion on how price elasticity is identified. To do this, we take advantage of the high-frequency data available to us. Our assumption is that, during this brief period around repricing events, 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.

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. On the other hand, supply shocks are independent of residual demand but still contribute to price shifts. After determining the distribution of demand shocks, the variance of supply shocks rationalizes the distribution of prices, which we capture by employing the square of price as a moment.

The parameter φ is identified by the difference in average sales under high demand states relative to low demand states. Given the value of φ, we infer sellers' beliefs from their pricing levels (e.g., Aguirregabiria, 2021) specifically, the belief 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 $\phi_0^{\text{Competitor}}$.

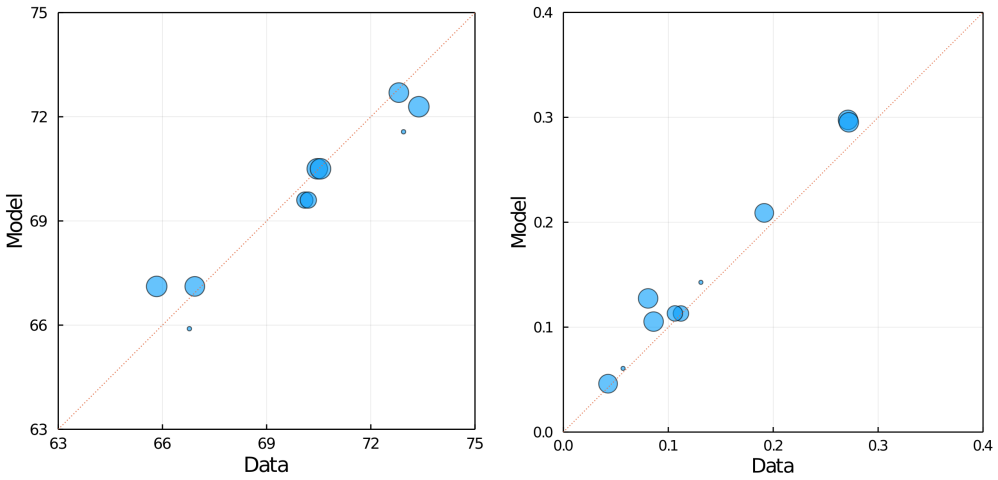
7 Results

In Figure 5, we assess the fit of our model by comparing the empirical and predicted prices and sales. The model adequately explains the sales and prices observed in the data.

Columns 1 to 3 of Table 2 display the demand estimates. The average own-price elasticity is -18.96, a magnitude consistent with previous literature (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. Based on the estimates, 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 than

Figure 5: Moment Fit



(a) Price Moment

(b) Sales Moment

Note: Figure 5 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 5a, while the fit of the sales moment is presented in Figure 5b.

FBM sellers. Note that we normalize the price to \$70 in the model; the estimated premiums over FBM sellers are \$7.15 and \$6.12 respectively, and 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 2 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 compared to FBA sellers. This difference help rationalize the empirical observations, where FBA prices are, on average, more competitive than Amazon prices (\$41.72 for FBA vs. \$66.89 for Amazon) in the data.

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 $p_0^{\text{competitor}} = 0.52$.

Overall, our structural estimations 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.

7.1 Counterfactual Simulations

We conduct two counterfactual analyses of symmetric information, corresponding to proposals Amazon put forward in response to EU antitrust investigations, to examine the impact of information advantage on price competition. These proposals involve (1) preventing Amazon from using marketplace data for its own retail business and (2) sharing marketplace data with third-party sellers. To highlight the role of the vertically integrated market structure, we also vary the commission rate and examine how it mediates the effect. Detailed information regarding the welfare calculation can be found in Appendix J.

7.1.1 Restricting Information Advantage

In the first counterfactual scenario, Amazon loses its information advantage as it is prevented from using marketplace data for its pricing decisions.[38] To simulate this, we set Amazon's belief 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 3 show that, under the current 15% commission rate and restricted information access for Amazon, Amazon's sales decrease and profits decrease by 0.48% and 0.30%, respectively, while third-party sellers experience a 3.10% increase in sales and a 3.04% profit gain. While more information increases efficiency, superior information also grants market power, allowing the advantaged party to better discern when to raise prices. This leads to an anti-competitive effect that outweighs the efficiency gains. Consequently, by removing the information

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

Table 2: 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 2 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, while we normalize price by the mean price of each product and set the average price across the sample to be \$70. 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.

advantage, and thus the market power associated with it, the average price decreases, leading to improvements in consumer welfare and social welfare by 0.45% and 0.35%, respectively.[39]

We further emphasize the role of vertical market structure in our findings by examining the scenario where Amazon does not charge the commission fee, thereby eliminating vertical incentives.[40]

Columns 4 to 6 of Table 3 display the changes in equilibrium outcomes when Amazon has no information advantage, compared to the case where Amazon has an information advantage under a 0% commission rate. Amazon's sales and profits drop significantly by 4.22% and 3.71%, respectively. On the other hand, third-party sellers gain more in sales and profits, increasing by 9.76% and 10.11%, respectively. The average increases in consumer welfare and social welfare across the two markets are approximately 0.50% and 0.78%, respectively. The larger effect under a 0% commission rate compared to a 15% commission rate suggests that, in the absence of vertical incentives, market power stemming from informational advantage may pose a more significant concern. We also present consistent findings when commission rate changes continuously, which further shape vertical incentives, as detailed in Section 7.1.3.

7.1.2 Sharing Information

In the second counterfactual, we explore a scenario where Amazon shares its marketplace information with third-party sellers, thereby eliminating their information disadvantage. The shared information is non-personal, so privacy is not a major concern. We simulate this situation by providing third-party sellers with the same access to belief parameters as Amazon. Specifically, we set $p_0^{\text{competitor}} = p_1^{\text{competitor}}$.[41]

Columns 1 to 3 of Table 4 present the results for this counterfactual with the current 15% commission rate remains unchanged. There is a notable 19.78% increase in sales for third-party sellers, while Amazon's sales decline by 4.58%. Additionally, information sharing leads to meaningful increases in consumer welfare and social welfare, with average increases of 2.41% and 2.01%,

[39] In theory, the equilibrium price of the advantaged party may increase or decrease when the information advantage is removed. For instance, in the AMZ-vs-FBA market, where Amazon may possess less market power due to the cost difference, the average price increases.

[40] 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 platform-Amazon from retailer-Amazon).

[41] 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.

Table 3: Counterfactual Analysis of Restricting Superior Information

	15% Commission Rate			0% Commission Rate		
	AMZ-vs-FBM	AMZ-vs-FBA	AMZ-Average	AMZ-vs-FBM	AMZ-vs-FBA	AMZ-Average
Panel A: Equilibrium Outcomes						
%Δ Amazon Price	-0.29	0.33	-0.22	-0.43	1.97	-0.19
%Δ Amazon Sales	0.08	-7.86	-0.48	-1.69	-33.93	-4.22
%Δ 3rd-Party Price	-0.05	-0.04	-0.05	-0.18	-0.63	-0.22
%Δ 3rd-Party Sales	2.37	5.24	3.10	6.19	18.06	9.76
Panel B: Profit and Welfare Change						
%Δ Amazon Profit	-0.30	-0.34	-0.30	-2.18	-29.73	-3.71
%Δ 3rd-Party Profit	1.95	6.53	3.04	5.99	19.99	10.11
%Δ Consumer Welfare	0.58	-0.58	0.45	0.46	0.74	0.50
%Δ Social Welfare	0.30	0.69	0.35	0.33	3.42	0.78

Note: Table 3 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% commission rate, and Columns 4 to 6 correspond to the case when the commission rate is reduced to 0%.

respectively.

With more information being shared, prices better reflect demand states, improving the matching quality between supply and demand. This results in a slight decrease in average prices but a significant increase in overall sales. Consumers and third-party sellers benefit, while Amazon indirectly gains from increased commission fees due to market expansion.[42] Columns 4 to 6 of Table 4 present the results under a 0% commission rate. In this scenario, Amazon's average profit decreases by 4.78% due to the loss of first-party sales, with no additional commission fees being collected.

Comparing the two counterfactual scenarios of symmetric information, achieved either by restricting the information advantage or by sharing information, shows that information sharing leads to a greater gain in social welfare; higher sales occur despite a smaller average price reduction compared to the restriction case, as the increased information improves the match between supply and demand.

Markets with Only Third-Party Sellers. Information sharing among third-party sellers can also influence markets where these sellers compete with one another, and Amazon does not function as a seller. The counterfactual results for these markets are discussed in detail in Appendix K. Appendix L presents back-of-the-envelope calculations on how the effect may aggregate across all market types.

7.1.3 The Role of Vertical Incentives

To further understand the discussions in Section 7.1.1 and Section 7.1.2, we examine the interaction between competition incentives—shaped by local variations in vertical incentives—and different information structures. The commission rate, which can vary continuously, serves as a natural instrument for vertical incentive variations. As the commission rate increases, Amazon and third-party sellers become more aligned, leading to reduced price competition. Figure 6 compares various measures representing Amazon, third-party sellers, consumers, and the social planner's payoff changes under different counterfactual information structures, with commission rates varying continuously from 0% to 30%.

[42] This finding highlights the role of commission rates in aligning the owner's incentives to share information and is consistent with recent observations regarding Amazon's information sharing (see Appendix A).

Table 4: Counterfactual Analysis of Sharing Superior Information

	15% Commission Rate			0% Commission Rate		
	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	-0.19	-0.01	-0.94	-0.92	-0.94
%Δ Amazon Sales	-1.80	-27.35	-3.58	-0.43	-24.78	-2.33
%Δ 3rd-Party Price	-0.03	-0.02	-0.03	-1.02	-0.64	-0.99
%Δ 3rd-Party Sales	12.81	40.10	19.78	16.96	34.50	22.23
Panel B: Profit and Welfare Change						
%Δ Amazon Profit	0.08	4.75	0.60	-3.33	-29.44	-4.78
%Δ 3rd-Party Profit	2.72	19.63	6.76	2.02	11.47	4.81
%Δ Consumer Welfare	1.38	10.29	2.41	4.55	15.10	6.17
%Δ Social Welfare	0.91	9.69	2.01	1.70	9.09	2.79

Note: Table 4 presents the counterfactual analysis when Amazon shares its superior 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% commission rate, and Columns 4 to 6 correspond to the case when the commission rate is reduced to 0%.

Amazon. Figure 6a shows Amazon's profit under different information structures and vertical incentives. While Amazon is consistently worse off with restricted information compared to having an information advantage, this effect diminishes as vertical incentives increase. Amazon is not always worse off from sharing information; its payoff also increases with higher vertical incentives. When commission rates are higher than 11%, Amazon can benefit from sharing information due to commission fee revenue.

Third-Party Sellers. In Figure 6b, third-party sellers consistently benefit more from shared information than from information asymmetry, especially with increasing vertical incentives. While third-party sellers generally benefit from restricting information, this gain diminishes with higher vertical incentives. At sufficiently high commission rates, as the market becomes less competitive, restricting information advantage can reduce third-party sellers' profits compared to having it.

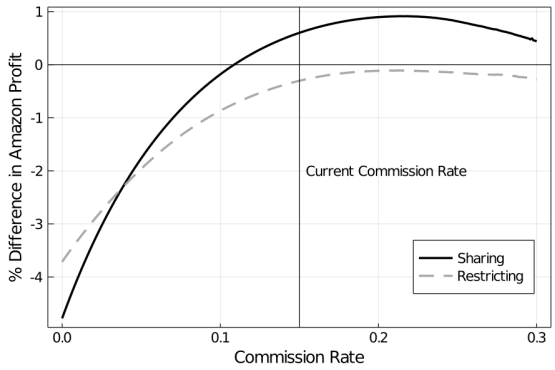
Consumers. As shown in Figure 6c, consumers are consistently better off when Amazon's information advantage is restricted, and this benefit is relatively inelastic to commission rate changes. Consumer gain usually is higher from sharing information than from restricting information, but it diminishes as the commission rate increases. When the commission rate exceeds 25%, information sharing can make consumers worse off, leading them to prefer restricting information.

Social Planner. Social welfare generally increases with more symmetric information but diminishes as the commission rate increases, as shown in Figure 6d. At very high commission rates, social welfare may decrease if Amazon's information is restricted, as the loss in information efficiency (leading to decreased third-party sellers' profits) outweighs the gains in consumer welfare. Additionally, sharing information always results in higher social welfare compared to restricting it.

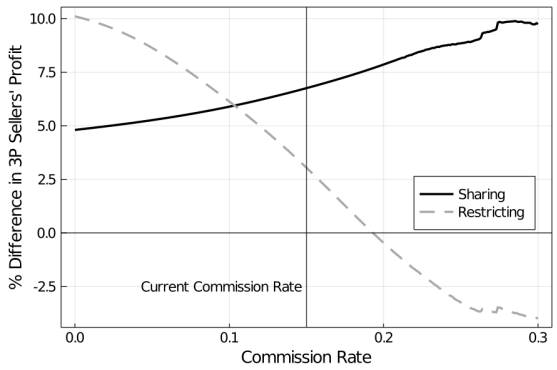
Endogenous Commission Rate. Our model abstracts from the endogenous commission rate, as it is not the primary focus. Empirically, Amazon is known for rarely changing them. With this caveat, we explore how our results interact with a hypothetical scenario where Amazon reoptimizes the commission rate. Figure 7a and Figure 7b display Amazon's profit and social welfare across various commission rates, ranging from 0% to 30%, under different information structures. We then calculate the optimal commission rate that maximizes Amazon's profits.

Under the current regime, where Amazon has an information advantage, the optimal commission rate implied by our estimates is 24.0%. This number is reasonable in the context of general platform

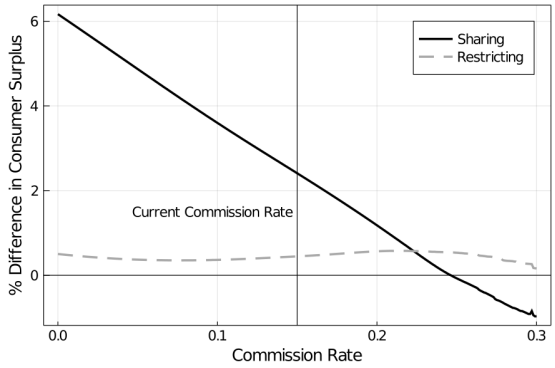
Figure 6: Interaction Between Vertical Incentives and Information Structure



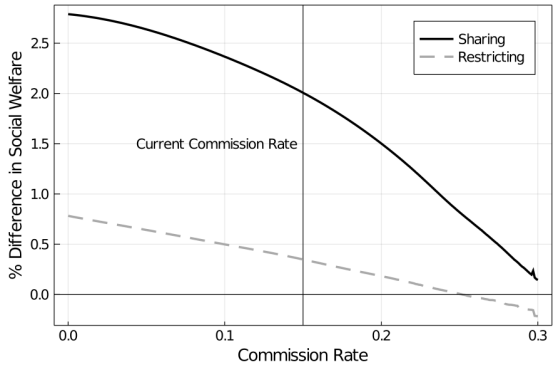
(a) % Change in Amazon Profit



(b) % Change in 3P's Profit



(c) % Change in Consumer Surplus

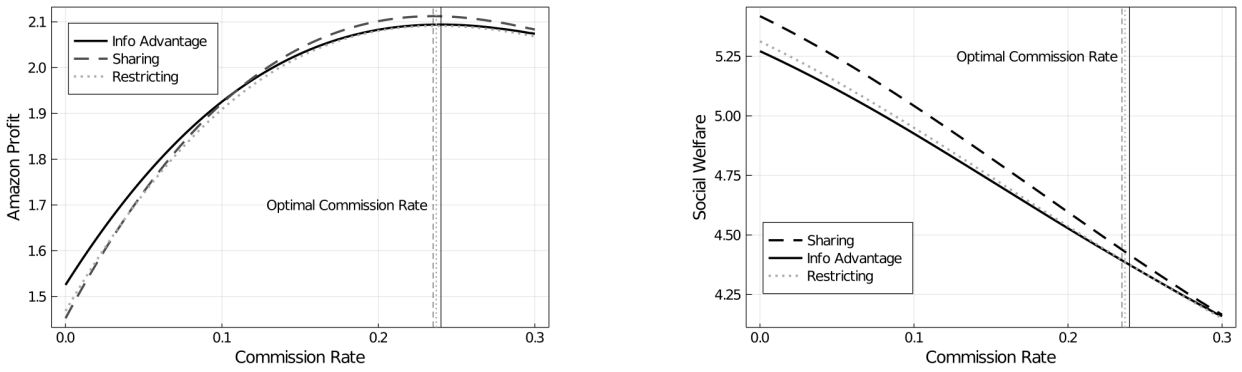


(d) % Change in Social Welfare

Note: Figure 6 illustrates the percentage difference in measures along with commission rates from 0% to 30% under two counterfactual simulations relative to the case with an information advantage. The outcomes are averaged across AMZ-vs-FBM and AMZ-vs-FBA markets, weighted by the number of observations under different information structures. The vertical line marks the current commission rate of 15%. Figure 6a, Figure 6b, Figure 6c, and Figure 6d show Amazon's profit, third-party profit, consumer surplus, and social welfare, respectively.

commission rates but is admittedly higher than Amazon's current commission rate. When Amazon shares information, the optimal commission rate is 23.5%, while it is 23.7% when its information is restricted. The magnitude of these changes is not significant, but the direction is notable, suggesting that symmetric information would lead to a lower commission rate if reoptimized. Therefore, the positive welfare impact of sharing information or restricting Amazon's information advantage would be amplified, as a lower commission rate reduces double marginalization and increases social welfare (see Figure A.3 for a clearer visualization).

Figure 7: Optimal Commission Rate and Information Structure



(a) Amazon Profit

(b) Social Welfare

Note: Figure 7 illustrates Amazon's profit and social welfare along with commission rates from 0% to 30% under different information structures. The outcomes are averaged across AMZ-vs-FBM and AMZ-vs-FBA markets, weighted by the number of observations. Figure 7a displays Amazon's total profit under three information structures, with vertical lines indicating the optimal commission rates that maximize Amazon's profit: a solid line at 24.0% for the current information advantage regime, a dashed line at 23.5% for sharing information, and a dotted line at 23.7% for restricting information. Figure 7b shows social welfare under the same three information structures, with corresponding optimal commission rates marked by vertical lines.

8 Conclusion

This research examines the impact of asymmetric information access on price competition within a vertically integrated market structure. We identify empirical patterns of supply-side information asymmetry, where individual sellers gain insights into demand through their own sales, while the platform owner accesses competitor sales information, allowing it to better understand demand and set prices. We estimate a structural model to quantify the importance of this information asymmetry. Counterfactual analysis shows a significant impact of symmetric information on equilibrium outcomes and efficiency, highlighting information advantage as a potential source of monopolistic market power. In addition, the impact of the information structure interacts with vertical incentives.

Our findings highlight the competitive effects of information access in a vertical market structure, which is relevant for competition policy, particularly in evaluating anti-competitive effects of information asymmetry in vertical merger guidelines and addressing concerns about monopolizing data-driven competitive advantages in digital markets. In today's data-driven world, there are

many types of information with wide-ranging applications across contexts; for example, search data can be used to understand consumer preferences. We note that the information-advantage-induced market power we quantify likely represents a lower bound, as the platform owner may have more information than sales, for example, personalized data and search data, which may also contribute to a greater extent of information advantage.

References

Acemoglu, D. (2021). Harms of ai. Technical report, National Bureau of Economic Research.

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.

Arrow, K. J. (1975). Vertical integration and communication. *The Bell Journal of Economics*, pages 173–183.

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*.

Cabral, L., Haucap, J., Parker, G., Petropoulos, G., Valletti, T. M., and Van Alstyne, M. W. (2021). The eu digital markets act: a report from a panel of economic experts. *Cabral, L., Haucap, J., Parker, G., Petropoulos, G., Valletti, T., and Van Alstyne, M., The EU Digital Markets Act,*

Publications Office of the European Union, Luxembourg.

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.

Crawford, G. S., Courthoud, M., Seibel, R., and Zuzek, S. (2022). Amazon entry on amazon marketplace. *CEPR Discussion Papers*, (17531).

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.

Hortaçsu, A., Natan, O. R., Parsley, H., Schwieg, T., and Williams, K. R. (2024). Organizational structure and pricing: Evidence from a large us airline. *The Quarterly Journal of Economics*, 139(2):1149–1199.

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. (2024). Insider imitation. *Journal of Political Economy*.

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.

U.S. Department of Justice and The Federal Trade Commission (2020). Vertical merger guidelines. Accessed: 2024-09-06.

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.

Online Appendix

The appendix is intended for online publication.

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, and as of October 2023, it still appears to be 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.[43]

B 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 A.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 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

[43] 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). This seems to fit in a plausible trend of Amazon sharing more data with third-party sellers.

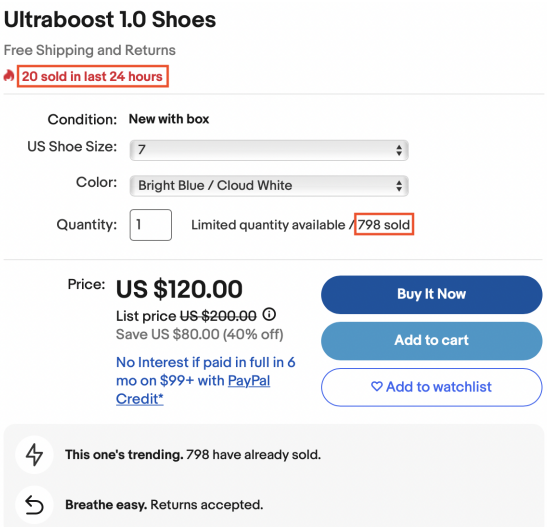
Table A.1: Disclosure of Sales Data by E-commerce Websites

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

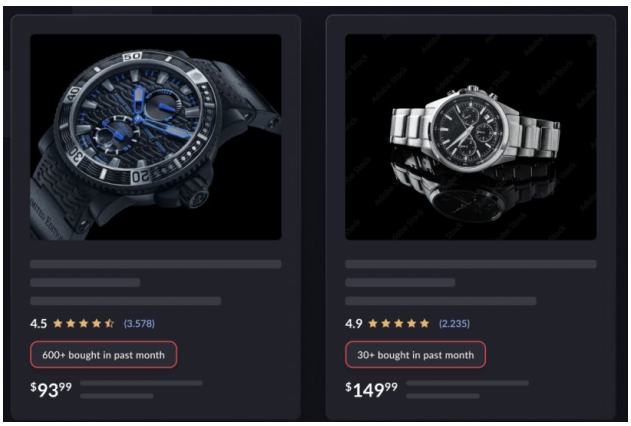
Note: Table A.1 lists the 19 e-commerce websites, their primary operating regions, and whether these websites disclose sales measures or not.

more likely to adjust its prices and make larger changes.

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

C Sales Error Probability from Restocking

Restocking of products presents a potential source of measurement error in our sales data. We track the stock levels of approximately 2,400 products on an bi-hourly basis. Using the high frequency data, we can more effectively differentiate between sales occurrences and product restocking. 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

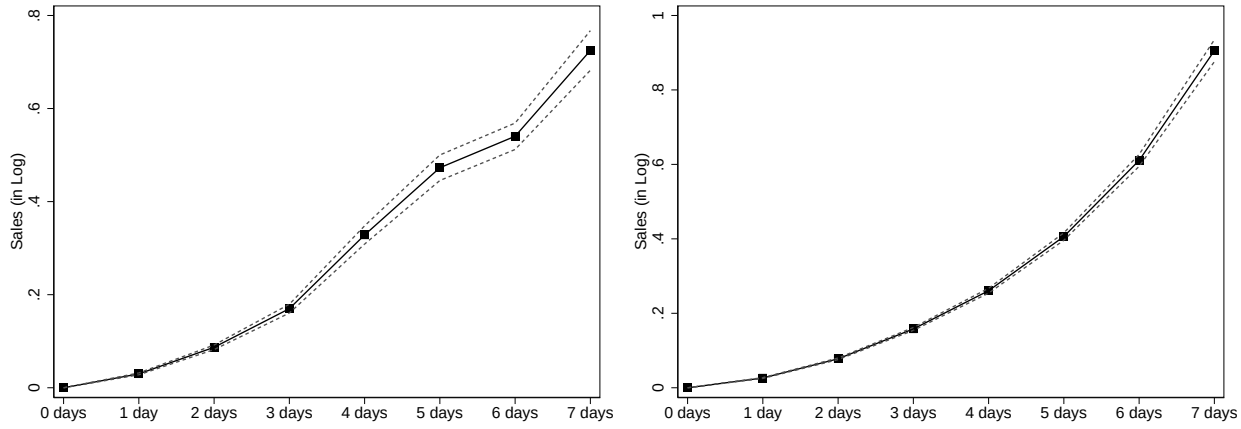
Stock Data Frequency	Probability (%)
1 Day	0.98
2 Days	2.15
3 Days	3.12
4 Days	4.14
5 Days	5.22
6 Days	6.06
7 Days	7.12

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

D Autocorrelated Demand

The estimates of γ_τ using sales $q_{m,j,t}$ as the dependent variable are displayed in Figure A.2, 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.2: Sales Over Number of Days With Sales During The Past Week



(a) Amazon

(b) Third-Party Sellers

Note: Figure A.2 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.

E Automated Pricing Services

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."[44] 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.

Furthermore, we conduct a survey of leading providers of automated pricing services in the market.[45] For each of these providers, we summarize the data used in their technology in Columns 2 and 3 of Table A.4. apart from considering competitors' prices, most automated pricing service providers offer adjustments based on sellers' own sales, often referred to as "Sales Velocity Repricing." However, none of the companies claim to possess the capability to set prices based on competitors' sales.

Column 4 of Table A.4 presents our survey results regarding the repricing frequency offered by each provider. With the exception of one provider primarily offering A/B testing services, the rest

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

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

claim to reprice instantaneously. In practice, this means repricing is typically completed within seconds, 2 to 3 minutes, or as fast as Amazon allows.

Table A.4: 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.4 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.

F Buy Box Changes Following Competitor Sales

We compare the change in buy box winning probability three days before and after a competitor's sales event using the following:

$$B_{it} = \beta \times \mathbb{1}(\tilde{t} > 0) + \psi_s + \epsilon_{s,\tilde{t}},\ \forall\ s \in \mathbb{S}_3^{\text{sales}} \wedge \tilde{t} \in \{-3, -2, -1, 1, 2, 3\}. \tag{A.1}$$

where $\mathbb{1}(\tilde{t} > 0)$ indicates whether the observation is after the competitor sales event. Equation A.1 compares the probability of the focal seller's buy box status before and after three days of competitor's sales.[46]

Column 1 of Table A.5 shows the third party's buy box status when their competitors make sales. The buy box changes by a small magnitude after competitor sales. Column 2 of Table A.5 controls for the log of prices of sellers in the market and includes fixed effects for the price gap

[46] We exclude the date of the sales event because it is likely to occur when the the buy box features the competitor.

(defined as $Price_j - Price_{-j}$). We find that price changes explain all the variation in buy box status for third-party sellers.

Column 3 of Table A.5 shows Amazon's buy box status when competitors make sales. Amazon is more likely to hold the buy box after competitors make sales; however, similar to the third-party seller case in Column 2, these changes are mostly driven by price adjustments, as indicated in Column 4 of Table A.5.

Table A.5: Buy Box Changes Following Competitor Sales

	Third-Party Sellers		Amazon	
$\mathbb{1}(\tilde{t} > 0)$	0.001***	0.000	0.004***	0.001
	(0.000)	(0.000)	(0.001)	(0.001)
$\log(p_j)$		-0.367***		-0.253***
		(0.028)		(0.043)
$\log(p_{-j})$		0.188***		0.513***
		(0.021)		(0.051)
Mean of Dep. Var	0.398	0.398	0.807	0.807
Event Fixed Effects	✓	✓	✓	✓
Price Gaps		✓		✓
No. of Observations	1,229,651	1,228,470	282,548	281,483
Adjusted R-squared	0.773	0.788	0.835	0.852

Note: Table A.5 displays the estimates from Equation A.1 using whether the seller is the featured seller (buy box) as the dependent variable. $\mathbb{1}(\tilde{t} > 0)$ indicates the periods after a competitor's sales event. The variable p_j denotes the focal seller's prices, while p_{-j} represents the prices of their competitors. The robust standard errors are clustered at the product level. Significance levels: *(p < 0.10), **(p < 0.05), ***(p < 0.01).

G Estimation of 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:[47]

$$q_{m,j,t} = \eta \times p_{m,j,t} + \psi_{m,j} + \epsilon_{m,j,t}. \tag{A.2}$$

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

We include the product-seller fixed effects, denoted by $\psi_{m,j}$, 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 address the potential bias in the coefficient of η caused by the endogeneity of prices when estimating Equation A.2, 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.6, we present the results of the first-stage regression, which indicate that prices tend to increase with higher log inventory levels. In Column 2 of Table A.6, we provide the estimates from the IV regression. As expected, the coefficient of the price is negative.

Formally, we define the demand states as follows, using the residuals $\hat{\epsilon}$ obtained from Equation A.2:

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

H Estimation of Best-Predicted Demand States

To estimate the ex-ante best predicted demand states, we use past demand states of the sellers' own sales and those of their competitors:

$$h_{m,j,t} = \sum_{\tau=-T}^{-1} \kappa_{\tau,j} \times h_{m,j,\tau} + \sum_{\tau=-T}^{-1} \kappa_{\tau,-j} \times h_{m,-j,\tau} + \epsilon_{m,j,t}. \tag{A.3}$$

In practice, we choose $T = -7$, which corresponds to using demand states from the previous week. Extending the timeframe beyond a week has minimal effect on prediction accuracy. Similarly, we discretize the best predicted demand states. Formally, we define the best-predicted demand states using the fitted values $\hat{h}$ obtained from Equation A.3:

$$h^{\text{BP}}_{m,j,t} = \begin{cases} 1, & \text{if } \hat{h}_{m,j,t} > \text{median}\left(\hat{h}_{m',j',t} \mid \mathcal{M}(m') = \mathcal{M}(m), a(j') = a(j)\right), \\ 0, & \text{if } \hat{h}_{m,j,t} \leq \text{median}\left(\hat{h}_{m',j',t} \mid \mathcal{M}(m') = \mathcal{M}(m), a(j') = a(j)\right). \end{cases} \tag{A.4}$$

Note that in practice $h^{\text{BP}}_{m,j,t}$ and $h^{\text{BP}}_{m,-j,t}$ are correlated, given the structure described in Equation A.3 and Equation A.4.

Table A.6: Estimation of Demand States

	(1)	(2)
	First Stage	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.6 shows the results from Equation A.2. 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).

I 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:

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

$$\boldsymbol{H}_{m,j,\tilde{t}} = \begin{bmatrix} \mathbf{1}(\hat{h}_{m,j,t} = 1) \\ \mathbf{1}(\hat{h}_{m,j,t} = 0) \\ \mathbf{1}(\hat{h}_{m,-j,t} = 1) \\ \mathbf{1}(\hat{h}_{m,-j,t} = 0) \end{bmatrix} \otimes \begin{bmatrix} \mathbf{1}(h^{\text{BP}}_{m,j,t} = 1) \\ \mathbf{1}(h^{\text{BP}}_{m,j,t} = 0) \\ \mathbf{1}(h^{\text{BP}}_{m,-j,t} = 1) \\ \mathbf{1}(h^{\text{BP}}_{m,-j,t} = 0) \end{bmatrix}, \text{and } \boldsymbol{T}_{m,j,\tilde{t}} = \begin{bmatrix} \mathbf{1}(\tilde{t} = 1) \\ \mathbf{1}(\tilde{t} = -1) \end{bmatrix}.$$

J Welfare Calculation

For each scenario, we calculate sellers' profits based on Equation 5. We compute consumer welfare using the following equation:

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

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 commission fees in the calculation of social welfare. The formula is as follows:

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

K 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. 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 outcomes for this scenario are delineated in Table A.7.[48] 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 commission 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.

When competitors' sales information is available, FBA-vs-FBA and FBM-vs-FBM markets become more competitive, 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.

[48] Amazon, as the platform owner, earns profits from commission fees without engaging in direct market participation. In contrast to markets involving Amazon, vertical incentives are less significant because the commission rate primarily serves as a static modifier to sellers' marginal costs.

Overall, the results suggest 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.

Table A.7: Information Sharing with Third-Party Sellers, Third-Party-Only Markets

	15% Commission Rate			0% Commission Rate		
	FBA-vs-FBM	FBA-vs-FBA	FBM-FBM	FBA-vs-FBM	FBA-vs-FBA	FBM-FBM
Panel A: Equilibrium Outcomes						
%Δ FBA Price	1.09	-3.30		1.27	-3.94	
%Δ FBA Sales	0.70	34.99		1.44	33.74	
%Δ FBM Price	0.07		-0.83	-0.19		-0.78
%Δ FBM Sales	-16.48		2.22	-14.58		2.08
Panel B: Profit and Welfare Change						
%Δ Amazon Profit	-2.15	27.93	1.56	0.00	0.00	0.00
%Δ 3rd-Party Profit	2.18	-24.53	-1.67	2.31	-25.60	-1.53
%Δ Consumer Welfare	-2.35	36.17	2.51	-1.66	36.13	2.26
%Δ Social Welfare	-1.14	18.95	0.89	-0.02	11.34	0.45

Note: Table A.7 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% commission rate, and Columns 4 to 6 correspond to the case when the commission rate is reduced to 0%.

L Counterfactual: All Types of Markets

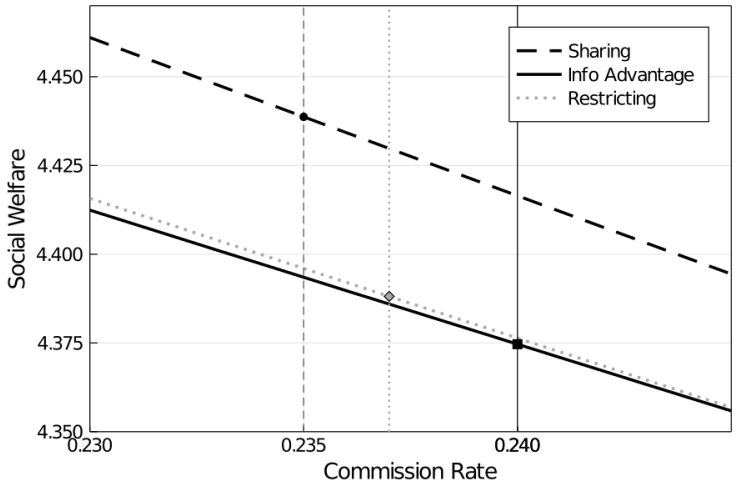
We perform back-of-the-envelope calculations on how the effect may aggregate across all market types. When considering all market types, including those with only third-party sellers, in our dataset and using their empirical weights, the implementation of information sharing leads to a significant improvement in both consumer welfare and social welfare, with respective increases of

4.02% and 2.39% when the commission rate 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. However, we suggest interpreting these numbers with caution, mindful of the numerous simplifications made in our model. Among these is the fact that Amazon carries a wide array of products. How multiproduct pricing interacts with informational advantage and impacts Amazon's market power could present an interesting avenue for future research.

M Counterfactual: Re-optimization of Commission Rate

Figure A.3: Commission Rate Re-optimization in Response to Changes in Information Structures



Note: This figure is a zoomed-in version of Figure 7b, showing social welfare under the three information structures, with optimal commission rates marked by vertical lines.

Online Appendix II

The appendix is intended for online publication.

A Theoretical Model

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

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

Assuming that firm 1 charges a commission 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.

A.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: $CW^A \coloneqq CW_1^A + CW_2^A$.

A.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:

$$p_1^P = \frac{2A_1a_2 - A_1k_2(b_1 + rb_2) + A_2(b_1 - 2a_2k_1 + rb_2)}{4a_1a_2 - b_2(b_1 + rb_2)} = p_1^A + \underbrace{\frac{(\eta_{A_1} - A_1)k_2(b_1 + rb_2)}{4a_1a_2 - b_2(b_1 + rb_2)}}_{\text{Increased price competition}}$$
$$\coloneqq p_1^A + \Delta p_1^{A,P},$$
$$p_2^P = \frac{-2a_1A_2 - A_1b_2 + A_2b_2k_1 + 2a_1A_1k_2}{-4a_1a_2 + b_2(b_1 + rb_2)} = p_2^A + \underbrace{\frac{2a_1(\eta_{A_1} - A_1)k_2}{4a_1a_2 - b_2(b_1 + rb_2)}}_{\text{Improved pricing response to competitors' information}} \coloneqq p_2^A + \Delta p_2^{A,P}.$$

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:

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

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

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

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 1a. 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 commission 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 1a represents the profit change for other k_2 values. The expected profit for Firm 2 can be written as follows:

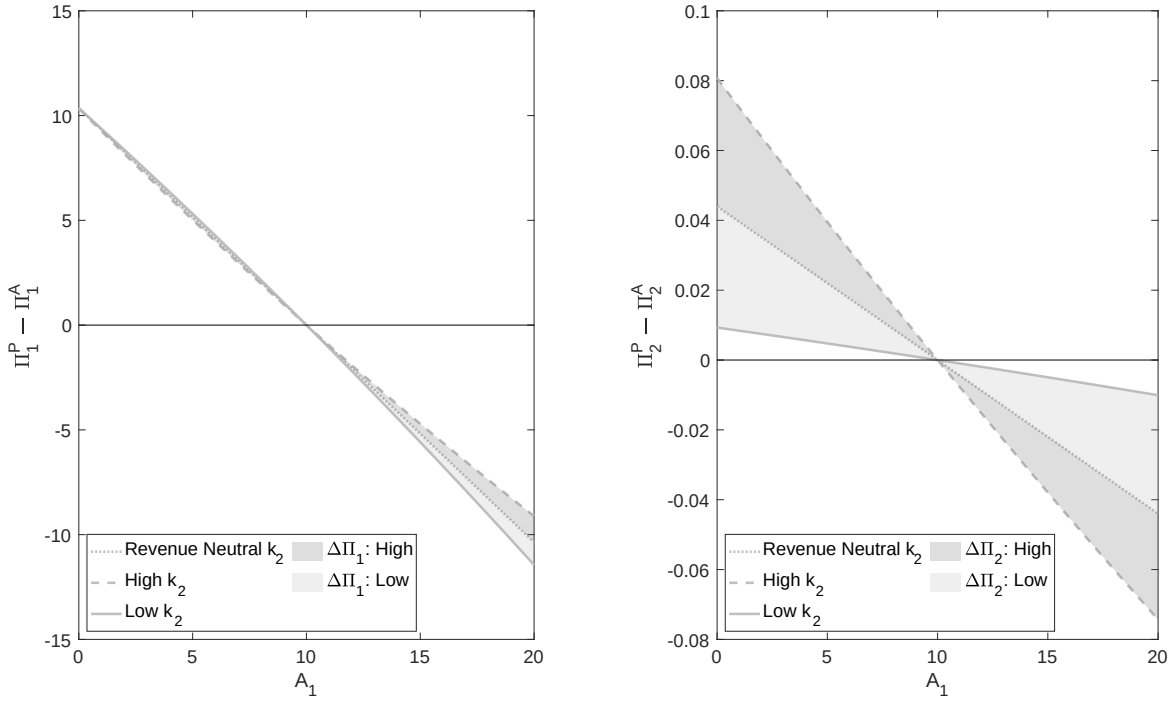
$$\begin{aligned}
\mathbb{E}[\Pi_2^P] &= \mathbb{E}[\Pi_2^A] + \int_{\mathcal{A}_1}\int_{\mathcal{A}_2} \left(p_2^A \Delta q_2^{A,P} + \Delta p_2^{A,P} a_2 p_2^P\right) dF(A_2) dF(A_1) \\
&= \mathbb{E}[\Pi_2^A] + \frac{k_2 \left(\overbrace{4a_1^2a_2k_2}^{\text{Gain from better pricing}} - \overbrace{b_2^2(b_1 + rb_2)}^{\text{Loss from intensified competition}} \right)}{\left(4a_1a_2 - b_2(b_1 + rb_2)\right)^2} \left(\int_{\mathcal{A}_1} (A_1 - \eta_{A_1}) A_1 dF(A_1) \right) \\
&> \mathbb{E}[\Pi_2^A] \iff k_2 > \frac{b_2^2(b_1 + rb_2)}{4a_1^2a_2}.
\end{aligned}$$

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 1b 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 1: Effect of Variable k_2 on Profits



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

Note: Figure 1 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:

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

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 2a, 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.

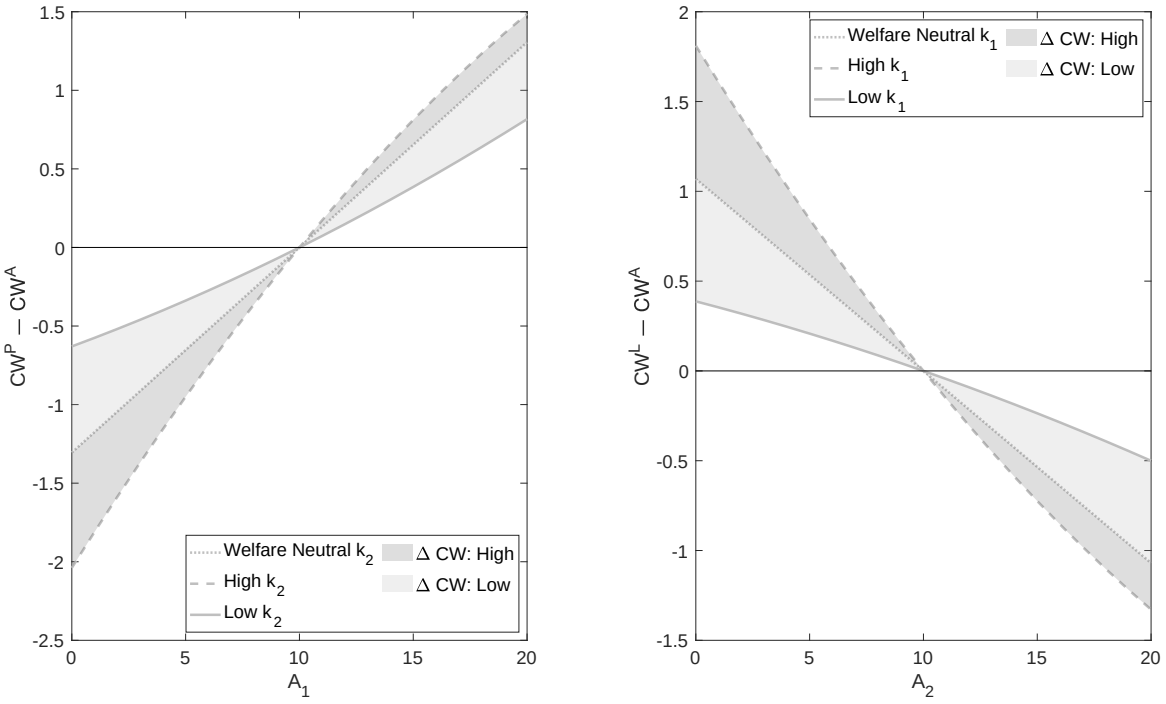
A.3 Limited Information

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:

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

Firm 1 no longer undercuts Firm 2 when its demand state is high (High A_2) but lowers 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.

Figure 2: Effect of Variables k_1 and k_2 on Consumer Welfare



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

Note: Figure 2 displays a simulation with parameters $a_1 = a_2 = 8, E[A_1] = E[A_2] = 10$, and $r = 15\%$. Additionally, Figure 2a has $b_1 = 1$, $b_2 = 5$, $k_1 = 0.2$, and $A_2 = 10$, while Figure 2b 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.

The quantity meets the following conditions:

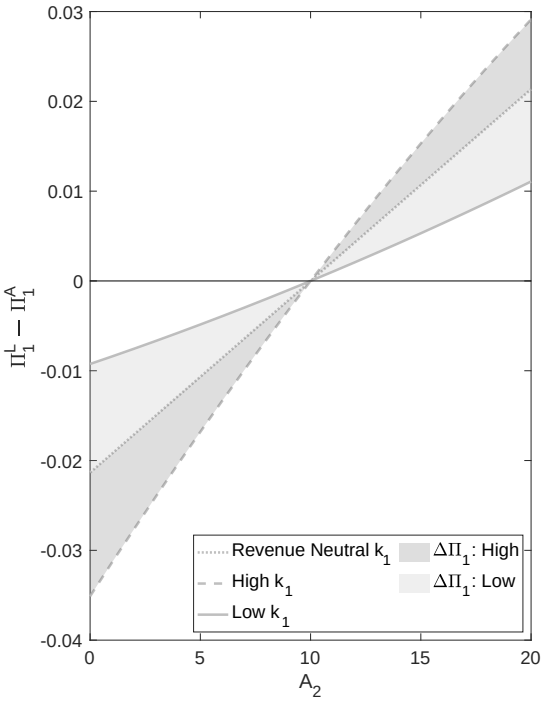
$$q_1^L = q_1^A + \frac{(2a_1a_2 - b_1b_2)(\eta_{A_2} - A_2)k_1}{4a_1a_2 - b_2(b_1 + rb_2)} := q_1^A + \Delta q_1^{A,L},$$
$$q_2^L = q_2^A + \frac{a_2b_2(A_2 - \eta_{A_2})k_1}{4a_1a_2 - b_2(b_1 + rb_2)} := q_2^A + \Delta q_2^{A,L}.$$

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

$$\begin{aligned}\mathbb{E}[\Pi_1^L] &= \mathbb{E}[\Pi_1^A] + \int_{\mathcal{A}_2}\int_{\mathcal{A}_1}(p_1^A \Delta q_1^{A,L} + \Delta p_1^{A,L} q_1^L) dF(A_1) dF(A_2) \\ &= \mathbb{E}[\Pi_1^A] + \frac{k_1\big(\overbrace{b_2(b_1^2 + b_1 r b_2 + 2a_2 k_1 r b_2)}^{\text{Gain from soften competition}} - \overbrace{4a_1a_2(a_2k_1 + r b_2)}^{\text{Loss without competitor information}}\big)}{(4a_1a_2 - b_2(b_1 + r b_2))^2} \\ &\left(\int_{\mathcal{A}_2}(A_2 - \eta_{A_2}) A_2 dF(A_2)\right).\end{aligned}$$

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 3, 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.

Figure 3: Effect of Variable k_1 on $\Pi_1^L - \Pi_1^A$



Note: Figure 3 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 1b has $A_1 = 10$, and $E[A_2] = 10$, while Figure 3 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.

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.

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:

$$\begin{aligned}CW_2^L &= CW_2^A + \frac{1}{a_2}\int_{\mathcal{A}_2}\int_{\mathcal{A}_1}(q_2^A + q_2^L)\Delta q_2^{A,L} dF(A_1) dF(A_2) \\ &= CW_2^A + \frac{a_2b_2k_1(4a_1 - b_2k_1)}{2(4a_1a_2 - (b_1 + rb_2)b_2)^2}\int_{\mathcal{A}_2}(A_2 - \eta_{A_2})A_2 dF(A_2) \geq CW_2^A.\end{aligned}$$

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 2b, 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.

B 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 decreases $\zeta^{\downarrow}_{s,\tilde{t}}$ as

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

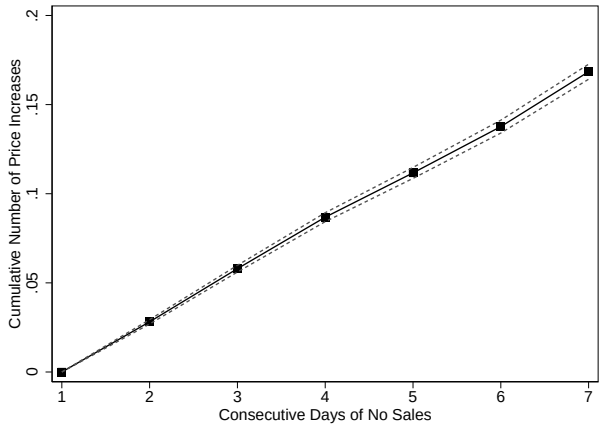
B.1 No-Sales Events

In Figure 4, 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 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.

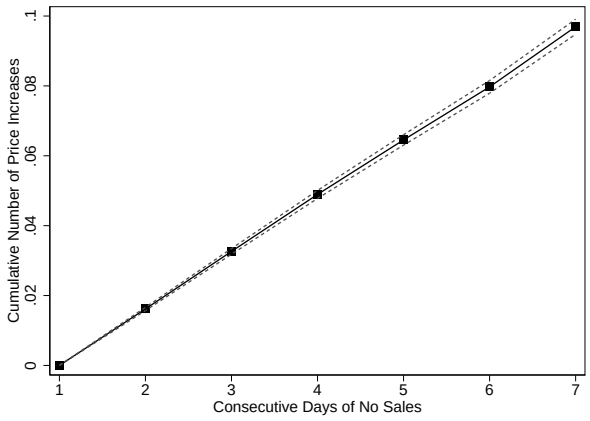
B.2 Sales Events

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 5, 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

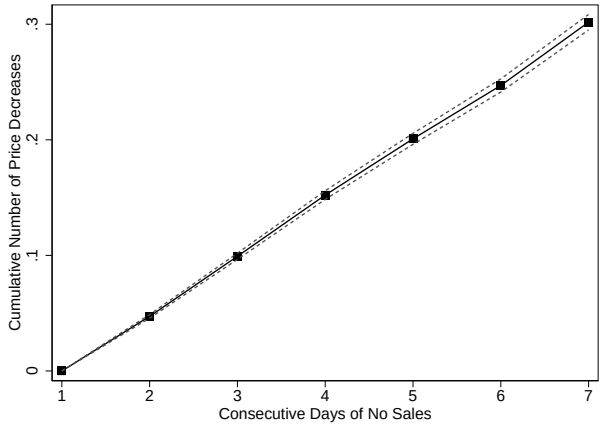
Figure 4: Cumulative Number of Price Adjustments during Consecutive Days of No Sales



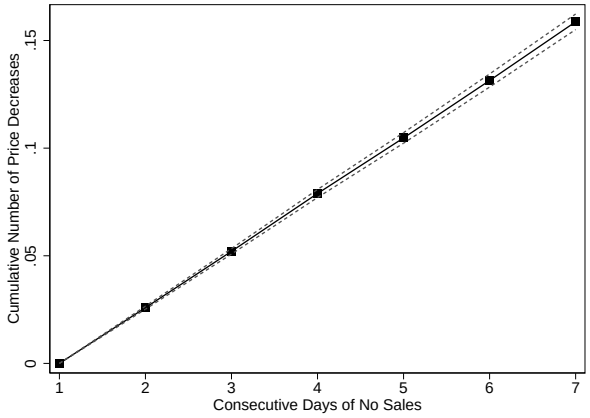
(a) Price Increase: Amazon



(b) Price Increase: Third-Party Sellers



(c) Price Decrease: Amazon



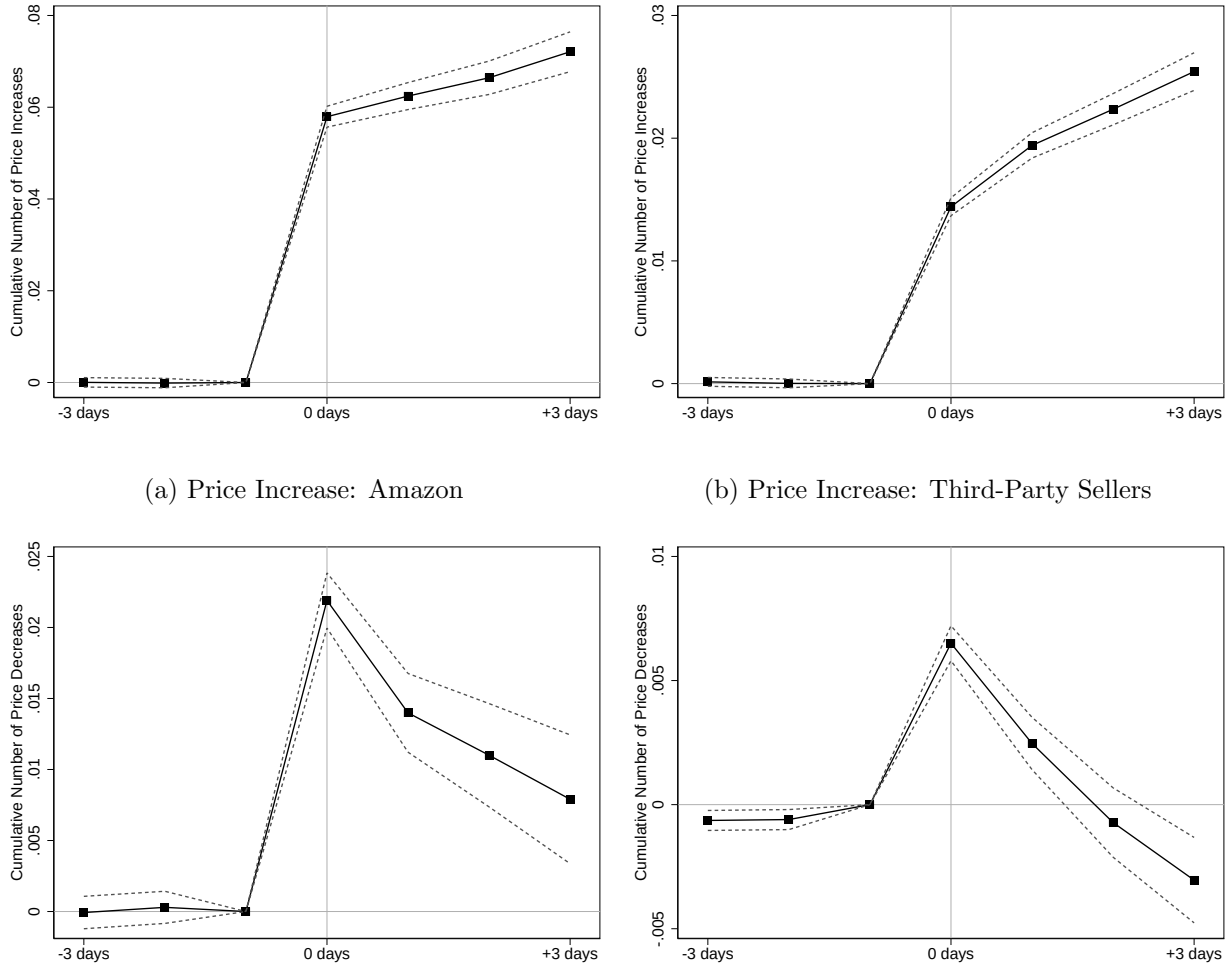
(d) Price Decrease: Third-Party Sellers

Note: Figure 4 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 4a and Figure 4b present the cumulative number of price increases and Figure 4c and Figure 4d present the cumulative number of price decreases. The robust standard errors are clustered at the product level.

event.[49]

[49]The outcomes in Figure 6c and Figure 6d 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 5: 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 5 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 5a and Figure 5b present the cumulative number of price increases and Figure 5c and Figure 5d 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.

B.3 Sellers Price On Their Own 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. Specifically, we define the

number of days with sales in the past as the following:

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

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 decreases $\zeta^{\downarrow}_{m,j,t}$, as defined in Equation a.1, in Figure 6. 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.

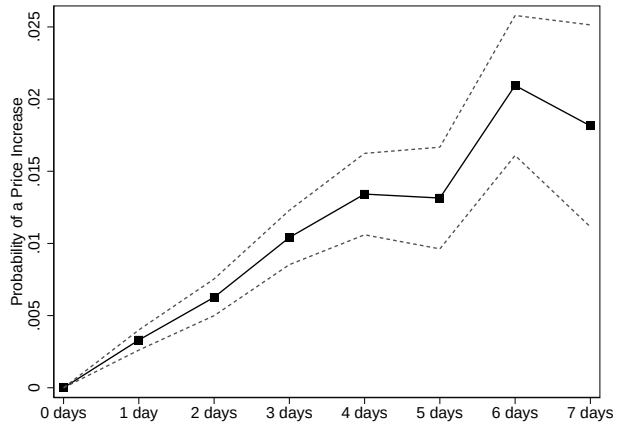
Furthermore, we present the estimates of γ_τ in Figure 7, using the price $p_{m,j,t}$ as the dependent variable in Equation a.2. The average price increase is larger when a seller has had more days with sales in the past, consistent with our findings for the probability of price adjustments.

B.4 Competitor's Sales Events

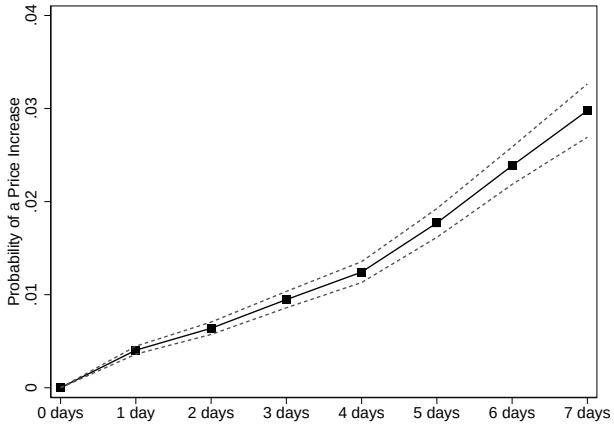
In Figure 8, we plot the estimates of γ_τ using the cumulative number of price adjustments as the dependent variable for Amazon and third-party sellers, respectively. According to our findings in Section 4.2, after a competitor's sales event, the competitor increases its price. In equilibrium, Amazon 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-

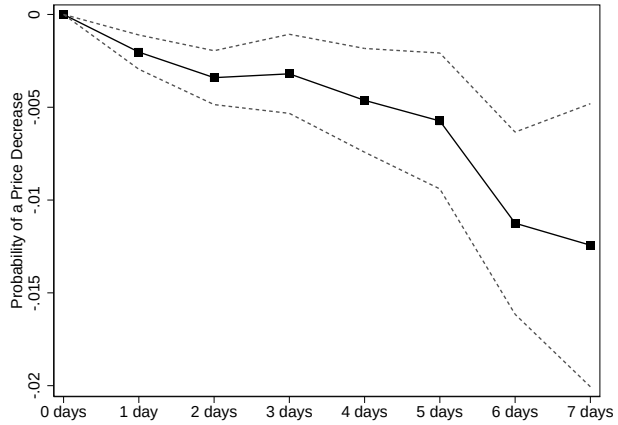
Figure 6: Price Adjustments Over Number of Days With Sales During The Past Week



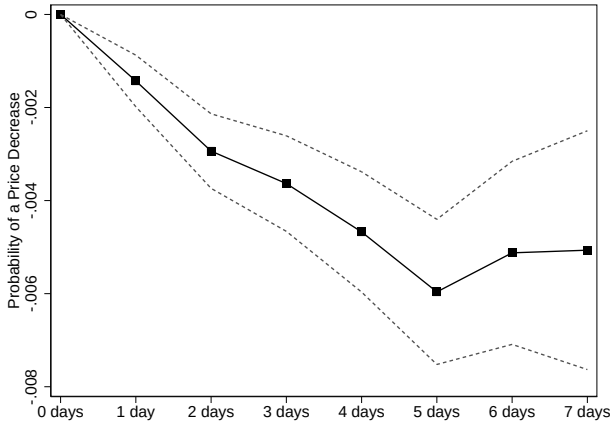
(a) Price Increase: Amazon



(b) Price Increase: Third-Party Sellers



(c) Price Decrease: Amazon

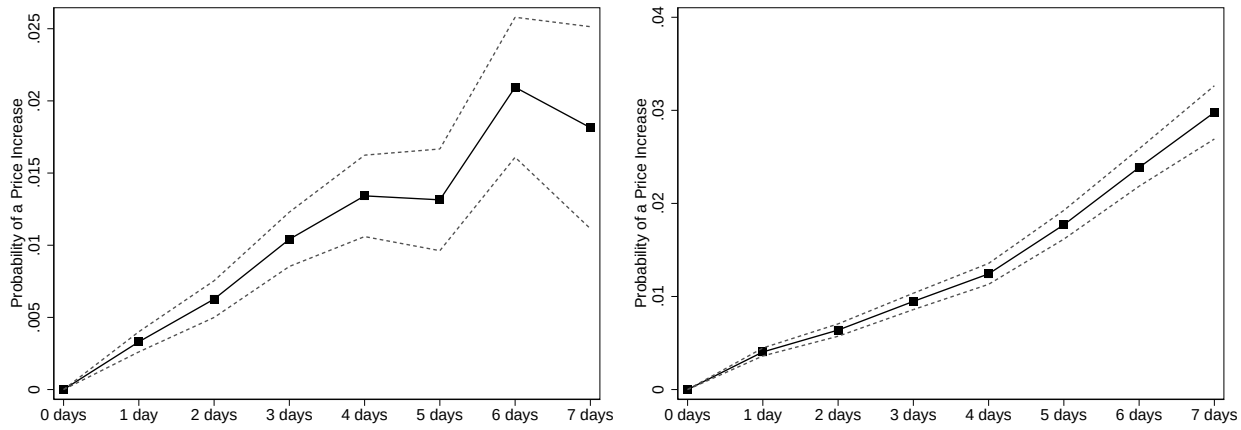


(d) Price Decrease: Third-Party Sellers

Note: Figure 6 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 6a and Figure 6b present the probability of a price increase, and Figure 6c and Figure 6d present the probability of a price decrease. The robust standard errors are clustered at the product level.

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 7: Price Levels Over Number of Days With Sales During The Past Week

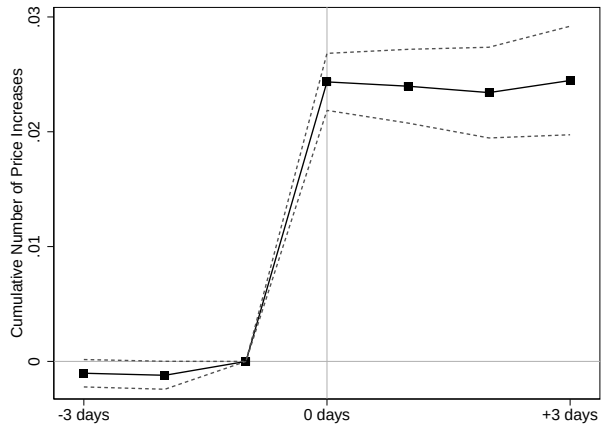


(a) Amazon

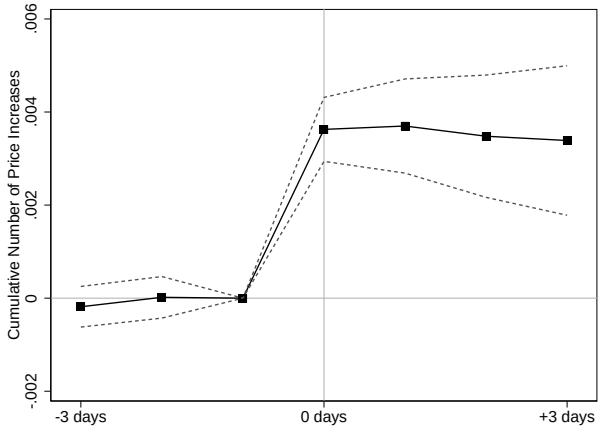
(b) Third-Party Sellers

Note: Figure 7 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.

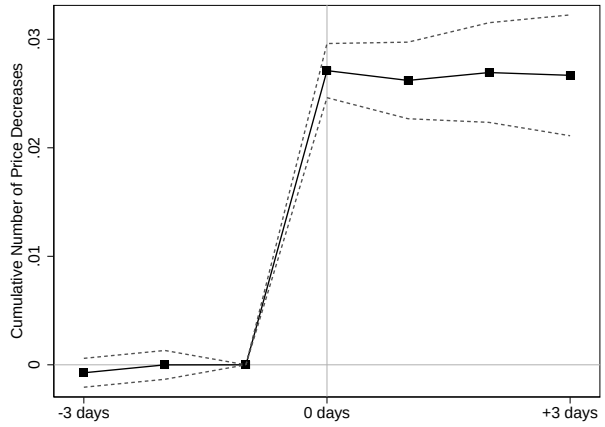
Figure 8: Price Adjustments After a Competitor's Sales Event



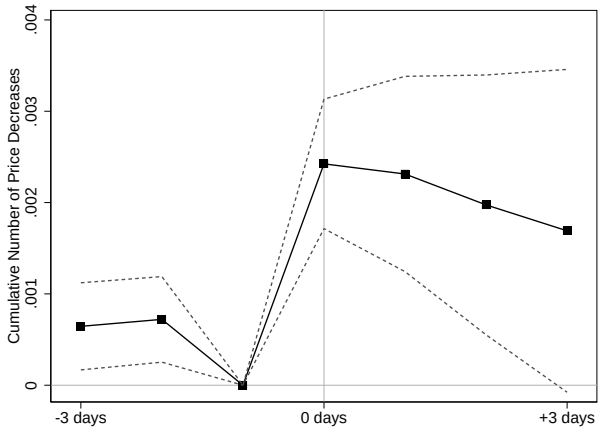
(a) Price Increase: Amazon



(b) Price Increase: Third-Party Sellers



(c) Price Decrease: Amazon



(d) Price Decrease: Third-Party Sellers

Note: Figure 8 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 8a and Figure 8b present the cumulative number of price increases and Figure 8c and Figure 8d 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.