Essays on Pricing Dynamics:
Evidence from the Brewing Industry and from Amazon Marketplace

Yanchi Yu
Daqing, China

M.A. Economics, Duke University, 2013
B.A. Economics, Zhejiang University, 2011

A Dissertation presented to the Graduate Faculty
of the University of Virginia in Candidacy for the Degree of Doctor of Philosphy

Department of Economics

University of Virginia
May, 2018

Abstract

This dissertation studies pricing dynamics using evidence from the US brewing industry and from the Amazon marketplace. In the first chapter, I analyze the relationship between market structure and inter-temporal price discounts in the U.S. brewing industry. Most studies assume that consumers face constant product prices within a month or a quarter. However, consumers can respond to price discounts and strategically adjust their shopping behavior. Firms exploit consumers' responses to temporary price discounts to inter-temporally price discriminate across consumers. A change of market structure may affect firms' price-discount strategies. I use the case of 2008 Miller/Coors joint venture to investigate how the change in market structure affects the dynamics of price-discount strategies of firms and quantify its welfare effects. I begin by documenting an empirical pattern that competing firms provide simultaneous promotions at stores in the pre-merger periods, while the merged firm alternates promotions after the merger. I then use autoregressive regressions to verify this empirical pattern statistically. To quantify the welfare effects of the change in price-discount strategies, I develop a structural model to characterize heterogeneous demand functions of consumers who stockpile (storers) and consumers who lack storage capacities (non-storers). I infer that a substantial number of consumers stockpile at the promotional prices. The percentages of sales to storers differ by brands and range from 13 percent to 26 percent. Storers are more price-sensitive and more likely to switch between brands. On the supply side, I model firms price-discount strategies using a two-stage game: in the first stage, firms consider whether to use an inter-temporal price discrimination strategy; in the second stage, firms simultaneously determine the product prices. If competitors used constant price strategies, firms can

increase their profits by at least 8 percent when switching to an inter-temporal price discrimination strategy. Likewise, if competitors used inter-temporal price discrimination strategies, firms can increase their profits by at least 6 percent when switching to inter-temporal price discrimination. In equilibrium, firms, therefore, choose inter-temporal price discrimination. After the market-structure change, the merged firm (with two close-substitute products) can and does increase its profit by 9 percent by staggering products on sale. I simulate the post-merger product prices and determine the difference of welfare effects with/without considering the promotion-strategy adjustment. Static models of competition ignore this effect which leads to a substantial under-estimation of the welfare impact of market mergers.

The second chapter, which is joint work with Denis Nekipelov, studies the pricing dynamics of algorithmic agents in the Amazon marketplace. Availability of algorithmic tools allowed many small retailers manage multi-product inventory and dynamically price their products in online marketplaces. At the same time, when the price is determined by an automated algorithm rather than retailer's own decision, the link between the product's marginal cost and the price traditionally studied in Industrial Organization is lost. The information regarding the marginal cost is implicitly communicated through the automated price updates generated by the automated algorithm. In this chapter, we use the ideas from the online learning literature in Computer Science to restore the link between the observed price changes and the marginal costs of retailers. The methodology developed in Nekipelov et al. (2015) uses the notion of regret to evaluate the automated algorithm. Regret measures the relative performance of the algorithmic dynamic strategy relative to the benchmark which corresponds to the best-fixed price in hindsight. This idea allows us to recover the identified set that contains the retailer's marginal cost as well as the expected

regret of her automated price strategy. We apply this methodology to study dynamic pricing on Amazon's marketplace. We find that expected regret for most retailers is close to zero. As a result, despite the simplicity of their algorithmic tools, they have good dynamic performance. At the same time, the estimated markups of retailers imply demand elasticities that are compatible with traditional retail markets. This may indicate that online marketplaces where small retailers use algorithmic tools may have good performance while achieving similar outcomes for consumers as the traditional retail.

JEL Classifications: D12, D22, L41, M31, L81, L66, L81

Keywords: Price discrimination; Market structure; Horizontal merger; Grocery retail; Beer; Algorithmic pricing; Online marketplaces; Price competition

Acknowledgements

I am deeply indebted to my advisors, Simon Anderson, Denis Nekipelov, Andrew Kloosterman, and Natasha Zhang Foutz for being a constant source of motivation, encouragement, support, and guidance through my Ph.D. journey. They have always been there to help me throughout my Ph.D. Their advice is not limited to research, but also towards holistic development of a student. Without their encouragement and support, I would not have gotten nearly this far.

I have benefited greatly from insightful comments and conversations with many faculty members and graduate students at the University of Virginia. I would like to specifically thank Gaurab Aryal, Federico Ciliberto, Kerem Cosar, Kenneth Elzinga, Maxim Engers, Charles Holt, Leora Friedberg, David Mills, Eric Young, and many participants at the IO research meetings. I also would like to thank my friends, who made the last five years at UVA enjoyable.

I acknowledge the Bankard Fund for Political Economy and the Radulovacki Summer Research Fund for the financial support and Kilts Marketing Data Center for providing access to AC Nielsen Retail Scanner Data.

To my parents, Xiaoyan and Changjiang, and my husband He, for unconditional love, unwavering support, and faith in me — thank you.

Contents

Chapter 1

Market Structure and Product Pricing Dynamics: Evidence from the U.S. Brewing Industry

1.1 Introduction

Economists have devoted substantial effort to understanding the consequences of market structure changes in determining the extent of market competition. The market structure provides information about a firm's competition environment, and it has the potential to influence a firm's strategic decisions, including product pricing, product quality, the variety of products, and other aspects. Of particular interest are the price effects of horizontal mergers in the concentrated markets, which influences consumer surplus and helps to shape antitrust, regulatory, and trade policies. The existing literature assumes that consumers face constant product prices within a certain period; however, Hendel and Nevo (2006a) and Hendel and Nevo (2006b) have shown that consumers can respond to price discounts and strategically adjust their shopping behavior. Firms exploit consumers' responses to price discounts and use temporary price discounts to inter-temporally price discriminate consumers. This work investigates how firms change their price-discount strategies as the markets become more concentrated, and quantify the welfare effects regarding the change of price-discount strategies.

To study the relationship between market structure and firms' promotion decisions, I use the 2008 Miller/Coors joint venture as an exogenous market-structure change to investigate how firms change their price-discount strategies. In 2008, the second largest brewing company SABMiller and the third largest brewing company Molson Coors in the US market formed a joint venture MillerCoors. The joint venture was approved on June 5th, 2008 by the DOJ, and was completed and operated as a combined entity on June 30th, 2008. After the market structure change, the US beer market became highly concentrated by two giant firms, Anheuser-Busch InBev (AB InBev) of 48.2 percent market share and MillerCoors (MC) of 29.5 percent market share. Miller and Weinberg (2016) documents a static retail price increase of major flagship beer brands after the market structure change. However, firms' price-discount strategies are more complicated in a dynamic setting.

One key factor that determines firms' responses in the price-discount strategies is consumers' responses to temporary sales. Temporary sales may only attract consumers who frequently switch brands and exploit the low sale prices (brand switching effect); temporary sales may expand brand market shares by attracting new consumers and encouraging consumption by existing customers (brand expansion effect); it may also induce consumers to stockpile for future consumption (stockpiling effect). Heerde et al. (2003) and Steenburgh (2007) show that stockpiling effect dominates brand expansion and brand switching effects in temporary sales periods. Firms exploit consumers' stockpiling behavior and the heterogeneity in consumer storage abilities to implement an inter-temporal price discrimination strategy (Hendel and Nevo (2013)). Using this strategy, firms set promotional prices in the promotion periods and regular prices in the non-promotion periods to induce consumers with good storage capacities (storers) only to shop in the promotion periods. In the promotion periods, firms carrying close-substitute products can benefit from setting only one product on sale to serve the storers. With this in mind, the primary aim of present research is to explore how multi-product firms adjust price-discount strategies as the markets become more concentrated.

In this paper, I focus on AB InBev and MillerCoors' price-discount strategies (pricing dynamics) from pre- to post-merger periods. I use auto-regressive regressions to document whether the pricing dynamics of AB InBev and MillerCoors have changed from pre- to post-merger periods. I find that firms set simultaneous promotions in the pre-merger periods while setting alternate promotions in the post-merger periods. The change of pricing dynamics results from the competition environment change. In the pre-merger periods, firms compete for consumers and choose to match competitors' promotional prices. In the post-merger periods, firms with multiple products switch to only set one product at the promotional price and capture the profit margin of other products. From the pre- to post-merger periods, consumers who are price sensitive and easy to switch still enjoy the promotion benefits, while consumers with a strong brand preference suffer from paying regular prices in the post-merger periods.

To measure the welfare effects, I develop a structural model to characterize heterogeneous demand functions of consumers with stockpiling capacities (storers) and those lacking storage capacities (non-storers) and estimate the model using Nielsen

Scanner data set. I find that data is consistent with a substantial number of consumers who stockpile at the promotional prices. The percentages of sales by storers differ by brands and range from 13.1 percent to 25.8 percent. Storers are more price-sensitive and more likely to switch between brands. These estimation results support the motivation of changing promotion decisions as the market becomes more concentrated.

Based on the demand side estimates, I model firms' price-discount strategies using a two-stage game: in the first stage, firms consider whether to use an inter-temporal price discrimination strategy; in the second stage, firms simultaneously determine the product prices. If competitors used constant price strategies, firms can increase their profits by at least 8 percent when switching to an inter-temporal price discrimination strategy. Likewise, if competitors used inter-temporal price discrimination strategies, firms can increase their profits by at least 6 percent when switching to inter-temporal price discrimination. In equilibrium, firms, therefore, choose inter-temporal price discrimination.

In the counterfactual analysis, I simulate the predicted prices after the market structure change considering this change of promotion decisions and compare with the case ignoring the change of price-discount strategies. Without considering the adjustment of promotion decisions, consumers are willing to pay around $30 per store to avoid the joint venture. Considering the reduced number of simultaneous promotions, consumers are willing to pay additional $72 per store to avoid this market structure change. Ignoring the adjustment of promotion decisions can lead to substantial bias when calculating consumer surplus.

This paper contributes to three strands of literature. First, it extends the literature on horizontal mergers. A group of papers analyzes the unilateral price effects of mergers in a particular industry. Examples include Dafny et al. (2012) in the health insurance industry; Allen et al. (2014) and Sapienza (2002) in banks; and Borenstein (1990) and Kim and Singal (1993) in airlines. These papers emphasize the unilateral price effects induced by a market structure change. The intuition is that firms internalize the business stealing effect and increase the prices as the market becomes more concentrated. Another group of papers incorporate firms' endogenous product choice and allow firms to endogenize both price and product characteristics after the market structure change. Examples include Sweeting (2010) in the broadcasting industry and Fan (2013) in newspaper. This work adds to the literature by considering the coordinated effects in promotions after horizontal mergers. There are few empirical works on the coordinated effects of horizontal mergers. Miller and Weinberg (2016) also structurally estimate a conduct parameter which rejects the Nash-Bertrand Equilibrium assumption after the horizontal merger, which implies collusive behavior after the merger. Compared with their work, this paper focuses on the change of price-discount strategies (pricing dynamics) after the horizontal merger.

Second, this paper also contributes to the literature on inter-temporal price discrimination and heterogeneous consumer responses. Hong et al. (2002) models heterogeneous consumer responses through shopping behavior assumptions, where all buyers have the same valuation for the goods but differ in price searching and storage abilities. There are two groups of consumers, "shoppers" and "captives." "Shoppers" always seek to buy at the lowest price and stockpile, while "captives" are loyal to firms. Each oligopolist owns a monopoly market of "captives" and competes for "shoppers." The paper incorporates the heterogeneous features of consumers and estimates the demand functions of "captives" and "shoppers". I find that "captives" are less price sensitive and less likely to switch between brands, while "shoppers" are more price sensitive and more likely to switch between brands. In addition, this work extends Hendel and Nevo (2013)'s work by considering a firm carrying two horizontally differentiated products. Based on the demand estimates, a firm carrying two close-substitute products benefits from charging the promotional price only for one product and the regular price for another product to gain an extra profit margin.

Third, this paper contributes to the literature on the brewing industry. About consolidation in the brewing industry, Tremblay and Tremblay (2005) document a consolidation trend in the post-World War II period. Elzinga (2011) describes the years 1950-1980 as the period of consolidation and the years from 1980 forward as the period of fragmentation. In recent years, some works are related to the market structure change in 2008, when SAB Miller and Molson Coors formed a joint venture to operate all beer business in the US market. Ashenfelter et al. (2015) estimate the effects of increased concentration and efficiencies on pricing and find that the predicted price increase was offset by a nearly equal cost efficiency effect. Chandra and Weinberg (2018) study the relationship between local market concentration and advertising behavior and find a significant positive effect of local market concentration on advertising expenditures. Sweeting and Tao (2016) show that if firms have uncertainty about each others costs, and they play a dynamic signaling game, the consolidation can lead to prices above the Nash-Bertrand Equilibrium. Compared with their works, this paper focuses on the promotion decisions in the US brewing industry after the horizontal merger.

The rest of the paper proceeds as follows. In section 2, I review the literature related to my research. In section 3, I provides the industry background on the brewing industry and retail promotions. Section 4 describes the data set I use and show the preliminary results. Section 5 provides the reduced-form regression analysis.

Section 6 and Section 7 present the structural model of demand and supply. Section 8 studies the welfare effects from the counterfactual analysis. Section 9 concludes.

1.2 Literature Review

This research is closely related to four sets of literature: effects of sales promotion; time-series modeling in price promotion research; intertemporal price dispersion; effects of horizontal mergers; and literature on the brewing industry. In this section, I discuss the existing research in these areas and how it relates to my research.

1.2.1 Effects of Sales Promotion

There is an extensive literature on sales promotion and the role it plays in changing consumer demand. Usually, a distinction is made between primary demand effects (timing acceleration and quantity increases) and secondary demand effects (brand switching), where primary demand effects account for borrowing consumer demand from other periods and increasing demand for the current period, and secondary demand effects represent brand-switching consumer demand. Decomposing promotion sales bump and measuring these effects are important in understanding firms' promotion strategies and provide policy implications.

Much of the work studying sales promotion effects find its basis in a seminal work by Gupta (1988) in which he distinguishes and measures three components of consumer responses: timing acceleration, purchase quantity, and brand choice. In this paper, he finds that approximately 74% of sales promotion elasticity results from the secondary demand effects (brand switching) and the remainder is attributed to the primary demand effects (timing acceleration and quantity increases). Following

Gupta (1988), Chiang (1991), Chintagunta (1993), and Bucklin et al. (1998), Bell et al. (1999) adapt, extend and generalize Gupta (1988)'s approach to multiple categories and brands. They all find that the percentage of secondary demand effects (brand switching) varies across categories but is dominant. Based on this approach, Pauwels et al. (2002) measure the long-term effects of sales promotion. However, a methodological issue is how to translate elasticity-derived decomposition into sales unit decomposition, what percentage of sales bump comes from brand switching, or stockpiling?

Heerde et al. (2003) clarify this issue by transforming the elasticity decomposition into sales volume decomposition. They find that, instead of 75% of the sales volume, only 33% of the sales volume is attributed to losses by other brands (brand switching), and the remainder (67%) is attributed to timing acceleration and quantity increases. However, these decomposition approaches are based on household data, decomposition of sales bump at the store level is absent in the literature. van Heerde et al. (2004) use store-level scanner data sets and find that brand switching, time acceleration, and quantity increases each account for a third of sales volume bump. Steenburgh (2007) reconciles the elasticity-based approach to the unit-based approach and provides an improved understanding of effects of sales promotion. Chan et al. (2008) develop a methodology to decompose the effects of sales promotion into brand switching, stockpiling, and change in consumption by allowing consumer heterogeneity. Their conclusion is compatible with the previous literature in which a large share of sale increases is attributable to stockpiling. In addition, they find that consumers with heterogeneous brand preference and usage respond differently to sales promotion. Most recently, Yoon and Tran (2011) revisit the relationship between consumer loyalty and price sensitivity by investigating the role of consumers' deal-proneness.

They find that consumers within the same loyalty segment exhibit different levels of price sensitivity.

The substantial purchase acceleration and stockpiling effects motivate the study of consumer stockpiling behaviors. Most notably, Hendel and Nevo (2006a) and Hendel and Nevo (2006b) estimate and compare price elasticities with/without consideration of purchase acceleration and stockpiling, and find that ignoring consumers' responses to promotion overestimate own price elasticities. The consumer demand elasticities are important inputs to study firms' optimal pricing strategy facing strategic consumers and have important policy implications.

1.2.2 Time-series Modeling in Sales Promotion

There is an extensive literature on time-series modeling, especially on sales promotion analysis. Dating back to Helmer and Johansson (1977), a number of research papers use time-series modeling to analysis sales and advertising[1]

Starting from Dekimpe and Hanssens (1995), researchers in marketing use the vector-autoregressive models with exogenous variables to study dynamic pricing tactics, the short-run and long-run effects induced by price promotions (Nijs et al. (2001), Srinivasan et al. (2004)), and the drivers of price promotions (Nijs et al. (2007)). Time-series models are well suited to model pricing dynamics for several reasons: first, retail prices are either mean- or trend- stationary, which satisfies the stationary condition for time-series modeling; second, the time-series model allows for endogeneity of retail prices, which can be explained by past prices and other exogenous variables; third, time-series models are designed to measure both direct (immediate/

[1] For instance, Leone (1983), Doyle and Saunders (1990), and Franses (1991). More details can be found in Dekimpe and Hanssens (2000).

or lagged) promotion responses and persistent responses by impulse response function (IRF). Recently, marketing researchers use the time-series models to measure the direct and indirect effects of advertising (Joshi and Hanssens (January 2010)) and quantify the effect of product recalls on brand value (Borah and Tellis (2015)).

Compared to structural models, a reduced-form time-series model is appropriate for "innovation accounting"(Enders (2004)), which provides descriptive insights on the patterns observed in the data (Nijs et al. (2007)). In this paper, the time-series models are appropriate for providing descriptive insights on whether firms change their pricing dynamics after a market structure change.

1.2.3 Intertemporal Price Dispersion

A large body of literature in economics and marketing has analyzed the intertemporal price dispersion. Intertemporal price dispersion occurs when an identical good is sold at different prices across different times in a given geographic market. There are several classes of theories explaining the existence of intertemporal price dispersion: intertemporal price discrimination, price competition of multiple firms facing stockpiling consumers, national and local brand competition, loss-leader, and prospect theory. In this section, I discuss relevant papers of each explanation theory.

The intuition of intertemporal price discrimination is that monopolists adjust their prices over time in order to discriminate between different types of consumers (Conlisk et al. (1984); Sobel (1984)). There are only a few papers that study intertemporal price discrimination for non-durable goods. Blattberg et al. (1981) introduce a model with two types of consumers differing in holding costs. The seller provides promotions to induce low holding cost consumers forward buy, while consumers with high holding cost buy every period. These periodic promotions help the seller to transfer inventory holding costs to consumers. In another paper, P.Jeuland and Narasimhan (1985) also consider two groups of consumers differing in holding costs. Additionally, they assume that high-cost consumers have a higher willingness to pay, and low-cost consumers have a lower willingness-to-pay. The sellers set periodic promotions to price discriminate over time.

Su (2010) extends P.Jeuland and Narasimhan (1985)'s work and incorporates multiple consumer groups differing in holding costs, fixed costs, consumption rates, and valuations. He has shown that sellers should use periodic promotions when frequent consumers have higher valuations, higher consumption rates, and lower fixed costs than occasional shoppers in a rational expectation equilibrium. Besbes and Lobel (2015) study a durable-goods firm's optimal pricing strategy facing a steady arrival of strategic consumers and establish the equivalence between the problem of pricing for a stream of heterogeneous strategic consumers and for heterogeneous stockpiling consumers. They prove that the firm restricts attention to cyclic pricing with at most twice the maximum willingness-to-wait length. This group of literature demonstrates that price promotions effectively inter-temporally discriminate consumers heterogeneous in many dimensions, including valuations, holding costs, fixed costs, and consumption rates. Consumers facing dynamic prices can stockpile to maximize their household utilities.

A class of theory papers explains the intertemporal price dispersion from firms' competitive pricing strategies. Salop and Stiglitz (1982) consider a market in which consumers and firms perform a Bertrand price competition, in equilibrium, firms mix regular prices and sale prices. Bell et al. (2002) consider a setting that consumers have flexible consumption and may re-enter the market with some probabilities and characterize a similar mixed-strategy equilibrium. Hong et al. (2002) extend Varian (1980)'s model by allowing consumers to carry inventories. In their model, all buyers have the same valuation for the goods but differ in price searchings and storage abilities. These differences make a distinction between two groups of consumers, "shoppers" and "captives", where "shoppers" always seek to buy at the lowest price and stockpile, while "captives" are loyal to firms. Each oligopolist owns a monopoly market of "captives" and compete for "shoppers". This setting obtains a mixed-strategy equilibrium, and mean prices increases cyclically. Gangwar et al. (2014) extend Hong et al. (2002)'s model by allowing endogenous stockpiling price thresholds and obtain a mixed strategy equilibrium with a multi-modal price distribution.

Anton and Varma (2005) explain the intertemporal price dispersion by a quantity competition for a homogeneous good with heterogeneous consumer valuations. All consumers can store the goods; consumers stockpile when price increases over time. An oligopolistic firm has a strong incentive to shift the future demand which is shared with its competitors to the current period by inducing consumers to stockpile. If the discount factor is sufficiently large, it obtains a pure-strategy equilibrium with storage, and it exhibits intertemporal price cycles. The equilibrium is independent of the rationing rule; besides the residual-demand rule, any rationing rule achieves the same equilibrium. Under the monopoly setting, it obtains a static monopoly price equilibrium with no storage. Guo and Villas-Boas (2007) study the price competition in differentiated markets facing heterogeneous stockpiling consumers using a two-period Hotelling setup. Consumers differ in valuations and propensities to stockpile. Consumers' valuations are positively correlated with consumers' propensities to stockpile. They assume that high-valuation consumers have a higher propensity to stockpile, then high-valuation consumers are out of the market in the second period.

In the second period, a lower degree of differentiation leads to more intense price competition, reducing the firms' incentives to lower prices in the first period. It obtains a no-storage equilibrium.

A third theory from the marketing literature (Lal, 1990) explains the intertemporal price dispersion as national brands collusively compete with local brands. Lal assumes a group of consumers are loyal to national brands and have a higher willingness to pay, while others can switch from national brands to local brands if the price difference is large enough. Under certain parameter value ranges, national brands collusively take turns to low price selling to switchers and loyal consumers and price high selling only to loyal consumers. Another theory in the marketing literature (Lal and Matutes, 1994) explains the intertemporal price dispersion as a loss-leader strategy selling products below marginal cost to attract consumers to stores. Recently, another group of work explains the intertemporal price dispersion by consumers' rational recognition process. Consumers perform the bounded rationality under the spirit of Prospect theory (Kahneman and Tversky, 1979). Villas-Boas and Villas-Boas (2008) explain the sale as an approach for uninformed buyers to be willing to experience the good and learn about its fit, while informed buyers may forget or change their preferences. Heidhues and Koszegi (2014) explain the regular price as an approach that firms introduce risks to a deterministic environment to loss-aversion consumers. Chen et al. (2016b) show that a monopoly seller chooses the optimal cyclic pricing strategy considering that consumers' valuations are subject to transitory satiety.

On the empirical side, there are many papers considering forward-looking consumers in the storable goods market. Hendel and Nevo (2006b) structurally estimate a dynamic model of consumer choice and find static estimation results overestimate own-price elasticity but underestimate cross-price elasticity. Hendel and Nevo (2013) empirically quantify the impact of intertemporal price discrimination on profits and welfare and find that sales increase profits and have a modest effect on consumer welfare. Wang (2015) studies the effectiveness of soda taxes based on a dynamic demand model. The author finds that static analysis overestimates price elasticity, and soda taxes are unlikely to substantially influence soda consumption. Hinnosaar (2016) studies the effectiveness of restrictions on alcohol sales hours or days allowing forward-looking consumers buy in advance. Wang et al. (2017) study the effect of obesity on demand for soda accounting for consumers' storing behavior. Ching and Osborne (2017) estimate a weekly discount factor using scanner data on laundry detergents and find that consumers show a medium level of forward-looking. Baker et al. (2017) study how households respond to the changes in sale tax rates. They find that households respond strongly by stocking up on storable goods. Perrone (2017) proposes a shortcut to estimate the long-run price elasticities using French data on food purchases.

1.2.4 Horizontal Merger and Brewing Industry

This work also contributes to the literature on horizontal mergers. A group of papers analyzes the unilateral price effects of mergers in a specific industry. Examples include Dafny et al. (2012) in health insurance industry, Allen et al. (2014) and Sapienza (2002) in banks, and Borenstein (1990) and Kim and Singal (1993) in Airlines. These papers emphasize on the unilateral price effects induced by a market structure change. The intuition is that firms internalize the business stealing effect and increases the prices as the market becomes more concentrated. There are a few empirical works on the coordinated effects of horizontal mergers. Miller and Weinberg (2016) also structurally estimate a conduct parameter which rejects the Nash-Bertrand Equilibrium assumption after the horizontal merger.

Another group of papers incorporates firms' endogenous product choice and allow firms to endogenize both price and product characteristics after the market structure change. Examples include Sweeting (2010) in the broadcasting industry and Fan (2013) in newspaper. This work adds to the literature by considering the price promotion decisions after horizontal mergers. Compared with their work, this paper focuses on the change of pricing promotions (dynamics) after the horizontal merger.

About consolidation in the brewing industry, Tremblay and Tremblay (2005) document this consolidation trend in the post-World War II period. Elzinga (2011) describes the years 1950-1980 as the period of consolidation, and the years from 1980 forward as the period of fragmentation.

In recent years, a number of works are related to the market structure change in 2008, which SAB Miller and Molson Coors formed a joint venture to fully operate beer business in the US market. Ashenfelter et al. (2015) estimate the effects of increased concentration and efficiencies on pricing, and find that the predicted price increase was offset by a nearly equal cost efficiency effect. Chandra and Weinberg (2018) study the relationship between local market concentration and advertising behavior and find a significant positive effect of local market concentration on advertising expenditures. Sweeting and Tao (2016) show that if firms have uncertainty about each others costs and they play a dynamic signaling game, the consolidation can lead to prices above the Nash-Bertrand Equilibrium.

1.3 Industry Background

1.3.1 Three-tiered Vertical Structure

The beer industry differs from other consumer product industries by its unique three-tiered vertical structure, with tiers being brewer, distributor, and retailer. The legislation prohibits firms from operating in multiple tiers. The product flow in the brewing industry consists of several processes: upstream brewing firms produce and distribute beer products to middle-tier distributors; middle-tier distributors smoothly connect to local downstream retailers; local retailers reach final-end consumers.

Distributors are heavily influenced by brewers. From upstream brewers to middle-tier distributors, brewers assign exclusive distribution territories to distributors serving their products. Within each assigned distribution territory for a certain brewer, the distributor has exclusive rights to serve beer brands from this brewer. The distributors' primary responsibilities include stock rotation and placement of point of sale materials. Distributors' assigned distribution territories are fixed in the relatively long time scope. In a contract between one national brewer and distributors, the exclusive territory does not contain any fixed expiration date, and the distributor owners can pass the business to their heirs and legatees without the brewer's prior approval. Though distributors can choose to serve beer brands from multiple national brewers, in practice, a distributor usually serves one of the big brewers (AB InBev, Miller or Coors) only.

From middle-tier distributors to local downstream retailers, brewers have substantial impacts on the wholesale prices charged by distributors to retailers. In practice, promotional activities and wholesale prices are not negotiated locally, instead, are concluded between sale representatives of brewers and retailer chains. The sale representatives from brewers will reach to retailer chains and conclude wholesale prices. Distributors serving the chain retailers are induced to supply at the achieved wholesale price (Asker (2016)). Though many states prohibit retail price maintenance and the three-tiered system protects the independence of each tier, brewers still have substantial influences over wholesale prices and promotional activities (Asker (2016)).

1.3.2 Temporary Price Promotions

Each year, packaged goods manufacturers spend around $75 billion dollars on promotions (Nijs et al. (2009)). Usually, temporary promotions induce short-run volume sales gains. Previous literature decomposes these sales gains and finds that these sales gains mainly result from stockpiling and brand switching effects. For instance, Blattberg et al. (1995) find that the short-run volume gains come from stockpiling, brand switching, and brand expansion effects. The brand expansion effects on the brand market shares are weak in the long run (Srinivasan et al. (2000)), which is referred as the dust settling effect (Srinivasan et al. (2004)).

The relative importance of stockpiling and brand switching is controversial. Gupta (1988), Chiang (1991), Chintagunta (1993), Bucklin et al. (1998), and Bell et al. (1999) find that a large fraction (about 75%) of demand responses come from brand switching, while Heerde et al. (2003) and Steenburgh (2007) revisit this question considering the sale volumes and find stockpiling demand effects dominate. In the economics literature, Hendel and Nevo (2006a) show that consumers respond to the sales by anticipating demand and stockpiling. Temporary price promotions (sales) induce consumers to stockpile and switch between brands.

Considering the stockpiling and brand switching effects on manufacturers, the previous literature shows that manufacturers have a motivation to provide sales. Temporary price promotions substantially increase the manufacturer's revenue, while having a mixed effect on the retailer's revenue (Srinivasan et al. (2004)).

In addition, by decomposing the drivers of retail price promotions, the previous literature shows that competitive prices from retailer competition only account for a small percentage of variation in the retail prices (Nijs et al. (2007)), while the main driving forces of retail price variation are supplier prices and promotional activity (Volpe et al. (2017)). Both supplier prices and promotional activity are brewers' choices, which depend on the market structure and demographics in the local markets.

1.3.3 Market Structure

The change of market structure indicates the change of firms' competition environment. In different competition environment, firms strategically compete in prices, product variety, quality, advertising, and temporary price promotions (sales). The US brewing industry was concentrated and became more consolidated after the joint venture of MillerCoors. Before the joint venture, Anheuser-Busch InBev (AB InBev) was the largest brewer with 42% revenue-based market share of the domestic beer market, SABMiller (Miller) is No.2 with about 18% revenue-based market share and Molson Coors (Coors) is No.3 with about 11% revenue-based market share. On Oct 9th, 2007, SABMiller and Molson Coors announced a joint venture of MillerCoors which will hold all of their operations in the U.S. On June 5th, 2008, DOJ Antitrust Division approved the joint venture MillerCoors. On June 30th, 2008, the joint venture was completed and operated as a combined entity.

Table 1.4 shows the annual revenue-based market shares over 2006-2010 of five major brewers including AB InBev, SABMiller, Molson Coors, Groupo Modelo and Heineken. From 2006 to 2010, the total revenue-based market shares decline around 1%, which indicates the growth of craft beer does not dramatically squeeze the major brewers' market shares. After the market structure change, AB InBev and Miller-Coors account for around 72% market share, which indicates the possibility of adjusting promotion decisions after the market structure change.

1.4 Data and Descriptive Analysis

The primary dataset is the Nielsen retail scanner dataset obtained from the Kilts Center for Marketing. This dataset includes weekly prices, sale volumes of each universal product code (UPC), and store environment generated by point-of-sale systems from participating retailers across all US markets. I choose the sample period from 2006 to 2010 to construct a balanced pre- and post-merger store-product panel. I restrict the attention to the market for light beers, which has higher retail sales across the country. Three of top four best selling beers in 2007 are light beers including Bud Light, Miller Lite, and Coors Light[2]. Also, consumers view light beers as a more healthy choice, and full-calorie beers are weak substitutes of low-calorie light beers (Sweeting and Tao (2016)).

I focus on five flagship light brands including Bud Light, Miller Lite, Coors Light, Michelob Ultra Light and Corona Extra Light[3]. These five brands are sold in at

[2]Bud Light, Miller Lite, and Coors Light have 42 million bbls, 18.4 million bbls, and 17.35 million bbls respectively.(Beer Marketer's Insights, 2008)

[3]Brands including Bud Light and Michelob Ultra Light belongs to AB InBev, SABMiller brews and operates products under the Miller Lite brand, Molson Coors carries Coors Light, and Crown Imports manages Corona Extra Light brand.

least 90% of 7894 representative stores in the dataset. These five flagship light beer brands are representative of major domestic and import light beers, Table 1.1 shows Bud Light, Miller Lite, and Coors Light are in the domestic beer price range, while Michelob Ultra Light's price falls into the premium domestic light beer price range. Corona Light is about 40 percent more expensive than the top three light beer brands, which stands for the high-end import light beer.

Beer products are sold in different packages size and containers, including 6-packs, 12-packs, 24-packs, and 30-packs. In the paper, I focus on the 12-packs flagship beer brands because it is the most widely sale package[4]. More than 90% stores carry 12-packs of all five flagship brands, while less than 10% stores carry 24-packs or 30-packs of all top three light beer brands. Including 24-packs and 30-packs light beers directly may lead to selection bias on stores. Less than 25 percent of stores carry 6-packs of all top three light beer brands. I merge 12-packs can and bottle beer products because 12-packs cans and bottles have the same price in most stores.

I do not aggregate 24-packs and 30-packs to 12-packs to avoid the possibility of creating additional variation over time.[5] For instance, when I use volume-weighted average price of 12-packs and 30-packs beer products, assuming the price of 12-packs and 30-packs beer products stay the same across time, the average price of 20 units sold in 12-packs and 19 units sold in 30-packs is different from the average price of 19 units sold in 12-packs and 30-packs. Aggregating beer products in different packages creates additional noise in classifying the inventory state.

There are a number of temporary price reductions (sales) over time at each store.

[4]12pk and 30pk beer have a similar frequency of sales. However, 30pk flagship beer products are missing in the majority of stores. Focusing on 30pk will leads to a potential selection bias in stores.

[5]The variation over time is used to classify the inventory state, merging with products in other volume using volume-weighted average creates more sale periods, which could lead to bias in classifying the inventory state.

A typical pricing pattern observed is a regular price with some temporary price cuts. Figure 1.1 and 1.2 represent a typical pricing pattern at a single store[6]. Figure 1.1 shows that Bud Light, Miller Lite, and Coors Light have promotions in the same periods, while Figure 1.2 displays that Bud Light, Miller Lite and Coors Light switch to non-simultaneous promotions in the post-merger periods. In addition, it is interesting to find that the gap between sale and non-sale prices shrinks after the market structure change. In the pre-merger periods, the sale price is around $7.5, and the regular price is around $11.25. In the post-merger periods, the sale price increases to $10.5 and the regular price increases to $12.5. One explanation for this shrinking gap is that firms adjust the expenses of promotions after the market structure change. This adjustment of promotion expenses influences both promotional prices and promotion patterns.

Besides the price increases in both sale and non-sale periods, the promotion patterns, especially whether all three products are on sale at the same time, change after the formation of the joint venture. Table 1.2 shows the promotion decisions in stores from pre- to post-merger periods. According to both panels in the table, the conditional probabilities of 2 or 3 flagship brands on sale, three brands on sale, and any combinations of 2 flagship brands on sale decrease from pre to post-merger periods. The differences in probabilities from pre- to post-merger periods are positive and significant. Another notable result is that around three out of five sales in the pre-merger periods are sales with 2 or 3 brands on sale, and four out of five sales in the pre-merger periods at the frequent sale stores are sales with 2 or 3 brands on sale. It is worthwhile to note that the conditional probabilities of 2 and three flag-

[6]The store is located at Sonoma, California. The store sells around 40 cases of beer per week in year 2006.

ship brands on sale decrease around 6.5% and 14%, respectively. In addition, I find that the conditional probabilities of 3 flagship brands on sale and any combinations of 2 flagship brands on sale decrease from the pre- to post-merger periods. Table 1.2 indicates that the promotion strategies change from pre- to post-merger periods statistically.

1.5 Reduced Form Results

1.5.1 Time-series variation

To study the pricing dynamics, I use a time-series model with exogenous shocks to document the change of pricing dynamic patterns. The autoregressive models with exogenous shocks were used in the marketing literature (Dekimpe and Hanssens (1995), Nijs et al. (2001), Srinivasan et al. (2004), Enders (2004), Nijs et al. (2007)). Economists use the autoregressive models with exogenous variables to study the dynamic pricing tactics, the short-run and long-run effects induced by price promotions (Nijs et al. (2001), Srinivasan et al. (2004)), and the drivers of price promotions (Nijs et al. (2007)). Compared to structural models, a reduced-form time series model is appropriate for "innovation accounting"(Enders (2004)), which provides descriptive insights on the patterns observed in the data (Nijs et al. (2007)). In this paper, a time series model is appropriate for providing descriptive insights on whether firms change their simultaneous promotion decisions after a market structure change. In this section, I exploit the time-series variation to analyze the change of simultaneous promotion decisions.

The time series regression specifies the relationship between a product's current price, previous prices, and other exogeneous variables.

$$p_{jmt} = c + \boldsymbol{\alpha}\mathbf{p}_{-jmt} + \boldsymbol{\beta}\mathbf{p}_{-jmt}\mathbb{1}\{post\} + a_1 p_{jmt-1} + a_2 p_{jmt-2} + ... + a_p p_{jmt-p} + HD_t + \varepsilon_{jmt}$$

where p_{jmt} is the product j's volume-weighted average price at week t in market m, $\mathbf{p}_{-jmt}$ is a vector of volume-weighted flagship light beer product prices other than j at market m, HD_t is a dummy variable controlling for national holidays[7]. Each market m is a designed market area (DMA) region defined by Nielsen[8]. In this analysis, p_{jmt} is the Miller Lite price at market m week t, $\mathbf{p}_{-jmt}$ is a vector of volume-weighted average product prices other than Miller Lite. The Miller Lite price p_{jmt} is correlated with its own past price histories $a_1 p_{jmt-1} + a_2 p_{jmt-2} + ... + a_p p_{jmt-p}$ and competitors' current price $\mathbf{p}_{-jmt}$. I use dummy variable $\mathbb{1}\{post\}$ to indicate the post-merger periods. In the post-merger periods, $\mathbf{p}_{-jmt}\mathbb{1}\{post\}$ captures the change of contemporaneous effects from its competitors.

When deciding the optimal lag for the model, I reply on the widely-used AIC and BIC criteria. For the majority of DMA regressions, AIC criteria and SBIC criteria agree on the same optimal lag numbers. In the cases that AIC and SBIC differs, I reply on the SBIC criteria, which strongly penalizes free parameters. It means that I lean on a more conservative choice of lag numbers. The optimal numbers of lags are 1 or 2 in most of the DMA areas.

The coefficient vector of competitor's current price vector $\mathbf{p}_{-jmt}$, $\boldsymbol{\alpha}$ captures the effects of competitors' prices in the pre-merger periods. If $\boldsymbol{\alpha} > 0$, that is, if product

[7]National holidays include Superbowl week, Independence day, Thanksgiving, and Christmas holiday.

[8]A DMA region is a group of counties that form an exclusive geographical area in which the home market television stations are dominant.

j's competitors have promotion prices, product j will likely have a promotion price in the same period (simultaneous promotion). If $\boldsymbol{\alpha} < 0$, the product j sets a regular price while its competitors promote, the product j sets a promotional price while its competitors set a regular price (non-simultaneous promotion). The coefficient $\boldsymbol{\beta}$ before the interacted term $\mathbf{p}_{-jmt}\mathbb{1}\{post\}$ captures whether Coors Light and Bud Light change the promotion strategies after the market structure change. A statistically significant negative $\boldsymbol{\beta}$ indicates firms tend not to use simultaneous promotions in the post-merger periods.

To identify the parameters of interest, I exploit the time-series variation of prices and use competitors' current prices as exogenous shocks. In the pre-merger periods, Miller Lite, Coors Light, and Bud Light belong to different firms, and the price information from competitors is private information, which supports the assumption of exogenous price shocks from competitors. In the post-merger periods, if Miller Lite's price and Coors Light's price are positively correlated, I may overestimate the contemporaneous effect of Coors Light. If Miller Lite's price and Coors Light's price are negatively correlated, I may underestimate the contemporaneous effect of Coors Light. Another parameter of interest is the coefficient before Bud Light price interacted term. In both pre and post-merger periods, Bud Light's price information is private to its competitors, which can be treated as exogenous shocks. After controlling national holidays, if Bud Light's price and Miller Lite's price are positively correlated, I may overestimate the contemporaneous effect of Bud Light's price, which indicates a more negative estimate before the Bud Light interacted term.[9]

[9]Table 1.2 supports the positive correlation between Bud Light and Miller Lite and the positive correlation between Miller Lite and Coors Light.

1.5.2 Results

Table 1.5 shows the time series regression results for Miller Lite at the DMA-market level. Each column represents the regression result for a sample DMA. In column 1, Miller Lite and Coors Light were on promotion in the same sale weeks before the merger, while the simultaneous promotions are less likely in the post-merger periods. In column 2, Miller Lite, Coors Light, and Bud Light were on sale at the same time in the pre-merger periods, while Bud Light tends not to have simultaneous promotions with Miller Lite afterward.

Many other markets share this similar change of simultaneous promotion patterns observed in the Table 1.5. I run regressions for each market and label the markets with statistically significant reducing simultaneous promotions between Bud Light and Miller Lite as 1 and otherwise as 0. Similarly, I label the markets with statistically significant reducing simultaneous promotions between Coors Light and Miller Lite as 1 and otherwise as 0. I summarize the regression results for all markets in Table 1.6. Table 1.6 shows that 43% of stores have a statistically significant change in simultaneous promotions between Bud Light and Miller Lite. 14% of stores tend to have less simultaneous promotions between Coors Light and Miller Lite, and 8% of stores reduce simultaneous promotions with both Bud Light and Coors Light. [10] The main finding of the reduced simultaneous promotion in the post-merger periods still holds. The distribution of stores with less simultaneous promotions are across the country.[11]

The static price assumption ignores the firms' promotion decisions and consumers'

[10] I also run regressions using store-level market definition and classify the regression results in the same approach.

[11] Store-level regression suffers from missing data issue.

responses to promotions, which implies homogeneous effects on consumers with heterogeneous response abilities. For instance, consumers may differ in storage abilities. The increased average price may have little effects on consumers who can shop for deals and stockpile while having large effects on consumers who have to shop regularly. The static price assumption may underestimate or overestimate the effects on consumer surplus depending on the distribution of stockpiling and non-stockpiling consumers.

In addition, the static price assumption ignores firms' adjustment of promotion decisions after the change of corporate structure and market structure. After the market structure changes, firms may reduce promotion expenses and adjust the frequencies of simultaneous promotions of close-substitute products. As I document in this section, the joint venture reduces the frequency of simultaneous promotions by two close-substitute products Miller Lite and Coors Light. Considering consumers' heterogeneous responses to promotions and firms' adjustment in promotional decisions, it is not clear that static merger simulation can provide a complete answer to the welfare analysis. To study the welfare effects, I structurally model consumers' demand and firms' adjustment of promotion decisions and analyze the welfare effect after the market structure change.

1.6 Empirical Model

To study the welfare effects of a market structure change, I structurally model consumers' dynamic demand and firms' decisions to adjust promotional patterns. In this section, I model consumers' demand using a dynamic demand model with two types of consumers - storers and non-storers, and I model firms' price-discount strategies using a two-stage game: in the first stage, firms consider whether to use an intertemporal price discrimination strategy, and in the second stage, firms simultaneously determine the product prices. In a finite period repeated game, I solve a Subgame Perfect Nash Equilibrium.

1.6.1 Demand Model

In a static model, one common assumption is that consumers' current purchases are independent of previous purchases and consumptions. In each period, consumers only consider current product prices, decide how much to purchase and consume all their purchases. However, Hendel and Nevo (2006a) find that consumers stockpile at the prices for future consumption, which indicates the static consumption assumption may oversimplify consumer shopping behavior. In a dynamic setting, consumers take both current product prices and past inventories into consideration when deciding when and how much to purchase products.

To incorporate the dynamics of consumers' shopping behavior, I employ the demand model by Hendel and Nevo (2013). Specifically, I model consumers' two decisions: timing of purchase and the amount of purchase in each period. Compared with the dynamic demand model in Hendel and Nevo (2006a), this demand model does not require strong assumptions on the functional form of inventory cost and price expectations, and products' past prices are sufficient to indicate the inventory state (Hendel and Nevo (2013)). I follow Hendel and Nevo (2013) and make the following assumptions on consumer storage costs and information.

Assumption 1. *There are two types of consumers: storers and non-storers.*

Because consumers are heterogeneous in their stockpiling abilities, Assumption 1 uses two types of consumers to represent heterogeneous consumers. Some consumers have very high storage costs and will choose not to stockpile, while other consumers have low storage costs and choose to stockpile for future usages. For instance, some consumers have small refrigerators and choose not to stockpile, while other consumers have large refrigerators and are more likely to stockpile.

I define the consumer type $h \in \{S, NS\}$, in which I allow storers (type-S) and non-storers (type-NS) to have different preferences. Absent storage, in R finite periods with discount factor $\delta = 1$[12], type-h consumer solves the utility maximization problem as follows:

$$\max_{q_t} \sum_{t=1}^{R} E\left(u^h(q_t) + m_t\right) \quad s.t. \quad p_t q_t + m_t \leq y_t$$

where $u^h(q_t)$ is a quadratic utility function of beer consumption q_t, u' is the marginal utility from consuming q_t units of beer, and u'' is the diminishing utility per unit of beer consumption. Since I consider the beer consumption, which is a relatively small portion in the consumer household budget, I assume $m_t > 0$ holds for all t. Consumers always consume products other than beer. I allow different demand functions for different consumer types as follows:

$$q_t^h = Q^h\left(p_t\right), \quad \text{where} \quad h \in \{S, NS\}$$

Besides storage-absent demand functions, non-storers and storers differ in storing behavior, which is reflected in the differences of shopping frequencies and prices they actually faced. Consumers' shopping frequencies depend on whether consumers

[12]Since the time periods are weeks, I ignore the role of discounting, which is reasonable in a short time length.

consume the entire purchases per period. I assume that non-storers consume all the current purchases in each period. However, storers can choose to consume less than current-period purchases and stockpile for future consumption; storers can consume more by using both previous inventory and current purchases. To simplify inventory carryover and predict storers' purchase behavior, I make another assumption as follows:

Assumption 2. *Storage is free, but inventory lasts only for T periods.*

Assumption 2 implies that storers will only purchase products for T periods ahead. For instance, storers can have a fixed space for beer inventory. Storers may shop and stockpile at promotional prices until the storage is fully occupied. When the storage space is fully occupied, consumers can consume the beer inventory for at most T periods. Though beer products are not perishable like other groceries, Assumption 2 ensures that storers are not able to stockpile beer for the rest of their lives. In addition, this assumption simplifies the storage decisions of different products and assumes products purchased in the same period will expire after T periods. Also, it ensures that past prices are sufficient to indicate the inventory state.

Assumption 3. *Consumers know their future demand at least T periods ahead.*

Assumption 4. *Consumers have perfect foresight on future prices.*

Assumption 3 is reasonable for consumers since the time periods are weeks. Assumption 4 simplifies the demand by allowing storers to have complete information on the prices in the next T periods and only purchase during the lowest-priced periods. It is reasonable to make this assumption based on Mela et al. (1998), which found that the increased long-term exposure to promotions reduces the probabilities for consumers to purchase in the off-deal periods. When these consumers decide to buy, they "lie in wait" for especially good deals and tend to buy more of a good. In my model, for the group of consumers who pay attention to promotions and who have been exposed to promotions, it is reasonable to assume these consumers can wait for good deals and buy. In addition, I can think of these consumers as bargain shoppers who make more extensive research into the timing of sales. It is reasonable to assume a group of consumers who pay attention and have been exposed to deals, or a group of consumers who make extensive research in deals have perfect foresight on future prices and can wait for good deals and buy more in the deal periods.

Based on Assumption 3 and Assumption 4, storers only shop at the promotional periods while non-storers shop at all periods. Because storers only shop at the promotional periods, the price options for storers are not the actual prices in each period, instead, the lowest price within $T+1$ periods. Based on this perfect foresight assumption (Assumption 4), I define the effective price as the minimum price in the relevant $T+1$ periods. Define the effective price of product j at period t as follows:

$$p_{jt}^{ef} = min\{p_{jt-T}, ..., p_{jt}\}$$

Under Assumption 2, Assumption 3 and Assumption 4, both non-storers and storers face a static utility maximization problem. Specifically, non-storers do not stockpile, purchase and consume within each period. Storers only make purchases during promotional periods, which simplifies storers' problem to purchases during promotional periods. Instead of the actual prices p_t, storers face the effective prices p_t^{ef}, which is the lowest price vector among current period t and T storage periods ahead. Under the perfect foresight assumption (Assumption 4), storers only purchase the products at the effective prices. In R finite periods without discounting, storers solve the utility maximization problem as follows:

$$\max_{q_t} \sum_{t=1}^{R} E\left(u_t^S(q_t) + m_t\right) \quad s.t. \quad p_t^{ef} q_t + m_t \leq y_t$$

Because storers only purchase at the effective prices, storers' optimization problem is equivalent to R static optimization problems. In each static optimization problem, storers' optimal consumption in period t is $q_t^S = Q_t^S(p_t^{ef})$, and non-storers' optimal consumption is $q_t^{NS} = Q_t^{NS}(p_t)$.

Since storers only purchase at the lowest price within $T+1$ periods, storers' purchases are the sum of their current and future consumptions. Consider one-week storage case ($T = 1$) for example, whether storers purchase in the period t for the current consumption depends on p_{jt-1} and p_{jt}. If $p_{jt-1} > p_{jt}$, storers need to purchase at period t for consumption at the period t; if $p_{jt-1} < p_{jt}$, storers do not purchase for the current consumption and consume the inventory from period $t-1$. Similarly, whether storers purchase in the period t for the future consumption depends on p_{jt} and p_{jt+1}. If $p_{jt} < p_{jt+1}$, storers purchase and leave an inventory for the future consumption. If $p_{jt} > p_{jt+1}$, storers choose not to purchase for future. For ties like $p_{jt-1} = p_{jt}$ or $p_{jt} = p_{jt+1}$, I assume that consumers purchase immediately if the price is below a certain threshold p_j^m, or wait until the last opportunity to buy if the price is above a certain threshold[13]. Consider storage period as T, storers compare the current price p_{jt} to T preceding period prices and purchase for the current consumption only if p_{jt} is the lowest. Then storers compare the current price p_{jt} with $T-1$ preceding

[13]The certain threshold used is the store specific median price of product j. I follow Hendel and Nevo (2013) and make this assumption to break the price ties.

prices and $t+1$ price to decide the purchase for $t+1$ consumption. Storers repeat this comparison until $t+T$ period, and the purchase at the period t is the sum of current and future consumptions as follows:

$$x_{jt}^{S}\left(p_{t-T},...,p_{t+T}\right)=\sum_{r=0}^{T}Q_{jt+r}^{S}\left(p_{jt},p_{-jt+r}^{ef}\right)\mathbb{1}\{p_{jt}=p_{jt+r}^{ef}\}$$

Considering that storers only purchase at the lowest price within $T+1$ periods, sellers can predict the inventory state and adjust promotion decisions accordingly.

Predicted Demand (T=1)

I present the storers' predicted purchase for $T=1$ case to illustrate the intuition of purchase timing adjustment. Define the period t with $p_{jt}<p_{jt+1}$ when storers purchase for future consumption as sale periods S_t, and $p_{jt}>p_{jt+1}$ when storers don't purchase for future consumption as non-sale periods N_t. Define $p_{jt}=p_{jt+1}>p_j^m$ as non-sale periods N_t, and $p_{jt}=p_{jt+1}<p_j^m$ as sale periods S_t. Storers' purchases at period t are determined by the sale states in previous and current periods. For instance, if the last period is sale S_{t-1}, storers consume previous inventory and don't purchase for current consumption. If the current period is sale S_t, storers purchase for future consumption and leave an inventory. Based on this intuition, I define four events: a sale preceded by a sale ($S_{t-1}S_t$), a sale preceded by a non-sale ($N_{t-1}S_t$), a non-sale preceded by a sale ($S_{t-1}N_t$), and a non-sale preceded by a non-sale ($N_{t-1}N_t$).

Storers' purchases x_{jt}^S at period t are as follows:

$$x_{jt}^S\left(p_{t-1},p_t,p_{t+1}\right)=\begin{cases} Q_j^S\left(p_{jt},p_{-jt}^{ef}\right) \quad + \quad 0 & N_{t-1}N_t \\ Q_j^S\left(p_{jt},p_{-jt}^{ef}\right) \quad + \quad Q_j^S\left(p_{jt},p_{-jt+1}^{ef}\right) & N_{t-1}S_t \\ 0 \quad + \quad 0 & S_{t-1}N_t \\ 0 \quad + \quad Q_j^S\left(p_{jt},p_{-jt+1}^{ef}\right) & S_{t-1}S_t \end{cases}$$

where $Q_j^S\left(\cdot\right)$ is storers' storage-absent consumption demand. Because storers have perfect foresight and only purchase at the lowest price within $T+1$ periods, storers' optimal consumption depends on the effective price p_{-j}^{ef}. If the last period is non-sale N_{t-1}, storers do not have inventory carryover and have to purchase for the current consumption $Q_j^S\left(p_{jt},p_{-jt}^{ef}\right)$. In the events $N_{t-1}N_t$ and $N_{t-1}S_t$, the first component $Q_j^S\left(p_{jt},p_{-jt}^{ef}\right)$ stands for the need to purchase for the current period. In the events $S_{t-1}N_t$ and $S_{t-1}S_{t-1}$, the first component of purchase x_{jt}^S is 0, in which storers consume the inventory from the past period $t-1$. If the current period is sale S_t, storers purchase for the future consumption $Q_j^S\left(p_{jt},p_{-jt+1}^{ef}\right)$. In the events $N_{t-1}S_t$ and $S_{t-1}S_t$, the second component $Q_j^S\left(p_{jt},p_{-jt+1}^{ef}\right)$ is the purchase for future consumption. In the events $N_{t-1}N_t$ and $S_{t-1}N_t$, the second component of purchase is 0, where storers don't stockpile for the next period.

1.6.2 Supply Model

To study the welfare effects of a market structure change considering the adjustment of promotion decisions, I model and estimate storers' and non-storers' demand functions from the demand side. On the supply side, I model firms' decisions to adjust promotion patterns in an R-period finite repeated game. In each period, firms participate in a two-stage game: in the first stage, firms consider whether to use an inter-temporal price discrimination strategy; in the second stage, firms simultaneously determine the product prices.

Model

In a finite R-period repeated game, the set of players includes major brewing companies. Each brewing company j belongs to the set of players $\phi = \{B, M, C\}$[14]. In each period, a two-stage game models firms' pricing decisions. In stage 1, firms simultaneous choose whether or not to use a price discrimination strategy. The first-stage pure-strategy space s_j for each player j includes constant regular price N and regular price with temporary sales S. In stage 2, firms commit to their chosen price strategies (N or S), and simultaneously decide sequences of prices ($\{p_1, p_2\}$). For instance, if firm j chooses price strategy N in the first stage, firm j commits to the constant regular price strategy and set $p_{j1} = p_{j2}$. Firm j's payoff per period is measured by the total two-stage profits ($\Pi_{jt} = \pi_j(p_{jt_1}, p_{-jt_1}) + \pi_j(p_{jt_2}, p_{-jt_2})$) received by each strategy profile. Firm j's total payoff is measured by the sum of discounted total profit per period ($\sum_{t=1}^{R} \delta^t \Pi_{jt}$).

Timing

In each period, at stage 1, firm j decides to use temporary promotions strategy S or constant price strategy N. In stage 2, firm j commits to the promotion strategy chosen in stage 1 and sets a sequence of prices $\{p_{j1}, p_{j2}\}$. In stage 1, firms simultaneously make decisions on promotion strategies. In stage 2, firms observe all players' previous

[14] B stands for AB InBev, M stands for SABMiller, and C is Molson Coors

promotion strategies in stage 1, simultaneously decide sequences of prices, and commit to their price strategies.

Best responses

In a finite R-period repeated game, I solve the game by backward induction. In each period, firm j makes the price strategy (constant price N, or constant price with temporary sales S) in stage 1, and chooses a price sequence $\{p_{j1}, p_{j2}\}$ in stage 2.

In stage 1, firms' price strategy influences the timing of consumers' purchase decisions. If firm j chooses a constant price strategy N, consumers face a constant price and shop in both week t_1 and week t_2. If firm j chooses temporary promotions strategy, consumers who are heterogeneous in storage capacity alter their timing of purchases. Under the perfect foresight assumption (Assumption 4), consumers have complete knowledge of the price sequence and can adjust their purchase decisions accordingly. Non-storers shop every period and have a static demand $Q_j^{NS}(p_t)$ in the period t. Storers only shop at the lowest price and can stockpile for further consumption.

Absent other players' price strategies, firm j chooses to use an inter-temporal price strategy S if there are enough storers, and the demand of storers are relatively strong. Considering other players' price strategies, firm j chooses the best response strategy which maximizes its own profit. Firm j's profit function depends on the competitor firm m's price strategy s_m and the competitor firm n's price strategy s_n:

$$\Pi_j = \Pi_j\left(s_j, s_m, s_n\right)$$

Holding competitor firm m and competitor firm n strategies fixed, firm j chooses the regular price with temporary promotions strategy S if this strategy achieves a higher profit than the constant price strategy N:

$$\Pi_j\left(S, s_m, s_n\right) > \Pi_j\left(N, s_m, s_n\right)$$

In stage 2, firms simultaneously choose their price sequences $\{p_1, p_2\}$ that are compatible with their pre-selected price strategies in stage 1. I assume that firms have constant marginal cost mc, and firm j's profit function using constant regular price strategy N is

$$\Pi_j\left(N, s_m, s_n\right) = \max_{p_j} 2\left[Q_j^S\left(p_j, p_{-j}\right) + Q_j^{NS}\left(p_j, p_{-j}\right)\right]\left(p_j - mc_j\right)$$

where firms serve both non-storers and storers in both periods. The total profit is the sum of the profits serving both types of consumers. However, when firm j employs an inter-temporal price discrimination strategy $p_{j1} < p_{j2}$. Storers only purchase in the week with promotional price. Firm j serve both non-storers and storers in the promotional week, while only serve the non-storers in the regular price week. The total profit function is as follows:

$$\begin{aligned}\Pi_j\left(S, s_m, s_n\right) = \max_{p_{j1}, p_{j2}, p_{j1} < p_{j2}} & \left[2Q_j^S\left(p_{j1}, p_{-j1}\right) + Q_j^{NS}\left(p_{j1}, p_{-j1}\right)\right]\left(p_{j1} - mc_j\right) \\ & + Q_j^{NS}\left(p_{j2}, p_{-j2}\right)\left(p_{j2} - mc_j\right)\end{aligned}$$

where the first component stands for the profit of promotional week, and the second component is the profit of regular price week. The first component consists of storers' current consumption $Q_j^S(\cdot)$ and inventory $Q_j^S(\cdot)$, and non-storers' current consumption

$Q_j^{NS}(\cdot)$. The second component is only non-storers' purchase $Q_j^{NS}(\cdot)$. Firm j's best response function can be derived through the first order conditions. Solving for the vector of price under strategy s in period t, where $s \in \{N, S\}$ and $t \in \{1, 2\}$, provides

$$p_{s,t} = mc + \Delta(s,t)^{-1} Q(s,t)$$

where the purchase $Q(s,t)$ under strategy s in period t, $s \in \{N, S\}$ and $t \in \{1, 2\}$, is given by:

$$Q_j(s,t) = \begin{cases} 2Q_j^S + Q_j^{NS}, & \text{if } s = S, t = 1 \\ Q_j^{NS}, & \text{if } s = S, t = 2 \\ Q_j^S + Q_j^{NS}, & \text{if } s = N \end{cases}$$

where $\Delta(s,t)$ is given by:

$$\Delta_{jr}(s,t) = \begin{cases} -\frac{\partial Q_r(s,t)}{\partial p_j}, & \text{if } r \text{ and } j \text{ are produced by the same firm;} \\ 0, & \text{otherwise.} \end{cases}$$

The vector of markups depends only on the parameters of the demand system and the equilibrium price vector. Under the constant marginal cost assumption, I obtain the pricing equation as follows:

$$p_{s,t} - \Delta(s,t)^{-1} Q(s,t) = c + \omega$$

Estimates of the parameters can be obtained under the orthogonal conditions between ω and constant term.

Equilibrium

In a finite R-periods repeated game, I use backward induction and solve the Subgame Perfect Nash Equilibrium (SPNE). In the last period game, firms choose non-discriminating price strategy N or intertemporal price discrimination strategy S in stage 1, and commit to the pre-selected price strategies and choose the price sequences in stage 2. In stage 2, equilibrium is determined by the condition that all players choose the actions which are best responses to the anticipated play of their opponents.

In stage 2, define firm j's strategy s_j associated profit given competitors' constant price N as $\Pi_j^*(s_j, N, N)$. Assuming competitors using constant price strategy N and set constant prices, firm j chooses inter-temporal price discrimination strategy S if the following condition satisfies:

$$\Pi_j^*(S, N, N) > \Pi_j^*(N, N, N)$$

Assuming competitors using intertemporal price strategy S and set prices at $\{\underline{p}_{-j}, \bar{p}_{-j}\}$, firm j chooses inter-temporal price discrimination strategy S if the following condition satisfies:

$$\Pi_j^*(S, S, S) > \Pi_j^*(N, S, S)$$

The equilibrium solution depends on the conditions comparing the non-discriminating profit and the discriminating profit. If $\Pi_j^*(S, N, N) > \Pi_j^*(N, N, N)$, under the competitor's non-discriminating price strategy, firm j's best response is to intertemporally price discriminate. If $\Pi_j^*(S, S, S) > \Pi_j^*(N, S, S)$, under the competitor's cyclic pricing strategy, firm j's best response is also to inter-temporally price discriminate. Under these two conditions, intertemporal price discrimination by all firms is the unique Nash equilibrium in each period. If the game in each period has a unique equilibrium, backward induction shows that the unique perfect equilibrium of the finitely repeated game is to play the static equilibrium in every period of every subgame. Thus, under certain conditions, intertemporal price discrimination is the unique SPNE for the finite repeated game.

1.7 Identification and Estimation Strategy

I estimate the demand model employing the procedure of Hendel and Nevo (2013). This approach matches the predicted purchases to the observed purchases considering both consumers' responses to price changes and storers' stockpiling behavior. The observed purchase of product j in week t , x_{jst} is as follows:

$$x_{jst} = Q_{js}^{NS}(p_t) + x_{js}^{S}(p_{t-T}, ..., p_t, ..., p_{t+T}) + \epsilon_{jst}$$

where $Q_{js}^{NS}(p_t)$ is the purchase by non-storers at store s, x_{js}^{S} is the purchase by storers at store s, and ϵ_{jst} is an idiosyncratic error term. In this demand model, prices play two roles: influencing purchase amount and determining space state (non-sale N or sale S).

The product prices influence both non-storers and storers' absent-storage purchases through linear demand functions $Q_{js}^{NS}(p_t)$ and $Q_{js}^{S}(p_t)$. The amount purchased decreases as its price increases and increases as prices of other products increase.

In addition to consumers' absent-storage demand, prices determine the sale states:

non-sale N or sale S, which determines the storers' purchase timing decisions. For instance, I assume that storers can stockpile for one period to illustrate the intuition. If the last period is sale S_{t-1} and current period is sale S_t, storers have an inventory from last period $t-1$ and do not purchase for current period t. Storers choose to stockpile for consumption in the period $t+1$. The predicted purchase of storers $\hat{x}^S_{js}(p_{t-1}, p_t, p_{t+1})$ is the purchase for consumption in period $t+1$ as $Q^S_j(p_{jt}, p^{ef}_{-jt+1})$. The predicted purchase from both non-storers and storers as follows:

$$\hat{x}_{js}(p_{t-1}, p_t, p_{t+1}) = Q^{NS}_{js}(p_t) + Q^S_j(p_{jt}, p^{ef}_{-jt+1})$$

I match the predicted purchase $\hat{x}_{js}$ to the observed purchase x_{js} to identify the parameters of interest.

1.7.1 Identification

There are two sets of parameters of interest to be identified: own price effect β^h_j and cross price effects γ^h_{ij}, and share of sales by storers and non-storers ω^h. The joint distribution of $(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\omega})$ conditional on market-specific covariates is fixed across markets, with each market representing an independent draw from the distribution. I denote q_{jst} as the quantity sold of product j at store s in week t and p_{jst} as the price of product j at store s in week t. The econometrician always observes the quantity sold q_{jst}, the product retail price p_{jst}, and the histories of quantities sold $\{q_{js1}, ..., q_{jst-1}\}$ and product retail prices $\{p_{js1}, ..., p_{jst-1}\}$.

The model is identified if the joint distribution of products' own price effects, cross-price effects, and share of storers is uniquely identified by the joint distribution of observables. Formally, define a model as a pair $(\mathbb{P}, \Phi)$, where $\mathbb{P}$ is a set of joint distributions of the vector of $(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\omega})$, Φ is a set of mappings: $\phi : \mathbb{P} \rightarrow \mathbb{F}$, and $\mathbb{F}$ is a set of joint distributions of the vector of observables. Assuming the pair of set $(\mathbb{P}, \Phi)$ contains the true $(\mathcal{P}, \phi)$ which generates the observables.

Definition 1. *A model* $(\mathbb{P}, \Phi)$ *is identified if and only if for every* $(P, \hat{P}) \in \mathbb{P}^2$ *and* $(\phi, \hat{\phi}) \in \Phi^2$, $\phi(P) = \hat{\phi}(\hat{P})$ *implies* $(P, \phi) = (\hat{P}, \hat{\phi})$.

I consider identification of $(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\omega})$ which includes identification of consumer-type specific own price effects β_j^h and cross price effects γ_{ij}^h and the sale shares of type h consumers ω^h. The identification question reduces to whether the joint distribution of $(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\omega})$ can be determined when only quantities sold q_{jst} (not type h specific quantities sold q_{jst}^h) are observed.

Theorem 1.7.1. *The demand model is identified if (a)* ϵ_{jst} *has conditional mean zero:* $E(\epsilon_{jst}|\boldsymbol{p_{s1}}, ..., \boldsymbol{p_{sT}}, \boldsymbol{\alpha_s}) = 0$*; (b)* $(\boldsymbol{p_{s1}}, ..., \boldsymbol{p_{sT}}, \boldsymbol{\epsilon_{s1}}, ..., \boldsymbol{\epsilon_{sT}})$, $s = 1, ..., n$ *are i.i.d draws from the joint distribution; (c) large outliers are unlikely; (d) The independent variable matrix* $\boldsymbol{X}$ *has full rank.*

Proof. Under assumption (a) - (d), there exists a consistent estimator for a vector of parameters $(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\omega})$. Assumptions (a) - (c) are standard conditions for regressions which have been satisfied. The independent variable matrix consists of multiple independent variables with storage period τ. Each row of matrix $\boldsymbol{X}$ is

$$[\alpha_{js}, \sum_{k=0}^{\tau} 1(p_{jst} = p_{jst+k}^{ef})\alpha_{js}, p_{jst}, \sum_{k=0}^{\tau} 1(p_{jst} = p_{jst+k}^{ef})p_{jst}, p_{-jst}, \sum_{k=0}^{\tau} 1(p_{jst} = p_{jst+k}^{ef})p_{-jst+k}^{ef}]$$

Because $\sum_{k=0}^{\tau} 1(p_{jst} = p_{jst+k}^{ef})$ varies across time periods t, $\boldsymbol{\alpha_j}$ and $\sum_{\boldsymbol{k=0}}^{\boldsymbol{\tau}} \boldsymbol{1(p_{jt} = p_{jt+k}^{ef})\alpha_j}$, $\boldsymbol{p_{jt}}$ and $\sum_{\boldsymbol{k=0}}^{\boldsymbol{\tau}} \boldsymbol{1(p_{jt} = p_{jt+k}^{ef})p_{jt}}$, $\boldsymbol{p_{-jt}}$ and $\sum_{\boldsymbol{k=0}}^{\boldsymbol{\tau}} \boldsymbol{1(p_{jt} = p_{jt+k}^{ef})p_{-jt+k}^{ef}}$ are linearly independent vectors. Because $\boldsymbol{p_{jt}}$ and $\boldsymbol{p_{-jst}}$ varies across time periods under the

$\sum_{k=0}^{\tau} \mathbf{1}(\boldsymbol{p_{jt}} = \boldsymbol{p^{ef}_{jt+k}}) \neq \mathbf{0}$ condition, $\sum_{k=0}^{\tau} \mathbf{1}(\boldsymbol{p_{jt}} = \boldsymbol{p^{ef}_{jt+k}})\boldsymbol{\alpha_j}$, $\sum_{k=0}^{\tau} \mathbf{1}(\boldsymbol{p_{jt}} = \boldsymbol{p^{ef}_{jt+k}})\boldsymbol{p_{jt}}$ and $\sum_{k=0}^{\tau} \mathbf{1}(\boldsymbol{p_{jt}} = \boldsymbol{p^{ef}_{jt+k}})\boldsymbol{p^{ef}_{-jt+k}}$ are independent vectors. Because $\boldsymbol{p_{jt}}$ and $\boldsymbol{p_{-jst}}$ varies across time periods under the $\sum_{k=0}^{\tau} \mathbf{1}(\boldsymbol{p_{jt}} = \boldsymbol{p^{ef}_{jt+k}}) = \mathbf{0}$ condition, $\boldsymbol{\alpha_j}$, $\boldsymbol{p_{jt}}$ and $\boldsymbol{p_{-jt}}$ are linearly independent vectors. The independent variable matrix $\boldsymbol{X}$ is full rank. Assumptions (a) - (d) have been satisfied, and there exists a consistent GMM estimator for a vector of parameters $(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\omega})$.

Assuming $(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\omega})$ is not identifiable, there exist at least two different vectors of parameters $(\boldsymbol{\beta_1}, \boldsymbol{\gamma_1}, \boldsymbol{\omega_1})$ and $(\boldsymbol{\beta_2}, \boldsymbol{\gamma_2}, \boldsymbol{\omega_2})$. If $(\boldsymbol{\hat{\beta}_n}, \boldsymbol{\hat{\gamma}_n}, \boldsymbol{\hat{\omega}_n})$ is a consistent estimator of $(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\omega})$, then $(\boldsymbol{\hat{\beta}_n}, \boldsymbol{\hat{\gamma}_n}, \boldsymbol{\hat{\omega}_n})$ converges to both $(\boldsymbol{\beta_1}, \boldsymbol{\gamma_1}, \boldsymbol{\omega_1})$ and $(\boldsymbol{\beta_2}, \boldsymbol{\gamma_2}, \boldsymbol{\omega_2})$ in probability as $n \to \infty$. It is impossible since these are two different vectors, which contradicts. Thus, $(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\omega})$ is identifiable. $\square$

Consider storage period $\tau = 1$ and one product to illustrate the intuition, I consider two cases for current period: Sale (S_t) or Non-sale (N_t). When the current state is non-sale N_t, the state in previous period $t-1$ can be sale S_{t-1} or non-sale N_{t-1}. Under the $S_{t-1}N_t$ state, storers consume the inventory from last sale period and do not purchase in the period $(\sum_{k=0}^{1} 1(p_{jt} = p^{ef}_{jt+k}) = 0)$, while non-storers purchase for current consumption, the observed purchase is the non-storers' purchases, $x_{jt}(S_{t-1}N_t) = Q^{NS}_j(p_t)$. The variation in prices and quantities sold under state combination $S_{t-1}N_t$ identifies the own and cross price effects for non-storers.

Similarly, the variation in prices and quantities sold under state combinations $N_{t-1}N_t$ and $S_{t-1}N_t$ identifies the own and cross price effects for storers. Under the $N_{t-1}N_t$ state combination $(\sum_{k=0}^{1} 1(p_{jt} = p^{ef}_{jt+k}) = 1)$, both storers and non-storers purchase for current consumption, $x_{jt}(N_{t-1}N_t) = Q^{NS}_j(p_t) + Q^{S}_j(p_t)$. Subtracting $x_{jt}(S_{t-1}N_t)$ from $x_{jt}(N_{t-1}N_t)$, I find $Q^{S}_j(p_t) = x_{jt}(N_{t-1}N_t) - x_{jt}(S_{t-1}N_t)$. The variation in prices and quantities sold using the calculated $Q_j^S(p_t)$ identifies the parameters of price effects.

To identify the share of storers and non-storers, I exploit the variation in quantities sold across different state space. Consider storage period $\tau = 1$ and one product to illustrate the intuition, I consider two cases for current period: Sale (S_t) or Non-sale (N_t). When the current period is sale (S_t), the previous periods can be sale (S_{t-1}) or non-sale (N_{t-1}).

Under the condition that the current state is sale (S_t), if the previous period is sale (S_{t-1}), storers have inventory from the last period and only purchase for consumption in the next period $x_{jt}^S(S_{t-1}S_t) = Q_j^S(p_t)$. If the previous period is non-sale (N_{t-1}), storers purchase for both current and future consumption $x_{jt}^S(N_{t-1}S_t) = Q_j^S(p_t)+Q_j^S(p_t)$. Under both conditions, non-storers purchase for current consumption only, which means $x_{jt}^{NS}(S_{t-1}S_t) = x_{jt}^{NS}(N_{t-1}S_t) = Q_j^{NS}(p_t)$. Combining the purchases of storers and non-storers, $x_{jt}(S_{t-1}S_t)$ when $\sum_{k=0}^{1} 1(p_{jt} = p_{jt+k}^{ef}) = 1$ and $x_{jt}(N_{t-1}S_t)$ when $\sum_{k=0}^{1} 1(p_{jt} = p_{jt+k}^{ef}) = 2$ are as follows:

$$
\begin{aligned}
x_{jt}(S_{t-1}S_t) &= x_{jt}^S(S_{t-1}S_t) + x_{jt}^{NS}(S_{t-1}S_t) = Q_j^S(p_t) + Q_j^{NS}(p_t) \\
x_{jt}(N_{t-1}S_t) &= x_{jt}^S(N_{t-1}S_t) + x_{jt}^{NS}(N_{t-1}S_t) = 2Q_j^S(p_t) + Q_j^{NS}(p_t)
\end{aligned}
$$

The purchase increase from state $N_{t-1}S_t$ to state $S_{t-1}S_t$ comes from storers' anticipation of future consumption. The variation of quantities sold across state $N_{t-1}S_t$ and state $S_{t-1}S_t$ identifies the share of storers at the store level. $x_{jt}(S_{t-1}S_t)$ and

$x_{jt}(N_{t-1}S_t)$ imply the $Q_j^{NS}(p_t)$ and $Q_j^S(p_t)$ as follows:

$$Q_j^{NS}(p_t) = 2x_{jt}(S_{t-1}S_t) - x_{jt}(N_{t-1}S_t)$$
$$Q_j^S(p_t) = x_{jt}(N_{t-1}S_t) - x_{jt}(S_{t-1}S_t)$$

Using the same approach, under the condition that the current state is non-sale (N_t), I use the purchases from state combination $N_{t-1}N_t$ when $\sum_{k=0}^{1} 1(p_{jt} = p_{jt+k}^{ef}) = 1$ and $S_{t-1}N_t$ when $\sum_{k=0}^{1} 1(p_{jt} = p_{jt+k}^{ef}) = 0$ to identify the share of storers, where I find that:

$$x_{jt}(S_{t-1}N_t) = Q_{jt}^{NS}(p_t)$$
$$x_{jt}(N_{t-1}N_t) = Q_{jt}^{NS}(p_t) + Q_{jt}^S(p_t)$$

Similarly, the purchase increase from state $S_{t-1}N_t$ to $N_{t-1}N_t$ comes from storers' inventory from the last period. For longer storage period ($\tau > 1$), I follow the same intuition to condition on a longer price history to identify heterogeneous price effects of storers and non-storers and the share of storers.

1.7.2 Estimation

For estimation, if I assume all consumers are non-storers, the demand for product j at store s in week t is linear:

$$q_{jst} = \alpha_{sj} - \beta_j p_{jst} + \sum_{i \neq j} \gamma_{ji} p_{ist} + \epsilon_{jst}$$

where α_{sj} is a store-product level fixed effect, p_{jst} is product j's own price, p_{ist} is other product price, and ϵ_{jst} is an i.i.d shock. β_j captures the own-price effect, while γ_{ji} captures the cross-product price effect from product i. The product j's demand is a function of its own price p_{jst} and other products' prices p_{ist}.

If I assume consumers are heterogeneous in storage abilities, I assume that consumer type h's demand for product j at store s in week t is linear as follows[15]:

$$q_{jst}^h = \omega^h \alpha_{sj} - \beta_j^h p_{jst} + \sum_{i \neq j} \gamma_{ji}^h p_{ist} + \epsilon_{jst}$$

where β_j^h is the type h specific own-price effect, and γ_{ji}^h is the type h specific cross-product price effect from the product i. ω^h stands for the share of type h consumers $(h = S, NS)$. ω^h, β_j^h and γ_{ji}^h allow different consumer type has different intercepts, own-price and cross-price effects. I scale the parameters ω^S and ω^{NS} and assume that $\omega^S + \omega^{NS} = 1$. The parameters ω^S and ω^{NS} represent the fraction of sales by storers and non-storers when all prices are equal to zero. In addition, I impose one restriction on the symmetry of cross-product price effect: $\gamma_{ji} = \gamma_{ij}$[16].

Combining the predicted purchase behavior function of x_{jt}^S and consumer type h specific demand function, I minimize the difference between predicted purchases and observed purchases by the Generalized Methods of Moments approach. One main concern is the endogeneity of prices, where prices are correlated with the error term. For instance, some unobserved product characteristics are positively correlated with the product price and lead to higher demand for these products. The positive correlation

[15] I add the superscript h to indicate different consumer type

[16] The main findings on share of storers and difference in price sensitivity across storers and non-storers are robust without this assumption. I keep this assumption because of the symmetric switching behavior across different products.

between the error term and the product price leads to overestimating its own price effect, which I should use instrument variable for correcting the correlation between the product price and the error term on the demand side. One popular identification approach is to use a product's price in other markets as instrument variables, under the theory that cross-market correlation in the price of a given product will be due to common cost factors instead of unobserved features from the demand side. An alternative approach is to use the nonprice characteristics of other products as an instrument variable to capture the cost-side correlation. In this application, I exploit the panel structure of dataset and use store-product specific fixed effects to incorporate the unobserved constant factors from the demand side. In addition, I model the change of demand due to storers' participation to control for the demand change of promotions. By exploiting the panel structure of dataset and modeling demand changes across time, the endogeneity of prices (correlation between price and error term) is not the main concern in this application.

Another concern is that product quantities in each period are related to both current prices and previous prices, which violates the i.i.d assumption of pooled OLS approach. By incorporating fixed effects and correcting auto-correlations using GMM approach, Newey (1985) proved that the linearized GMM estimator is consistent and asymptotically efficient, which allows me to apply the GMM approach and estimate the parameters of interest.

1.8 Estimation Results

1.8.1 Demand Side

Table 1.7 reports the estimates for a static demand system. I derive linear demand functions from the quasi-linear utility function. The dependent variable is weekly sales of products (in cases) at the store level. All regressions include product-specific store fixed effects. Each observation is at the store-week-product level.

Table 1.7 presents the estimates for a static demand system of three major flagship brands, and Table 1.8 reports the estimates for a dynamic demand system of the same major flagship brands. Comparing the estimates from Table 1.7 and 1.8, estimates of the static demand system are more sensitive to its own and competitors' price changes. The intuition is that static demand model assumes that consumers consume all the purchases in current periods and interpret the stockpiling behavior as increasing consumptions, which leads to an overestimation of price effects.

In the dynamic demand system, actual product prices are used as continuous variables and to classify the states: sale (S) or non-sale (N). Table 1.8 displays the demand estimates for storers and non-storers. There are statistically significant shares of storers for all brands. It supports the argument that consumers are heterogeneous in purchasing behavior and storage costs. Ignoring the heterogeneity in consumer stockpiling behavior may overestimate the demand elasticities. In addition, I allow that three major brands have different shares of stockpiling consumers. The Bud Light has the largest share of storers (around 25.8 percent), while Miller Lite and Coors Light have similar shares of storers (each has around 13 percent). Allowing heterogeneity in the shares of storers enables to incorporate unobserved heterogeneity of beer brands.

For each brand, storers are more price sensitive than non-storers. It fits into the intuition that price-sensitive storers tend to shop in the sale periods and stockpile for future consumption. Firms serve both storers and non-storers in the sale periods and only serve non-storers in the non-sale periods. This difference in the price sensitivity supports the inter-temporal price discrimination motivation.

In addition, storers are more likely to switch between brands than non-storers. Considering the switching behavior of storers, firms with multiple close-substitute products can set one product on promotion to attract easy-switching storers, while setting other products on the regular price for relatively brand-loyal non-storers. This difference in brand switching behavior provides the possibility of asymmetric promotions for close-substitute products as firms own and operate more close-substitute brands. As Miller and Coors formed a joint venture, the joint venture MillerCoors own and operate more close-substitute products. As the markets become more concentrated, firms tend to use alternating promotions for close-substitute products.

The main findings on the price sensitivity and switching behavior between storers and non-storers are robust to different storage assumptions. I estimate the demand system under different storage assumptions. For instance, $T = 1$ means that storers can stockpile beer for one week, and shop twice a month. $T = 2$ means that storers can stockpile and leave the market for two weeks, and $T = 3$ implies that storers can stockpile for one month and can choose to purchase at most once a month. The regression results under different assumptions are similar. There exist statistically significant shares of stockpiling consumers for major flagship beer brands. Storers are more price sensitive than non-storers and are more easily to switch between brands. The main change is that the share of storers decreases, the price sensitivity of non-storers decreases, and the price sensitivity of storers decreases.

1.8.2 Supply Side

Table 1.9 compares the calculated marginal costs considering price discounts (Panel: Dynamic Estimates) and ignoring price discounts (Panel: Static Estimates). The marginal costs calculated ignoring price discounts are slightly higher than the marginal costs calculated considering price discounts for all three brands. The difference in the marginal costs can be explained by the existence of storers. Ignoring price discounts implicitly assumes that consumers face the same constant price over periods. Facing the same constant price over periods, there is no difference to purchase the products in any periods, which implies that consumers' current purchases are independent of the previous purchase history. When consumers do not purchase any products in a period, it means that consumers choose the outside option, not to buy anything. However, when I consider price discounts, consumers have incentives to shop in the period with promotional prices and stockpile for future consumption. When consumers do not purchase any products in a period, besides choosing not to buy anything, it could be the case that consumers are waiting for promotional prices. Therefore, ignoring price discounts leads to an overestimation of the percentage of sales decreased due to consumer switching to the outside option, which leads to an overestimation of demand elasticity and an overestimation of marginal costs. The marginal cost difference between static estimates and dynamic estimates for 12-pk Bud Light is \$1.11, while the differences in marginal costs for 12-pk Miller Lite and Coors Light are \$0.76 and \$0.58, respectively. The brand with a larger percentage of sales by storers has a larger bias in demand estimation, leading to a larger difference between static and dynamic demand estimates. Table 1.8 shows that Bud Light, Miller Lite, and Coors Light have 25.8 percent, 13.2 percent, and 13.1 percent of sales attributed to storers, which leads to various gaps between static and dynamic estimates. Omitting the dynamic feature in the demand estimation can lead to a substantial bias in markup and marginal cost calculations.

In addition, Table 1.9 presents the markups and markup ratios of sale prices and regular prices. Assuming constant marginal cost, Bud Light, Miller Lite, and Coors Light increase the markup ratios by at least 6.68 percent from sale weeks to regular price weeks. The main factor in the marginal cost of beer products is the shipping and distribution cost (Miller and Weinberg (2016)). The U.S. Department of Justice (DOJ) approved the joint venture between SAB Miller and Molson Coors, the second and third largest brewers in the U.S. domestic beer markets partially because it was expected to reduce the shipping and distribution costs (Heyer et al. (2008)). Ashenfelter et al. (2015) cite the reports of Bernstein Research claiming that the marginal cost will be decreased by 8 and 11 percent after the merger, but they find that the efficiency created by the merger leads to around 2% decrease in average prices, which was offset by the equal and opposite effect from market concentration. Whether the efficiency in marginal costs has been realized and the size of efficiency is not the main focus of this paper. To focus on the effect of the change in price discounts strategies, I will assume marginal costs are constant and stay the same from the pre to post-merger periods. Since the focus is welfare effects with/without considering price discounts, assuming constant marginal costs will not lead to any bias in consumer surplus calculation.

To solve the symmetric Subgame Perfect Nash Equilibrium[17], I compare the pos-

[17] In a finite R-period repeated game, I solve the game using backward induction and start from

sible profits for each strategy profile and use backward deduction to characterize the symmetric subgame perfect equilibrium. Based on these estimates of marginal costs in Table 1.9, I calculate and compare the firms' profits under the non-discrimination price strategy N and inter-temporal price discrimination price strategy S to study firms' best responses. Assuming the marginal cost during the sale and non-sale periods are the same, I first assume the competitors using a non-discriminating price strategy and calculate the firm's profits using Non-discrimination (N) and Sale (S) strategies (Panel: No Price Discrimination by Competitors). Then I assume that the competitor uses an inter-temporal price discrimination strategy S and calculate the firms' profits under N and S strategies (Panel: Price Discrimination by Competitors).

Table 1.10 presents the prices and profits regarding each strategy profile. The upper panel displays the comparison of profits given competitors using constant prices. All firms benefit from switching to price discount strategies S. For instance, Bud Light increases its profit by 17.18 percent by switching to sales strategy (S), Miller Lite increases around 19.10 percent profit by switching to sales strategy (S), and Coors Light benefits from switching to sales strategy (S) with an 18.96 percent profit increase. The lower panel presents the profits given competitors using the inter-temporal price discrimination strategy. By switching from non-discrimination strategy (N) to sales strategy (S), Bud Light increases the profit by 15.35 percent, Miller Lite has a 22.42 percent profit gain, and Coors Light increases the profit by 6.26 percent. In both panels, firms can increase their profits by using the intertemporal price discrimination strategy. Regardless of their competitors' pricing strategy, firms should use the inter-

the last period. Because the symmetric Subgame Perfect Nash Equilibrium consists of the equilibria which are Nash equilibrium in every subgame. If each period consists of a unique equilibrium, the unique equilibrium is the Subgame Perfect Nash Equilibrium. To solve the Nash equilibrium in each subgame, I solve the best-response function given competitors' strategies.

temporal price discrimination strategy. The main reason is the existence of relatively large percent of sales by price-sensitive storers. When firms choose the inter-temporal price discrimination strategy, firms can only serve price-insensitive non-storers, leading to higher prices and higher markups. Comparing the profits in the upper and lower panels, the profits of all firms using the inter-temporal price discrimination outweigh the profits that all firms only use the constant price strategy.

Table 1.10 shows that all firms choose the inter-temporal price discrimination strategy S in the equilibrium. Because of the existence of price-sensitive stockpiling consumers, non-discrimination price strategy weakens the firms' ability to charge higher prices for non-stockpiling consumers. When the group of storers are large enough and price sensitive enough, firms benefit from inter-temporally separating consumers and charging higher prices when price-sensitive consumers are absent in the market.

1.9 Counterfactural Analysis

To understand the role of only setting one of two close-substitute products on sale (asymmetric price discount strategy) in welfare analysis, I simulate the welfare effects of the horizontal merger under three scenarios: ignoring consumers' possibility in stockpiling (static demand setting), allowing consumers equipped with heterogeneous storage capacities (dynamic demand setting), and allowing consumers equipped with heterogeneous storage capacities and considering asymmetric price discount strategy (dynamic demand setting and asymmetric price discount strategy).

To motivate and quantify the incentives to use the asymmetric price discount strategy, I illustrate the profit gains of applying this strategy in the first subsection.

I find that both merged firm and its competitor benefit from the asymmetric price discount strategy by the merged firm.

I simulate and compare the welfare effects under three different assumptions. In the first two settings (static demand setting and dynamic demand setting), I only revise the ownership matrix to internalize the business stealing effects between merged firms. To simulate the welfare effects of the third setting (dynamic demand setting and asymmetric price discount strategy), I turn off the promotion for one of two close-substitute products and simulate the welfare effects. The welfare effect on consumers is the change in consumer surplus. Because I employ a quasi-linear utility function, compensation variation, equivalent variation, and consumer surplus are algebraically equivalent. I use the change in equivalent variation as the change in consumer surplus.

1.9.1 Asymmetric Price Discount

I illustrate the merged firm's benefits to use asymmetric price discount strategy for two close-substitute products by comparing the profits under different pricing strategies. Previous literature (Sweeting (2010), Fan (2013)) has shown that firms choose to re-position close-substitute products, and change product characteristics after the market structure changes. The intuition is to differentiate products and reduce competition between products under the same ownership. Firms use temporary price discounts to compete with other players in the market. After the joint venture formed, the merged firm owns two close-substitute products like Miller Lite and Coors Light. The merged firm may benefit from only setting one of two close-substitute products on sale (asymmetric price discount strategy). In addition, the merged firm may choose to set the products on sale back to the regular price to induce price-sensitive storers only shop in the promotion weeks. In regular price weeks, firms can only serve price-insensitive non-storers, which constitutes an intertemporal price discrimination strategy.

Table 1.11 presents the gains from using the asymmetric price discount strategy in the promotion periods compared with the symmetric price discount strategy for all three brands. In the promotion periods, if the merged firm sets Coors on promotion only, the total profit of Miller Lite and Coors Light increases by 9.34%. By setting Miller Lite at the regular price, the joint venture MillerCoors loses 11.21% of the profit from the Miller Lite, while it increases the profit from Coors Light by 24.79%, leading to a 9.34% gain in total profit. The merged firm benefits from only setting Coors on sale as the market becomes more concentrated. In both cases, the competitor of the merged firm benefits from this adjustment of simultaneous promotions. The intuition is that the competitor benefits from the sale periods because fewer products are on sale and the profit share of non-storers increases, leading to an increase in the total profit. The change of ownership not only increases the prices for sale and non-sale periods but also provides firms with the potential gains by adjusting simultaneous promotions.

1.9.2 Welfare Effects

In this section, I simulate and compare the welfare effects under different settings: static demand setting, dynamic demand setting, and dynamic demand setting with asymmetric price discount strategy. As I show in the previous section, the merged firm benefits from using asymmetric price discount strategy in the promotional weeks. In this section, I compare the consumer surplus under these three settings.

Table 1.12 presents the changes in consumer surplus (equivalence variation). When I ignore the adjustment in promotional strategy, consumers are willing to pay on average \$9.51 per store in a sale period, and on average \$19.43 per store in a regular period to avoid the merger. In both periods, product prices increase after the merger, which comes from the merged firms' internalization of business stealing effects. Since storers only shop in the sale period, while the non-storers shop in both sale and non-sale periods. Non-storers suffer from a larger consumer surplus loss. In total, storers and non-storers are willing to pay about \$30 per cycle per store to avoid the merger. When I consider the adjustment in promotion strategy, consumers are willing to pay on average \$81.60 per store in a sale period, and on average \$19.43 per store in a regular period. Both storers and non-storers suffer from losses in consumer surplus. When considering the adjustment in promotion strategy, consumers are willing to pay additional \$72.09 per store in a sale period to avoid the merger. Ignoring the change of promotion strategies leads to substantial bias in the calculation of consumer surplus.

Table 1.13 presents the proportion of consumer surplus loss by storers and non-storers in the sale period. When ignoring the adjustment of promotion strategy, storers only consists of around 5.04 percent of consumer surplus loss after the merger, while non-storers contributes to around 94.96 percent of consumer surplus loss in the sale period after the merger. However, when considering the promotion adjustment, storers' loss in consumer surplus increases both in quantity and percentage (from 5.04% to 10.08%). The reason is that there are fewer products on sale in the promotion periods. Non-storers' loss in consumer surplus increases in quantity but decreases in share (from 94.96% to 89.02%). Considering the change in promotion strategy, storers' loss in consumer surplus takes a larger share than the case ignoring promotion change.

Table 1.14 displays the price effects for the sale and regular periods after the merger. When ignoring the change in promotional strategies, the predicted prices increases \$0.36 per case on average, while regular periods have larger predicted price increases \$0.80 per case on average. However, when considering the change in promotional strategies, the sale periods have larger increases in predicted prices than regular periods. The simulated price changes considering the change of promotion strategies are compatible with observed data trends after the merger.

1.10 Conclusion

In this paper, I study how market structure affects firms' promotion decisions and the welfare implications, using the evidence from Miller/Coors joint venture in the US brewing industry. I document a change of promotion decisions from pre- to post-merger periods. The change of promotion decisions mainly comes from that firms switch from simultaneous promotions to non-simultaneous promotions.

To incorporate the change of promotion decisions, I develop a structural model of dynamic demand with consumer storage heterogeneity including storers and non-storers. Storers respond to the price changes by adjusting the purchase amount and the purchase timing, while non-storers respond to the price changes by adjusting purchase quantity only. On the supply side, firms compete in product prices and the timing to set promotions. Assuming a pure strategy equilibrium exists, firms choose to set simultaneous promotions and simultaneous regular prices in the pre-merger period. After the merger, firms adjust the promotion decisions by only setting one product on sale. Ignoring the change of price promotion decisions underestimates the consumer surplus loss. Consumers suffer from additional \$72 consumer surplus loss on average per store in a period. In addition, the Horizontal Merger Guidelines (2010) suggests that "the anti-trust agencies should consider whether possible anti-competitive effects vary significantly for different customers purchasing the same or similar products". When considering the change of promotion strategies, storers suffer from a larger share of consumer surplus loss than the case ignoring the change of promotion strategies. If storers are associated with specific social economic status, ignoring the change in promotion strategies may underestimate the price effects on this group after the merger.

1.11 Appendix

1.11.1 Figures

1.11.2 Tables

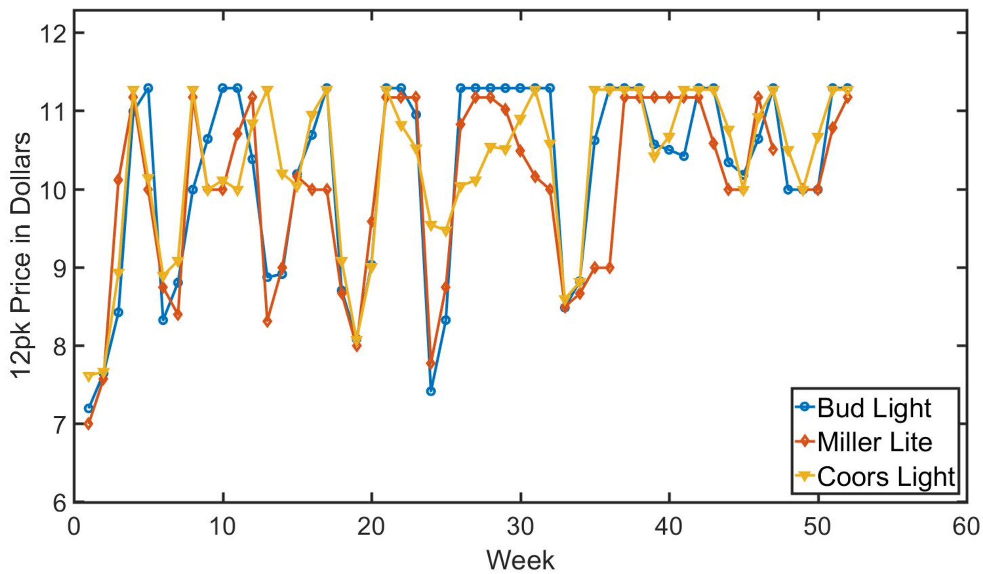

Figure 1.1: Pre-merger Simultaneous Price Promotions

Notes: The graph shows the prices of 12pk Bud Light, Miller Lite and Coors Light in 2006 (pre-merger) at an example store. The example store is in Sonoma, California. The store sold around 40 cases of 12-packs top three light beers per week in 2006. Each point is the observed product price per week at the example store.

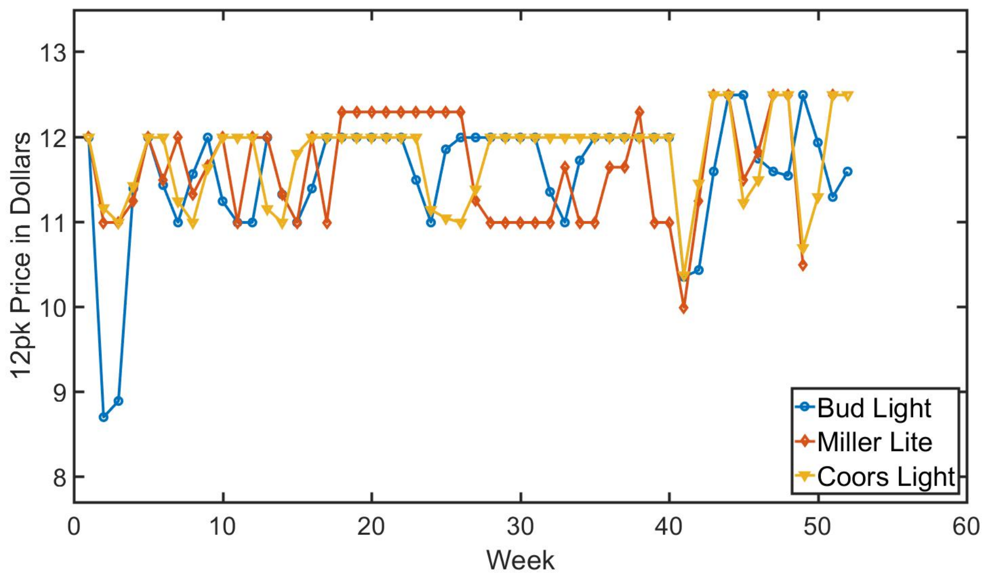

Figure 1.2: Post-merger Non-Simultaneous Price Promotions

Notes: The graph shows the prices of 12pk Bud Light, Miller Lite and Coors Light in 2009 (post-merger) at an example store. The example store is in Sonoma, California. The store sold around 23 cases of 12-packs top three light beers per week in year 2009.Each point is the observed product price per week at the example store.

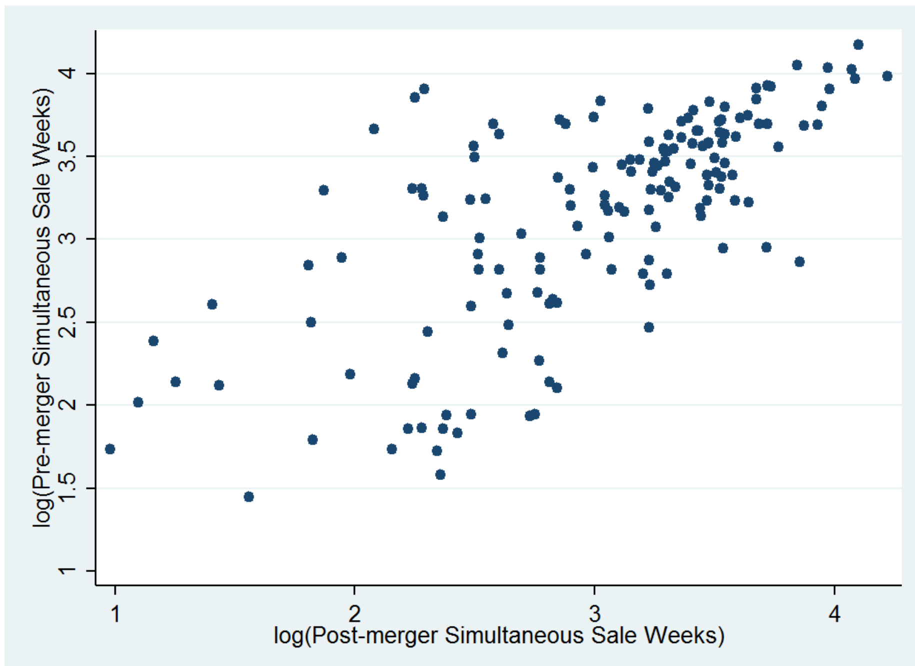

Figure 1.3: Scatter Plot of Pre and Post Merger Simultaneous Weeks(in log)

Notes: The figure shows the relationship between pre and post-merger simultaneous sale weeks in a log format. I define a week of at least one store in a DMA region with three top brands (Bud Light, Miller Lite, and Coors Light) labeling sale as a simultaneous sale week. Each observation corresponds to a pair of pre and post-merger simultaneous sale weeks in a log format by DMA. DMA classification is from Nielsen Scanner Data.

Table 1.1: Descriptive Statistics

	Price ($)			
	Mean	Std	1%	99%
Bud Light	10.19	1.15	7.80	13.25
Miller Lite	10.09	1.15	7.55	13.25
Coors Light	10.22	1.20	8.03	13.25
Michelob Ultra Light	10.98	1.28	8.32	14.69
Corona Light	14.41	1.53	10.99	17.99

Notes: The table presents the summary statistics on 1,820,520 observations from around 7800 stores from 2006 to 2010. Each observation is a 12pk product price in dollars.

Table 1.2: Pre vs Post Merger Conditional Probabilities

	All Stores		
	Pre-Merger Mean	Post-Merger Mean	Difference Mean
Pr(2 or 3 Brands on Sale | Sale)	0.5657	0.5008	0.0649***
Pr(3 Brands Sale | Sale)	0.2343	0.1985	0.0358***
Pr(Miller&Coors Sale | Sale)	0.3259	0.3025	0.0234***
Pr(Budweiser&Miller Sale | Sale)	0.3454	0.2808	0.0646***
Pr(Budweiser&Coors Sale | Sale)	0.3629	0.3145	0.0484***
	Frequent Simultaneous Sale Stores		
	Pre-Merger Mean	Post-Merger Mean	Difference Mean
Pr(2 or 3 Brands on Sale | Sale)	0.7971	0.6584	0.1387***
Pr(3 Brands Sale | Sale)	0.582	0.4065	0.1755***
Pr(Miller&Coors Sale | Sale)	0.6432	0.5093	0.1339***
Pr(Budweiser&Miller Sale | Sale)	0.646	0.4591	0.1869***
Pr(Budweiser&Coors Sale | Sale)	0.672	0.5029	0.1691***

Notes: The table presents the conditional probabilities of 2 or 3 brands on sale, three brands on sale, Miller&Coors on sale, Budweiser&Miller on sale, Budweiser&Coors on sale. The probabilities are calculated at the store level. In the upper panel, all stores are included. In the lower panel, only stores with frequent simultaneous sales are included ($Pr(\text{Simultaneous Sales}|\text{Sales}) \geq 0.5$). The *** indicates significant at 0.1% level.

Table 1.3: Pre and Post Merger Sale Frequency

	Pre-merger Sale Frequency
Post-merger Sale Frequency	1.009***
	(0.00540)
R^2	0.99
Observations	165

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: The table presents the relationship between pre-merger sale frequency and post-sale frequency by DMA. We define the weeks with at least one store in a DMA region providing price reductions as a sale, and define as a non-sale otherwise. Each observation is the pre and post-merger sale weeks per month by a DMA. The sample includes 165 geographically DMA in the Nielsen Scanner Data.

Table 1.4: Revenue Market Shares

Year	AB InBev	MillerCoors	Miller	Coors	Modelo	Heineken	Total
2006	0.4305	0.2955†	0.1801	0.1154	0.1072	0.0621	0.8953
2007	0.4246	0.2981†	0.1793	0.1188	0.1049	0.0617	0.8893
2009	0.4259	0.306			0.0946	0.0563	0.8828
2010	0.424	0.3002			0.0935	0.0558	0.8735

Notes: The table presents the revenue-based market shares of five largest brewers in the light beer market from 2006 to 2010. The revenue-based market shares marked with † are sum of Miller and Coors shares. Due to the Miller/Coors joint venture formed in the middle of 2008, we do not report the market shares in year 2008. Product ownership information is collected from the brewer's official website. Each firm's market share is calculated at the nation level. AB Inbev carries major flagship brands including Bud Light, and Beck's. Miller's major flagship beer brands are Miller Lite, and Miller High Life. Coors brews flagship beers including Coors Light. Modelo and Heineken produces flagship beer brands Model and Heineken respectively.

Table 1.5: Regression results at sample markets

	(1) Miller Lite	(2) Miller Lite
L.Miller Lite	0.928***	0.431***
	(0.0622)	(0.0547)
L2.Miller Lite	-0.176**	-0.148**
	(0.0588)	(0.0522)
Coors Light	0.311***	0.266***
	(0.0722)	(0.0629)
Coors Light*Post	-0.233*	0.171
	(0.0911)	(0.0940)
Bud Light	-0.00644	0.425***
	(0.0953)	(0.0678)
Bud Light*Post	0.0185	-0.272**
	(0.105)	(0.0912)
Michelob Light	0.0339	0.0209
	(0.113)	(0.0419)
Corona Light	-0.0615	0.0377
	(0.0653)	(0.0367)
Michelob Light*Post	0.103	0.0997
	(0.126)	(0.0807)
Corona Light*Post	0.0748	0.000245
	(0.0684)	(0.0474)
Holiday Dummies	0.0631	0.105
	(0.0629)	(0.141)
Constant	0.00396	-0.714
	(0.930)	(0.904)
Observations	260	260

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: The table presents the time series regression results at two representative markets. Each observation is at DMA-product-week level. The optimal lag number is selected through AIC and BIC criteria. There are 260 weeks from 2006 to 2010.

Table 1.6: Summary of Regression Results

		Bud Light Change 1	Bud Light Change 0
Coor Light Change	1	8.06%	13.71%
	0	42.74%	35.48%

Notes: The table presents the summary of DMA-level time series regression results. Each observation is classified by the DMA-level regression result. There are 124 DMAs with at least five stores. When the coefficient before Bud interacted term is negative and significant at ten percentage level, the DMA is coded as 1 in Bud Light Change part. When the coefficient before Coors Light interacted term is negative and significant at ten percentage level, the DMA is coded as 1 in Coors Light Change part.

Table 1.7: Estimates of Static Demand

	(1) Bud Q	(2) Miller Q	(3) Coors Q
Bud Light Price	-4.603***	1.683***	1.582***
	(0.0106)	(0.00822)	(0.00839)
Miller Lite Price	1.683***	-3.456***	0.225***
	(0.00822)	(0.0113)	(0.00943)
Coors Light Price	1.582***	0.225***	-1.786***
	(0.00839)	(0.00943)	(0.0116)
Observations	841,372	841,372	841,372

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: The table presents the estimates of products' static demands. All estimates are from least-squares regressions. The dependent variable is the quantity of Bud Light, Miller Lite or Coors Light sold at a store in a week. Each observation is at store-week-product level. All columns include store fixed effects. The sample includes weekly data from around 7800 geographically representative stores over 2006-2010.

Table 1.8: Estimates of Dynamic Demand

	(1) Bud Q	(2) Miller Q	(3) Coors Q
share of storers	0.258***	0.132***	0.131***
	(0.000762)	(0.000487)	(0.00168)
nonstorer bud price	-1.270***	0.917***	0.582***
	(0.0127)	(0.00998)	(0.0103)
nonstorer miller price	0.917***	-0.996***	0.116***
	(0.00998)	(0.0130)	(0.0104)
nonstorer coors price	0.582***	0.116***	-0.647***
	(0.0103)	(0.0104)	(0.0131)
storer bud price	-2.585***	1.413***	0.948***
	(0.00868)	(0.00740)	(0.00719)
storer miller price	1.413***	-2.196***	0.705***
	(0.00740)	(0.01000)	(0.00788)
storer coors price	0.948***	0.705***	-1.669***
	(0.00719)	(0.00788)	(0.00999)
Observations	841,372	841,372	841,372

Notes: The table presents the estimates of products' dynamic demand under the assumption T = 3. Assumption T = 3 indicates that consumers are flexible enough to stockpile for as long as three weeks. All estimates are from least-squares regressions. The dependent variable is the quantity of Bud Light, Miller Lite or Coors Light sold at a store in a week. Each observation is at store-week-product level. The sample includes weekly data from around 7800 geographically representative stores over 2006-2010.

Table 1.9: Estimates of Marginal Cost and Markup

	Static Estimates			
	MC($)	Price($)	Markup($)	Markup %
Bud Light	7.35	9.99	2.63	26.38
Miller Light	7.88	9.94	2.06	20.69
Coors Light	6.63	10.10	3.47	34.34
	Dynamic Estimates			
	MC($)	Sale Price($)	Markup($)	Markup %
Bud Light	6.24	9.11	2.87	31.51
Miller Light	7.12	8.90	1.78	20.02
Coors Light	6.05	9.25	3.20	34.58
	Dynamic Estimates			
	MC($)	Regular Price($)	Markup($)	Markup %
Bud Light	6.24	10.18	3.95	38.74
Miller Light	7.12	10.18	3.07	30.11
Coors Light	6.05	10.30	4.25	41.26

Notes: The table presents the estimates of marginal costs per 12pk case in dollars under static demand estimates and dynamic demand estimates. Demand estimates are from columns 1, 2 and 3 of Table 1.7 and columns 1, 2 and 3 of Table 1.8. Marginal costs are calculated at store level, and the weighted average of marginal costs are reported in the table. Prices are weighted-average prices over stores. The sample includes weekly data from around 7800 geographically representative stores over 2006-2010.

Table 1.10: Gains from Sales (per store per week)

Gains from Sales (No Price Discrimination by Competitors)

	Bud Light		Miller Lite		Coors Light	
	Price($)	Profit($)	Price($)	Profit($)	Price($)	Profit($)
Non-discrimination	9.17	66.32	9.02	23.00	9.03	41.12
Sales Strategy		77.73		27.40		48.92
Regular	11.92		10.95		12.25	
Sale	8.63		8.66		8.51	

Gains from Sales (Price Discrimination by Competitors)

	Bud Light		Miller Lite		Coors Light	
	Price($)	Profit($)	Price($)	Profit($)	Price($)	Profit($)
Non-discrimination	8.73	109.91	8.54	44.56	8.56	68.31
Sales Strategy		126.78		54.55		72.58
Regular	15.21		14.09		15.42	
Sale	8.20		8.19		8.05	

Notes: The table presents the gains from inter-temporal price discrimination given competitors use constant pricing and sales strategy of 12pk product per store per week. Demand estimates are from columns 1, 2 and 3 of Table 1.7 and columns 1, 2 and 3 of Table 1.8. Each firm column shows the price and profit in dollars per week per store under different regimes. Regular stands for the regular price, and Sale stands for the sale price. The column labeled Profit is the profit in dollars per store per week in each regime. The marginal costs are calculated at the store level, and the weighted average of marginal costs across stores and states are used for the optimal price calculation.

Table 1.11: Gains from Asymmetric Promotion

	Profits Change			
	Bud Light	Miller Lite	Coors Light	MillerCoors
Coors Promotion only	30.13%	-11.21%	24.79%	9.34%
Miller Promotion only	24.36%	21.97%	-21.89%	-3.07%

Notes: The table presents the gains from inter-temporal price discrimination. Demand estimates are from columns 1, 2 and 3 of Table 1.7 and columns 1, 2 and 3 of Table 1.8. Each firm column shows the profit changes in percentage under different regimes. Regular stands for the regular price, and Sale stands for the sale price.

Table 1.12: Dynamic Demand Merger Effect (per store per week)

	No Promotion Adjustment			Promotion Adjustment		
	SALE			SALE		
	Pre Price	Post Price	Δ EV	Pre Price	Post Price	Δ EV
Bud Light	8.20	8.44	-2.95	8.20	10.26	-23.37
Miller Lite	8.19	8.68	-2.85	8.19	15.03	-47.52
Coors Light	8.05	8.51	-3.71	8.05	9.82	-10.71
	NON-SALE			NON-SALE		
	Pre Price	Post Price	Δ EV	Pre Price	Post Price	Δ EV
Bud Light	15.21	15.79	-6.57	15.21	15.79	-6.57
Miller Lite	14.09	15.03	-6.49	14.09	15.03	-6.49
Coors Light	15.42	16.47	-6.37	15.42	16.47	-6.37

Notes: The table presents the simulated product prices of 12pk products after a merger. The change in Equivalent Variation is reported in dollars per store per week. Demand estimates are from columns 1, 2 and 3 of Table 1.8. Firms do not change the promotion decisions after a merger. Sale stands for the sale price and quantity sold, and Non-sale stands for the non-sale price and quantity sold. ΔEV shows the Equivalent Variation, which measures how much consumers are willing to pay not to have the merger. Each firm column shows the price in dollars and quantity sold in cases per week per store under different regimes. The column labeled ΔEV is the money consumers are willing to pay not to have the merger in dollars per store per week in each regime.

Table 1.13: Proportion of Loss in Consumer Surplus by Consumer Type (Sale Period)

	No Promotion Adjustment	Promotion Adjustment
Storers	5.04%	10.98%
Non-storers	94.96%	89.02%

Notes: The table presents the proportion of consumer surplus by each consumer type. Demand estimates are from columns 1, 2 and 3 of Table 1.8. Column "No Promotion Adjustment" means that firms do not change the promotion decisions after the merger. Column "Promotion Adjustment" means that firms use asymmetric price discount strategy after the merger.

Table 1.14: Predicted Price Changes

	No Promotion Adjustment	Promotion Adjustment
Change in Sale Price ($)	0.36	3.10
Change in Regular Price ($)	0.80	0.80

Notes: The table presents the change of average simulated prices after a merger. Demand estimates are from columns 1, 2 and 3 of Table 1.8. Column "No Promotion Adjustment" means that firms do not change the promotion decisions after the merger. Column "Promotion Adjustment" means that firms use asymmetric price discount strategy after the merger.

Chapter 2

Empirical Analysis of Algorithmic Retail Pricing

2.1 Introduction

Online retail platforms have significantly facilitated matching between consumers looking for specific products and retailers that offer those products. This feature of such platforms has made it possible for small retailers, that would not have sustained a "brick and mortar" store to operate and sell products to consumers. A much lower consumer search cost relative to the offline setting intensifies competition and requires retailers to engage in complex pricing strategies and quickly respond to price changes and stock-outs of competitors.

While large retailers have teams of experts that develop their competitive pricing strategy and manage inventory, hiring such experts is cost-prohibitive for small online retailers. At the same time, while such retailers are small in terms of the overall sales volume, they still offer a large set of products each which does require active inventory and price management. E.g. About 87% of retailers in the Amazon Marketplace carry at least 100 products (WebRetailer (2014)). Most retailers use automated tools that use algorithms to adjust prices to respond to changes in the inventory and prices of competitors. These tools are typically offered by the third-party firms that specialize in development and support of such tools.

The goal of this paper is to study a retail marketplace where retailers use algorithmic tools for the price adjustment. Our main question is the impact of automation of price adjustment on the efficiency of market outcomes. To answer this question, we will study the structure of the profit functions of online retailers from the observed prices set by the automatic algorithm and using those profit functions we will be able to study the range of possible market outcomes and corresponding revenues and welfare.

Our analysis is based on a large unique dataset of retailers on Amazon's marketplace. Amazon accounts for 44% of all online retail in the United States and, thus, offers a prime market for competition of small online retailers. Most algorithmic tools for automatic price adjustment (a.k.a. "re-pricing tools") offered by third-party firms are tailored for Amazon's marketplace. Consumers can search and browse for specific products on Amazon. The products offered by individual retailers can appear either as "sponsored" products, if a specific retailer buys advertising slots or they appear as "organic" products that best fit the consumer search query. This separation of results into sponsored and organic is similar to those in the traditional search engines such as Google or Bing. A unique feature of Amazon's marketplace is that multiple retailers can offer exactly the same product while a unique product can appear only once in the organic results. To make a decision on whose product is shown, Amazon uses the "buy box" mechanism that selects the retailer to display based on price, star rating, and other features. We describe this mechanism in more detail further.

The "buy box" mechanism has the similar structure to the first-price auction where the winner is selected based on the highest bid. Unlike standard auctions, the retailers on Amazon compete in continuous time and have limited inventory of products. As a result, simply setting the lowest price as in the traditional Cournot competition may not be optimal. In fact, if a retailer believes that her competitor will soon run out of stock, her optimal strategy will be to wait till her competitor stocks out and then sell the product at a higher price.

In this paper, we observe that dynamic re-pricing is performed using very simple algorithmic strategies that track very limited information such as the lowest and, possibly, second lowest price and set the price as a fraction of those prices. We use online learning theory to study such strategies. Online learning theory focuses on the task of minimizing an empirical loss over time as the data comes to the learner one bit at a time. Unlike standard inference in Econometrics, data generation in online learning is allowed to be *adversarial* to explicitly prevent the learner to effectively minimize the empirical loss function. This concept allows the data to be arbitrarily correlated with past actions of the learner and makes the online learning algorithms applicable to the environments with endogeneity such as dynamic games and auctions. The limitation of adversarial settings is that a given online learning strategy cannot be compared to the true optimal strategy because such a strategy would anticipate both the action of the learner and the arriving data and thus can make learner's loss arbitrarily different from that optimum.

To make the solution feasible, the comparison of a learning strategy with the optimal strategy is replaced with the comparison of that learning strategy with a fixed class of strategies. In this case, adversarial data generation affects both the learner's loss function and the loss function corresponding to the selected class of strategies. The success of the learning strategy is measured in the difference between the loss of the learner and the smallest loss of strategies from the selected class which is called the regret of the learner.

The idea of using the notion of regret for empirical study of markets with algorithmic agents has been proposed in Nekipelov et al. (2015). In this paper, we apply that methodology to estimate profit functions of retailers on Amazon's marketplace.

More broadly, our work lies in the general agenda of the analysis of pricing decisions in online retail. This area has received significant attention in the recent years. A group of literature focuses on firms' competition-based dynamic pricing. Fisher et al. (2017) answer this question by measuring price elasticity using a field experiment and estimating the competitors' significance by exploiting information on product stockouts. They find that a best-response pricing algorithm leads to an 11% increase in revenue. Li et al. (2017) propose a methodology with regularization to identify relevant competitors to help hotels price competitively. Bodur et al. (2015) find that online price comparison websites affect consumers' price evaluation and online retailer choice. Another group of literature focuses on constrained-capacity based price optimization in brick-and-mortar and online retailers(Gallego and van Ryzin (1994), Levin et al. (2009), Besbes and Zeevi (2009), Ferreira et al. (2016)) and in various industries[1]. Considering both online and brick-and-mortar retailers, Cavallo (2017) provides a large-scale comparison of online and offline prices by multichannel retailers and finds prices are identical for 72% of the time. Balakrishnan et al. (2014) study the retail-online competition by modeling consumers' browse-and-switch behavior and find it intensifies competition between online and offline retailers.

[1]Fisher et al. (2017) provide a more detailed review on this line of literature.

From the consumer perspective, previous literature investigates on consumers' strategic responses to the dynamic-changing prices. Consumers strategically respond by postponing or accelerating timing of purchase and learning(Gaur and Park (2007), Su (2007), Su and Zhang (2008), Cachon and Swinney (2009), Li et al. (2014), Cachon and Feldman (2015), Moon et al. (2017), Zhang et al. (2017)).

The rest of the paper proceeds as follows. In section 2, we describe the unique data set we use. Section 3 presents the background on Amazon's buy box selection mechanism and the simulation to predict winning sellers. Section 4 presents the model and inference conditions for the rationalizable set. Section 5 provides the empirical analysis on the inference on rationalizable sets, markup ratios, and relative prices across all markets. Section 6 presents the inference on repricing algorithms. Section 7 concludes.

2.2 Data

2.2.1 Background

Amazon.com is the largest e-retailer in the United States with 178 billion US dollars net revenue in 2017 (Statista (2017a)). The platform serves online shoppers who are looking to buy new or used products, especially media and electronics products, and retailers who are selling through the Amazon Marketplace platform to reach hundreds of millions customers. Amazon's business is based on online retail, retail third-party seller services, AWS, subscription service, and other segments. Retail third-party seller services category is the second largest revenue category (32 billion dollars) in annual sales, proceeded by Online stores category (108 billion) and followed by AWS category (17 billion) in 2017 (Statista (2017b)). Third-party retailers sell new or used products online through the e-commerce platform (Amazon Marketplace) to reach Amazon's huge customer base. Around 50 percent of merchandise unit sales in 2017 are attributed to third-party retailers (Statista (2017c)). Business owners reach Amazon's massive customer base through Amazon Marketplace and benefit from Amazon's reputation. On the other hand, Amazon profits from sales through the platform and broadens the variety of products on the platform to benefit customers.

To sell on Amazon as a professional seller (sell more than 40 items per month), third-party sellers pay a fixed monthly subscription fee, a percentage-based referral fee, a fixed-number closing fee, and a shipping fee if the delivery is fulfilled by Amazon. Third-party sellers can choose to fulfill orders by themselves (fulfilled by merchandise, FBM) or by Amazon (FBA). If orders are fulfilled by Amazon, Amazon is in charge of picking, packing, and shipping the products and takes care of customer service and returns. For Amazon Prime members, FBA provides shoppers with a free 2-day shipping and fast and reliable order fulfillment.

Since a unique product can only appear once in the organic search result, third-party sellers compete for being displayed (a.k.a winning the "Buy Box"). When a buyer selects "Add to Cart" on the product details page, the seller who has the "Buy Box" at the moment wins the sale. Among all Amazon sales, 82% of sales go through the Buy Box, and the percentage increases for mobile users (RepricerExpress (2018), Lanxner (2018)). Figure 1 displays a sample web page. The upper red box marks the Buy Box option. In this example, the third-party seller named "Global iq" wins the Buy Box and the order is fulfilled by Amazon. The lower red box links to the list of other sellers selling this product.

To become eligible for winning the Buy Box, the third-party sellers can choose fulfillment by Amazon (FBA), which will automatically enable them to win the Buy Box, or have a professional account, a good performance history (Order Defect Rate, Cancellation Rate, and Late Shipment Rate), and sufficient order volume. To win the Buy Box, there are several variables that have the highest effect: competitive pricing, using Fulfillment By Amazon(FBA), customer service, and available stock. For instance, Figure 2 shows the list of other sellers selling the video game "Star Wars Battlefront II". The seller "Global iq" who wins the Buy Box ranks in the fifth position on the list. The seller "Global iq" provides a competitive price at $29.98 with FBA shipping and 97% positive rating. In addition, Chen et al. (2016a) reveal that Amazon is much more likely to choose sellers who use algorithmic pricing in the Buy Box. These algorithmic sellers contribute to a third of third-party sales.

Though algorithmic pricing sounds complicated for third-party sellers, many companies such as Reprice Express(figure:2.2), RepriceIt, and Feedvisor(figure:2.3) provide automated pricing services, which allows sellers to choose from a menu of pricing-strategy options. These options include: find the lowest price offered and go above it (or below it) by X dollars or Y percentage, find Amazon's price and adjust up or down based on it, and other options. The automated pricing services will follow the selected strategy and keep competitive pricing to increase the chance of winning the Buy Box. Moreover, Amazon provides a free algorithmic pricing tool named "Amazon Selling Coach" for sellers with professional accounts. Amazon selling coach notifies the seller on lower prices and suggests to match low prices for single or multiple listings based on listing condition, fulfillment method, feedback rating, and handling time.

2.2.2 Data Description

From May 2016 to May 2017 and from Dec 2017 to March 2018, we collected product price, shipping, seller ratings, and other sellers list for 95 products in the video game category. It is ideal to study video games category due to its popularity on Amazon and its competitive environment. Usually, a large number of third-party sellers and Amazon participate in the competition for the Buy Box of video games, which might include many algorithmic pricing sellers.

We define a market as a combination of video game and video game console. For instance, a video game named "Star Wars Battlefront II" provides both PlayStation 4 and Xbox One versions, which constructs two markets: "Star Wars Battlefront II - PlayStation 4" and "Star Wars Battlefront II - Xbox One". Since video game consoles provide exclusive usage for video games, it is reasonable to define two separate markets even for one game. Within each market, multiple third-party sellers and Amazon compete for the Buy Box position for this game.

In order to track the dynamic-evolving competition environment, we collect product information on a daily base. From May 2016 to May 2017, we collected the list of sellers from the other sellers' link on the product details page. Figure 2.4 displays a sample list of other sellers on Amazon. To reduce the possible price differences due to conditions ("used" or "used like new"), we focus on the listings of products in the new condition. In the table shown in Figure 2.4, sellers were ranked in the ascending order of the sum of price and shipping. We collected price and shipping information from the first column, condition information from the second column, delivery method from the third column, and seller ratings from the fourth column. The table displays ten sellers on a page, and the length of tables varies by products.

From Dec 2017 to March 2018, besides the list of sellers collected, we also collect the information on the sellers winning the Buy Box. For instance, Figure 2.1 shows a seller named “Global iq” winning the Buy Box for product “Star Wars Battlefront II - PlayStation 4”.

Our data comes from 95 popular video games from Amazon. Figure 2.5 and Figure 2.6 show the price trends for two sample video games: “Fallout 4 - PlayStation 4” and “Mortal Kombat XL-Xbox One”. The prices in these figures are not stable across time. Instead, third-party sellers adjust their prices to match with their close competitors. The platform seller “Amazon” does not provide the best price across all time periods, and third-party sellers “media&more” and “Games 4 Life” beat Amazon in certain periods.

The repricing tools perform a simple matching with the close competitors' prices and prices at certain ranks to stay competitive. For instance, on each price list webpage as in figure 2.4, there are ten sellers listed per page. We choose the top price and the bottom price on each webpage as two benchmarks and show how sellers respond to the price sequences at certain ranks. Figure 2.7 displays the price sequences of a third-party seller “media&more”, top1 seller, and top 10 seller. The gray areas mark two areas of our interest. In the first grey area, the third-party seller's price is close to the lowest price; then the seller increases its price. In the second grey area, the third-party seller's price is far above the top 10 price; then the seller adjusts by decreasing its price. Figure 2.8 presents the seller's standardized price and rank trajectories and marks the two corresponding areas. From its own price and rank history, there is no clear linear relationship between price adjustments and rank changes.

One limitation of our data is that we lack the Buy Box winner information from May 2016 to May 2017. We use data with winner information from Dec 2017 to March

2018 and simulate the winning seller for the period from May 2016 to May 2017. We only include top 20 retailers for consideration. To simulate the winning sellers, we use observed information on price, shipping, and ratings to construct a number of variables to characterize the competition environment that third-party sellers face.

Table 2.1 presents the summary statistics in both data collection periods. The table compares the statistics for data without winners information from May 2016 to May 2017 and data with winning information from Dec 2017 to March 2018. To compare prices for different products, we re-scale price based on the lowest price available within a certain market. Besides observed variables FBA, rate, and Amazon, we construct a list of variables to characterize the competition environment. The variable "FBA_price" is the lowest price provided by an FBA seller within a market. The variable "FBA_rate" is the rate of the lowest priced FBA seller. These two variables characterize the competition pressure from the lowest-priced FBA seller. In addition, we use the variables "FBA_price_hr" and "FBA_rate_hr" to describe the price and rating of the highest-rating seller. Besides FBA sellers, the variable "top1_rate" measures the rating of the lowest-priced seller. To capture the effect of Amazon, the variable "Amazon_price" is the price provided by Amazon. If Amazon is not in the market, we use a dummy variable "No_amazon" to capture these cases. The table indicates that two periods have similar values of ratings and prices for multiple groups of sellers. However, the proportions of Amazon participating in the market differ in these two periods. Because Amazon has perfect metrics in all dimensions, Amazon has a dominant chance of winning the Buy Box. The low proportion of Amazon participating will not affect the simulation of winning probability.

2.3 The Buy Box Mechanism

2.3.1 Mechanism

Amazon's Buy Box selection mechanism is a black box for third-party sellers. However, Amazon and multiple repricing tools provide a list of approaches to win the Buy Box as follows:

- price items competitively
- offer prime and free shipping
- provide great customer service
- keep stock available

We discuss three components in detail and how we simulate the probability of winning Amazon's Buy Box.

To win the buy box, third-party sellers need to price products competitively. Amazon selling coach provides a feature named "Match Low Price" to stay competitive for single or multiple listings. The feature matches the landed prices (sum of price and shipping) of items based on listing condition, fulfillment method, feedback rating, and handling time. Third-party sellers can choose to only match listings in similar conditions or any listings based on their preferences. For instance, a third-party seller selling a new product can choose only to consider listings selling new products. A seller using FBA might only consider sellers using the same shipping method. To capture this effect, we rescale sellers' listing prices based on the lowest landed price available in the market and construct a variable named "FBA_price" as the lowest landed price provided by an FBA seller. Similarly, a seller with 97% positive ratings can only match listings from sellers with similar or better ratings. We construct a variable named "FBA_price_hr" as the price of the FBA seller with the highest rating.

The algorithm also uses shipping methods for Buy Box winner selection. Since Amazon considers its fulfillment service with perfect metrics, sellers using Fulfillment By Amazon (FBA) increase their probabilities of winning the Buy Box. Sellers using Fulfillment By Merchant (FBM) require high scores across all metrics and very low prices. We use a dummy variable "FBA" to indicate a seller using FBA service to capture the privilege of FBA shipping.

In addition to prices and shipping methods, the Buy Box algorithm also uses ratings as key inputs to select a winner. Figure 2.4 displays a table of listings with seller ratings. Seller ratings are out of 100 percent. For new sellers, Amazon marks them as "Just Launched". We construct a variable named "rate" from 0 to 100 and treat new sellers with rating 0 and Amazon with rating 100. To control for the competition from FBA sellers, we include the rating of lowest-priced FBA seller and the highest rating of FBA seller into the simulation. In addition, we add landed price and rating variables to control the competition pressure from non-FBA sellers.

2.3.2 Simulation

We use data with labeled winning sellers from Dec 2017 to March 2018 to train a model to simulate and predict the probabilities of winning Buy Box. The probability of winning Buy Box depends on seller i's price, shipping method, rating, and competition environment characterized by other sellers' lowest price, the highest rating, and other characteristics. We use a Support Vector Machine Approach to classify winning sellers and predict the probabilities of winning the buy box for all sellers in the market.

One potential concern for this data set is the imbalanced class problem. Imbalance class problem occurs when the class of interest has much fewer samples than the other classes. When the interested class is insufficiently represented, samples in this class are more likely to group into undiscovered idiosyncratic noise. However, the smaller class usually are the main interest of studies, and it is important for a classification model to have a good prediction rate on the smaller class. In our simulation, we are interested in simulating the winning sellers through a list of active sellers. The labeled training dataset consists of 5% labeled winning sellers while 95% of sellers are labeled in the unselected group not winning the Buy Box. The group of sellers winning the Buy Box is a minority class, while unselected sellers belong to a majority class. Imbalanced class data leads the classifier to bias towards the majority class and have a poor classification rate on the minority class. In our case, it leads to an underestimation of the probability of winning the Buy Box for each seller in the market.

Researchers have proposed various techniques to solve this problem (Longadge et al. (2013), Ali et al. (2015)). These techniques can be divided into two categories: data-level approach, and algorithmic-level approach. The techniques under the data-level approach undersample or oversample a certain class to reach a balanced distribution of classes.

For data-level approach, sampling is applied either to add new samples (over-sampling) or to remove existing samples (under-sampling). The over-sampling approach randomly over-samples the minority class to reach a balanced class distribution. By duplicating the minority class samples, it reuses the minority class data and exaggerates the information of the minority class. On other direction, the under-sampling approach randomly under-samples the majority class to balance the distribution. By removing majority class samples, it reduces the information of the majority class and may lose valuable information. We employ the oversampling approach to balance classes in the data-preprocessing step. For algorithmic approaches, cost-sensitive learning classification approaches adjust the costs for samples in the minority class and minimize the cost of misclassification.

The linear SVM classifier model classifies a new instance by computing the decision function $wX+b$, where X contains information on seller i and its competitors' prices, ratings, and shipping methods $(p_i, \boldsymbol{p_{-i}}, rate_i, \boldsymbol{rate_{-i}}, FBA_i, \boldsymbol{FBA_{-i}})$. If the result of decision function is positive, the predicted result is the positive class, otherwise, it is the negative class.

$$\hat{y} = \begin{cases} 0, & \text{if } wX + b < 0, \\ 1, & \text{if } wX + b \geq 0 \end{cases}$$

We train a linear SVM model to find the values of w and b which maximizes the decision margin and limits violations. The decision margins are the points which are parallel and at equal distance to the decision boundary. Formally, we define $t_i = -1$ for the negative class with $y_i = 0$ and $t_i = 1$ for the positive class with $y_i = 1$. The objective function of a soft margin linear SVM classifier is as follows:

$$\begin{aligned} &\underset{w,b,\zeta}{\text{minimize}} \quad \frac{1}{2}\left\|w\right\|^2 + C\sum_{i=1}^{m}\zeta_i \\ &\text{subject to} \quad t_i\left(wx_i + b\right) \geq 1 - \zeta_i, \quad \zeta_i \geq 0, \quad i = 1, \ldots, m. \end{aligned}$$

where ζ_i is a slack variable which measures how much instance i is allowed to violate the margin, and C is a hyperparameter which regularizes the weight on the loss of prediction violations. There is trade-off between maximizing margins and reducing the number of violations. When we increase the margin width, the number of violations also increases. The hyperparamter C helps control the trade-off between increasing margins and punishing violations. A smaller C means a model with wide margins and more violations, and a larger C means a model with narrow margins and less violations.

To find the optimal penalty factor C, we perform a grid search and evaluate the models based on 10-fold cross-validation. we first grid search possible C values over $\{0.001, 0.01, 0.1, 1, 10, 100\}$. Figure 2.9 reports the average F1 scores under different C values. The average F1 score increases marginally after the penalty factor reaches 1 (power of 10 at 0). When we increase the penalty factor from 1 to 100, the increases of the F1 score are marginal. Thus, we choose the penalty factor as 1 for the linear SVM model.

2.4 Model

2.4.1 Setup

We assume that the marketplace induces competition between N retailers offering the same product. In each period, the retailers can submit their prices, and the marketplace determines the retailer who is placed in the buy box via a mechanism which we denote G. We assume that the retailer who is placed in the buy box sells 1 unit of the product while retailers who are not placed in the buy box do not sell any units of the product. Each player is an algorithmic pricing algorithm. Each player i has a strategy space P_i. The utility of player i depends on its own strategy p_i, price strategies of other players $\boldsymbol{p_{-i}}$, and private cost information c_i only observable by seller itself. We assume that the price strategy space is the bounded segment $\mathcal{P}$.

We assume that price competition and mechanism G gets repeated in the sequence of discrete periods. In each period t, player i chooses a landed price p_{it}. We denote the sequence of price strategy profile as $\{\boldsymbol{p}\}_t$. Player i has a private information on unit cost c_i. We assume c_i is fixed through the sequence. In practice, FBA sellers usually send a batch of products to the Amazon warehouse in advance, and the unit cost is fixed once shipped to Amazon[2].

Mechanism G is as follows: Each player i is associated with a shipping method FBA_{it} and a rating $rate_i$ and chooses a landed price per unit p_{it}. We assume shipping method FBA_i is fixed for player i during the observation period because an algorithmic pricing tool does not equip the ability to alter shipping method. Amazon Buy Box algorithm considers the price strategy, shipping method, and rating of all players and select a winning seller for the Buy Box. Player i's probability of winning the Buy Box $Pr_i(\boldsymbol{p})$ is a monotone decreasing function of its own price p_{it} and competitors' ratings $rate_{-i}$ and a monotone increasing function of other players' prices p_{-it} and its own rating $rate_i$. We denote the probability of winning the Buy Box for seller i is as follows:

$$Pr(\boldsymbol{p}) = f(p_i, \boldsymbol{p_{-i}}, rate_i, \boldsymbol{rate_{-i}}, FBA_i, \boldsymbol{FBA_{-i}})$$

Figure 2.10 presents a sample Pr function. As the price increases, the probability of winning the "Buy Box" does not decrease continuously. The probability drops sharply around the price ranges associated with rank changes. In addition, Figure

[2]Note that the sequence of price strategy $\{\boldsymbol{p}\}_t$ is observable in the dataset.

2.10 displays a sample probability function for a non-FBA seller “media&more”, which has a relatively low probability of winning the “Buy Box” compared with FBA sellers and Amazon.

By $\pi_{it}(\boldsymbol{p}; c_i)$ we denote the profit function of player i. The profit at time t is defined as:

$$\pi_{it}(\boldsymbol{p}; c_i) = (p_{it} - c_i)Pr(\boldsymbol{p})$$

where a set of parameters $(\boldsymbol{p}, \boldsymbol{FBA}, \boldsymbol{rate})$ are observable in the data while player i's cost c_i is not observed. We denote the expected revenue of selling at price p_{it} as follows:

$$R_{it}(\boldsymbol{p}) = p_{it}Pr(\boldsymbol{p})$$

We rewrite the profit of each stage t is as follows:

$$\pi_{it}(\boldsymbol{p}; c_i) = R_{it}(\boldsymbol{p}) - c_i Pr(\boldsymbol{p})$$

The private information parameter c_i is player i's private unit cost.

2.4.2 No-regret Learning

We use the online learning theory to analyze the game of repeated rounds. Online learning is the process of optimizing given history of choices and answers in the previous rounds and possible additional available information. Online learning occurs in a sequence of successive rounds. In each round, the learner receives an instance and is required to make a choice. After making a choice, the correct answer is revealed at the end of each round. The learner suffers a loss defined by the difference of the value obtained by her choice, and the value would have been achieved by the correct answer. The learner's ultimate goal is to minimize the cumulative loss suffered along with her history. It means that the learner tries to infer using information from previous rounds and make few mistakes in the present and future rounds.

To learn from past rounds, there must exist some correlations between past and present. Unlike standard inference in Econometrics, data generation in online learning is allowed to be deterministic, stochastic, or even *adversarial* to explicitly prevent the learner to effectively learn from the experience. This concept allows the data to be arbitrarily correlated with past actions of the learner and makes the online learning algorithms applicable to the environments with endogeneity such as dynamic games and auctions. The limitation of adversarial settings is that a given online learning strategy cannot be compared to the true optimal strategy because such a strategy would anticipate both the action of the learner and the arriving data and thus can make learner's loss arbitrarily different from that optimum.

To make the solution feasible, the comparison of a learning strategy with the optimal strategy is replaced with the comparison of that learning strategy with a fixed class of strategies called a hypothesis class, which is known to the learner. When we assume that all data are generated by some target strategies from the hypothesis class, we call it *the realizable case*. This is analogous to the batch learning model. Alternatively, we can relax the realizable assumption by not assuming that all answers are generated by some target strategies from the hypothesis class. Instead, we require the learner to be compared with the best-fixed strategy from the hypothesis class. The difference between the value of the learner's strategies and the value would have been achieved by the best-fixed strategy is captured by *regret*. The *regret* measures how sorry the learner is in retrospect. In a dynamic and complex environment, the online learning algorithm is ideal to analyze agents using algorithmic tools. We will formally define the concept of regret further.

In the Amazon marketplace, we consider repeated discrete applications of mechanism G. We refer the sequence of price strategy profile $\{\boldsymbol{p}\}_t$ as the *sequence of play*. To infer players' private values in cost, we make a rationality assumption on how players choose price strategies. In this work, we employ the online learning algorithm and use the weaker *no-regret* assumption to model players' strategic behavior.

For the sequence of play $\{\boldsymbol{p}\}_t$, player i's total utility is $\sum_t \pi_{it}(\boldsymbol{p_t}; c_i) = \sum_t \left[R_i(\boldsymbol{p}_t) - c_i Pr(\boldsymbol{p}_t)\right]$. For the best fixed price strategy p', player i's cumulative proft is $\sum_t \pi_{it}(p', \boldsymbol{p_{-it}}; c_i) = \sum_t \left[R_i(p', \boldsymbol{p_{-it}}) - c_i Pr(p', \boldsymbol{p_{-it}})\right]$. We define the average difference between using the sequence of price strategy $\{\boldsymbol{p}\}_t$ and the cumultive profit-maximizing fixed price p' as the *regret*:

$$regret_i(\boldsymbol{p_i}) = \frac{1}{T}\left(\max_{p' \in \mathcal{P}} \sum_t \pi_{it}(p', \boldsymbol{p_{-it}}; c_i) - \sum_t \pi_{it}(\boldsymbol{p_t}; c_i)\right)$$

where *regret* is measured by comparing the cumulative utility of the best fixed price p' and the cumulative utility of a price sequence $\boldsymbol{p}_i$. The *no-regret* assumption assumes the cumulative utility using the sequence of price strategy is at least as high as the cumulative utility using a best fixed price strategy, given other players using the sequence of price strategies at each stage. This assumption is named as *no-regret assumption* and will be formalized in next section.

The *no-regret* assumption is weaker than the best-response assumption of the classical Nash equilibrium. In a dynamic and complex competition environment, players are learning how to play over time. Players might benefit from frequent adjusting prices to stay competitive and increase profit, which might lead to a negative regret. A negative regret still satisfies the *no-regret assumption*.

2.4.3 Inference

To infer the player's private information on cost, we assume the cost is fixed over the period and use no-regret learning as the behavioral model. Compared with the classical Nash equilibrium model, when a player changes its price, we can infer a different private cost c_i for player i. Instead of assuming players changing cost over time, it is more reasonable to use the no-regret learning model in a dynamic-evolving competition environment, especially for players who change price frequently. We employ the methodology of Nekipelov et al. (2015) to infer private values through *rationalizable set*.

Definition 2 (Average Regret)**.** *A sequence of play that we observe has a ϵ_i-regret for player i if:*

$$\forall p' \in \mathcal{P} : \frac{1}{T}\sum_{t=1}^{T} \pi_{it}(\boldsymbol{p_t}; c_i) \geq \frac{1}{T}\sum_{t=1}^{T} \pi_{it}(p', \boldsymbol{p_{-it}}; c_i) - \epsilon_i \tag{2.1}$$

This definition leads to the following definition of a *rationalizable set under no-regret learning*.

Definition 3 (Rationalizable Set)**.** *A pair of (ϵ_i, c_i) of a cost c_i and error ϵ_i is a rationablizable pair of player i if it satisfies equation (2.1). We refer to the set of such pairs as the rationalizable set and denote it with NR_i.*

In this model, we assume that the strategies $\{\boldsymbol{p}\}_t$ are chosen simultaneously, and player i is not able to adjust price strategy p_{it} based on the state of nature and competitors' price strategies p_{-it}. The rationality assumption models players who may adjust price strategies when participating the game.

For the model of our main interest, when player i's rationalizable set is (ϵ_i, c_i), equation (2.1) can be organized as:

$$\forall p' \in \mathcal{P} : \frac{1}{T}\sum_{t=1}^{T}\big(R_i(p', \boldsymbol{p}_{-it}) - R_i(\boldsymbol{p}_t)\big) \leq c_i \cdot \frac{1}{T}\sum_{t=1}^{T}\big(Pr(p', \boldsymbol{p}_{-it}) - Pr(\boldsymbol{p}_t)\big) + \epsilon_i$$

If we denote the increase in the average revenue for player i switching to a fixed price p' as

$$\Delta R_i(p') = \frac{1}{T}\sum_{t=1}^{T}\big(R_i(p', \boldsymbol{p}_{-it}) - R_i(\boldsymbol{p}_t)\big)$$

and denote the increase in the average probability of buying from player i switching to a fixed price p' as

$$\Delta Pr_i(p') = \frac{1}{T}\sum_{t=1}^{T}\big(Pr(p', \boldsymbol{p}_{-it}) - Pr(\boldsymbol{p}_t)\big)$$

We simplify the condition as:

$$\forall p' \in \mathcal{P} : \Delta R_i(p') \leq c_i \cdot \Delta Pr_i(p') + \epsilon_i \tag{2.2}$$

The rationalizable set NR is an envelope of the family of half planes obtained by varying $p' \in \mathcal{P}$ in equation (2.2). We explain how the setting in our model satisfies the required assumptions in Nekipelov et al. (2015) to infer private costs.

The rationalizable set NR derived from linear constraints with equality is a closed convex set. Since NR is a closed convex set, we can use functions $R_i(\cdot,\cdot)$ and $Pr_i(\cdot,\cdot)$ to define the set of support hyperplanes which stands for the rationalizable set NR. Since it is easy to have an upper bound on the maximum price a player can price, the support of prices is a compact set $\mathcal{P} = [0, \bar{p}]$. The expected revenue and the probability of purchase are decreasing functions of the player's price. The functions and linear combination of functions $R_i(\cdot,\cdot)$ and $Pr_i(\cdot,\cdot)$ are monotone and bounded on $\mathcal{P}$.

We note that the rationalizable set NR is not bounding for any error. As the error increases, the range of values which are compatible with rational behavior increases. However, for any bounded error, the rationalizable set NR is bounded. Since we are interested in the properties of rationalizable set NR and the set NR is a closed convex bounded set, we can fully characterize the set NR by its boundaries, we follow Nekipelov et al. (2015) to use the notion of *support* function to represent the boundary of the set NR. Recall the definition of the *support* function as follows:

Definition 4. *The support function of a closed convex set X is defined as:*

$$h(X, u) = \sup_{x \in X} \langle x, u \rangle$$

where the set $X = NR$ is a subset of $\mathbb{R}^2$ or cost and error pairs (c_i, ϵ_i), and u is also an element of $\mathbb{R}^2$.

An important property of the support function for closed convex bounded sets is characterizing the support function of NR is equivalent to characterizing NR itself. Recall the Hausdorf norm for subsets A and B of the metric space E with metric

$\rho(\cdot,\cdot)$ is defined as:

$$d_H(A,B) = \max\{\sup_{a\in A}\inf_{b\in B}\rho(a,b), \sup_{b\in B}\inf_{a\in A}\rho(a,b)\}$$

It turns out that $d_H(A,B) = \sup_u |h(A,u) - h(B,u)|$. Based on the functions $\Delta R(p)$ and $\Delta Pr(p)$, we characterize the rationalizable set NR by the theorem below:

Theorem 2.4.1. *Under the monotonicity of $\Delta R(\cdot)$, the support function of NR is*

$$h(NR,u) = \begin{cases} |u_2| \inf\limits_{u_1 - \Delta Pr(p)u_2 > 0} -\Delta R(p), & \textit{if } u_2 < 0 \\ & \textit{and } \frac{u_1}{|u_2|} \in [\inf_p -\Delta Pr(p), \sup_p -\Delta Pr(p)] \\ +\infty, & \textit{otherwise} \end{cases}$$

Proof. While the previous work Nekipelov et al. (2015) provides the methodology for support function, they rely on the continuity assumption of $\Delta R(\cdot)$ and $\Delta Pr(\cdot)$. In our case, both functions are discontinuous. As we decrease the prices, the probabilities of being selected in the Buy Box will increase. The probability of winning the Buy Box jumps when seller i beats any of sellers ranked before seller i. We follow the methodology used in Jalaly et al. (2017) which only requires the monotonicity assumption. Provided that the support function is positive homogenous, without loss of generality we can set $u = (u_1, u_2)$ with $\|u\| = 1$. To find the support function, we take u_1 to be dual to c_i and u_2 to be dual to ϵ_i. We re-write the inequality of the half-plane as: $c_i \cdot -\Delta Pr(p_i) - \epsilon_i \leq -\Delta R(p_i)$. We need to evaluate the inner-product

$$u_1 c_i + u_2 \epsilon_i$$

For $u_2 \geq 0$, $h(NR, u) = +\infty$. For any $u_2 < 0$, we note that the inequality for the half-plane can be re-written as: $c_i(-\Delta Pr(p_i)|u_2|) + u_2\epsilon_i \leq -\Delta R(p_i)|u_2|$.

For each $\boldsymbol{u} = (u_1, u_2)$, we find $p^u_{\boldsymbol{u}}$ and $p^l_{\boldsymbol{u}}$ such that $-\Delta Pr(p^l_{\boldsymbol{u}}) < -\frac{u_1}{u_2} < -\Delta Pr(p^u_{\boldsymbol{u}})$. The function $\Delta Pr(\cdot)$ is constant in the interval $[p^l_{\boldsymbol{u}}, p^u_{\boldsymbol{u}})$. All vectors $\boldsymbol{u}$ that satisfies the conditions above correspond to the hyperplanes across the vertex of the rationalizable set. Thus, the support function will be $-|u_2|\Delta R(p_{\boldsymbol{u}^l})$.

Now suppose that $\inf_p \Delta Pr(p) > 0$ and $u_1 > -|u_2| \inf_p \Delta Pr(p)$. We can re-write as follows:

$$u_1 c_i + u_2\epsilon_i = \left(u_1 + \inf_p \Delta Pr(p)|u_2|\right) c_i + \left(-\inf_p \Delta Pr(p)\right) |u_2| c_i + u_2\epsilon_i$$

Function $\left(-\inf_p \Delta Pr(p)\right)|u_2|c_i + u_2\epsilon_i$ is bounded by $|u_2|\left(-inf_p \Delta R(p_i)\right)$ for each pair $(c_i, \epsilon_i) \in NR$. At the same time, function $(u_1 + \inf_p \Delta Pr(p_i)|u_2|)c_i$ is strictly increasing in c_i. As a result, the support function at any vector u with $u_1/|u_2| > -\inf_p \Delta Pr(p)$ is $h(NR, u) = +\infty$. $\square$

2.5 Inference on Rationalizable Sets

We follow Nekipelov et al. (2015)'s approach to infer the costs and regrets for a set of players in the Amazon Marketplace. We focus on players who frequently change their prices thus are more likely to use algorithmic pricing tools. Each player corresponds to a price strategy for a listing in the market. The Amazon Marketplace is dynamic-evolving and highly competitive where players even change prices multiple times in a day. Within a certain time frame, price changes result from the changes in a competitive environment.

We construct the support function to fully characterize the rationalizable set NR which is the *identified set* of parameters for each player: player's cost per unit and the bound on the average regret. It is a one-dimensional function which can be estimated from data simulation.

To compute the empirical rationalizable set, we assume that prices and costs have an upper bound, thus the strategy space of each player is a finite set. For each price p' in the price strategy space, we compute the $\Delta R(p')$ and $\Delta Pr(p')$. Next, we discrete the space of regret error ϵ_i. For each error ϵ_i we use the equation 2.2 to compute the upper and lower bound on the cost for each discrete error term. We find the smallest error ϵ_0 where the lower bound is smaller than the upper bound. The smallest error ϵ_0 is the smallest rationalizable error.

We show examples of rationalizable set through Figure 2.11 and Figure 2.12. Figure 2.11 and Figure 2.12 present two long-history and high-frequency price changing sellers. The vertical axis is the cost values for players, and the horizontal axis is the average regret. The shape of rationalizable set NR has a convex shape. Both Figure 2.11 and Figure 2.12 have very small addictive regrets, which indicates the sellers in Figure 2.11 and Figure 2.12 adjust their prices efficiently. On the identification of rationalizable set, Figure 2.12 provides a narrow set of cost and regret pairs than Figure 2.11. We note a difference in the shape of the rationalizable set from the rationalizable set for bidders in advertising auctions Nekipelov et al. (2015). The shape of rationalizable sets is not a smooth convex shape but with some kinks.

At each average regret level, the lower and upper bounds correspond to the minimum and maximum of alternative prices in the finite price space. We derive the lower and upper bounds through equation 2.2. We discuss the details and economic meanings of these bounds. We group and distinguish cases based on the sign of

$\Delta Pr(p')$.

- When $\Delta Pr(p')$ is positive, the players can increase the chance of being selected by switching to a fixed price. If switching to a fixed price does not increase the revenue, this implies extra cost completely offsets the extra revenue and the lower bound for the cost. The player will not lower the price when the extra revenue cannot compensate the extra cost. Formally, when $\Delta Pr(p') > 0$, equation 2.2 implies $c \geq \max\limits_{p':\Delta Pr(p')>0} \frac{\Delta R(p')+\epsilon}{\Delta Pr(p')}$.

- When $\Delta Pr(p')$ is negative, the players can decrease the chance of being selected by switching to a fixed price. If switching to a fixed price does not increase the revenue, this implies extra cost completely offsets the extra revenue and the upper bound for the cost. The player will not increase the price when the additional revenue is not larger than the additional cost. Formally, when $\Delta Pr(p') < 0$, equation 2.2 implies $c \leq \min\limits_{p':\Delta Pr(p')<0} \frac{\Delta R(p')+\epsilon}{\Delta Pr(p')}$.

- Cost should not be negative or exceed the upper bound of the price. In our case, players are assumed to have positive costs for the merchandise. Whenever products are priced below the cost, players receive a negative revenue. Players will choose not to serve or exit the market. Formally, $c \in [0, \bar{p}]$.

We compute the smallest addictive regret across sellers and markets. We note that some sellers have negative regret, and we group negative regret and zero regret together. About 33% sellers have negative or zero regrets. As Figure 2.13 shows, when we categorize very small regret ($\epsilon < 0.001$) into the zero and negative regret group, there are about 61% of listings have negative, zero or almost zero regret. The rest 39% of listings have positive but small regrets. When we categorize small regret

($\epsilon < 0.01$) into the zero and negative regret group, about 93% of sellers have negative, zero or almost zero regrets. It indicates that a majority of players have reached a stable state and competed efficiently, while about 7% of players are converging.

We also report the computed markup ratios grouped by length of periods in the market, rating, and shipping method. Table 2.3 reports the quantiles of markup ratios by different seller groups. We use the length of periods a seller is in the market as criteria and group sellers into five categories. From category 1 to category 5, seller's life length (time periods staying in the market) increases. Category 5 sellers stay in the market for the longest time, while category 1 sellers participate in the market for the shortest time. Category 5 sellers have the highest markup ratios in all quantiles, while category 1 sellers have the lowest markup ratios in all quantiles.

Table 2.4 reports the quantiles of markup ratios by different rating groups. We group sellers into 4 rating categories, where a higher category stands for a higher rating. Category 1 sellers include sellers with 0 rating and sellers just launched on the platform. Category 2 sellers have ratings from 1 to 80, which captures the sellers with a medium rating. Category 3 sellers have ratings from 80 to 90, and category 4 sellers have ratings from 90 to 100. Category 4 sellers have the highest markup ratios. Sellers with excellent ratings can have a medium markup ratio, while new entrants have relatively higher markup ratios.

Table 2.5 reports the quantiles of markup ratios by different shipping methods. The markup ratios of sellers with/without FBA shipping are very similar. The markup ratios for sellers using FBA shipping are slightly lower than sellers without FBA shipping.

Table 2.2 reports the representative markets with minimum, median, and maximum median markup ratios in each segment. Across the markets of Xbox one video games, the upper panel presents the minimum (market 72), median (market 4), and maximum (market 71) markup ratios. The 50% markup ratios vary from 6% to 29%. Across the markets of PlayStation 4 video games, the lower panel displays the minimum (market 67), median (market 15), and maximum (market 60) of median markup ratio quantiles. The 50% markup ratios range from 4% to 26%.

We report the quantiles of markup ratio by market in the Table 2.7 and Table 2.8 (in Appendix). The markup ratios vary across markets. The 25% markup ratios vary from 5% to 20%, and the 75% markup ratios vary from 9% to 68%. Table 2.9 and Table 2.10 present the quantiles of relative revenue ratios by market.

2.6 Inference on Repricing Algorithm

The repricing tools use simple matching algorithms to stay competitive in the market. Repricing tools can track prices in top ranks and prices from close competitors. We use the Granger Causality test to detect how players respond to price sequences of certain ranks and prices from competitors. For any two retailers, A and B, if retailer A's price sequence can be used to forecast retailer B's price, we define it as retailer B responds to retailer A's price changes. If retailer B's price sequence can also be used to forecast retailer A's price, retailer A and retailer B forms a cycle of price adjustments.

Figure 2.14 presents the sample price trajectories for selected active sellers and price sequences in rank 1 and 2. Seller A, B, and C are sellers which adjusted their prices at least 10 percent of their histories and appeared at least once in the top 10 positions. Sellers R1 and R2 are the price sequences at rank 1 and rank 2 in this market. We run a pairwise Granger Causality test between possible pairs selected from the list of sellers and visualize the Granger Causality test results in Figure 2.15. In Figure 2.15, seller pairs are connected by arrowed lines. An arrowed line from A to B means that the price sequence of seller A can be used to forecast seller B's price sequence, in other words, seller B responds to seller A's price changes. In Figure 2.15, we find that both seller A and seller B respond to price changes in the top rank. Besides rank 1 price sequence, seller B also responds to price changes in rank 2 and price changes from seller A and seller C.

To analyze the matching algorithm used by sellers, we focus on sellers' responses to the lowest price, the second lowest price, and price in rank 10 (the bottom price on the first page). To accomplish this goal, we proceed in several steps: first, we select some active sellers. Active sellers were selected by two criteria, whether sellers frequently change the price and whether sellers ever stayed on the first page (top 10 ranks). The criteria one is defined as sellers adjusting their prices for at least 10 percent of their price histories. This criterion helps us pick sellers with frequent price changes. The criteria two is defined as whether sellers ever stayed on the first page. Because we are interested in sellers who actively compete for the "Buy Box", we set the second rule to focus on these active sellers.

We fill in the missing values of active sellers with the highest price in the market. The assumption is that sellers leave the market because they can not even provide the highest available price in the market. To test whether active sellers respond to the minimum price change, we construct a price sequence with the lowest price in the market and test whether minimum price sequence can be used to forecast active sellers' price sequence. The Granger Causality test is only performed in one direction. We do not include market 1, 8, and 60 due to the extremely short price histories.

Table 2.6 reports the distribution of sellers' responses to price sequences in certain ranks. Among sellers who respond to price sequences in certain ranks, about 40% of sellers adjust prices based on price sequences in rank 1, rank 2, and rank 10. About 50% of sellers respond to price changes in rank 1 and 2. About 85% of responding sellers adjust their prices based on price sequences in rank 1 or 2 or both. About 69% of responding sellers consider the price sequence in rank 10 for price adjustments. Table 2.11, 2.12, 2.13, 2.14, 2.15, and 2.16 in the Appendix provide detailed Granger Causality test results by market.

2.7 Conclusion

In this paper, we study the impact of automation of retail price adjustment on the efficiency of market outcomes and seller markups using the evidence from the Amazon marketplace. We use the online learning framework from the Computer Science literature to recover the link between observed automated price changes and the marginal costs of algorithmic retailers. Instead of the best-response assumption, we use a weaker no-regret assumption to model agents' dynamic pricing behavior, where regret measures the relative performance of the dynamic strategy compared with a best-fixed benchmark strategy in retrospect. We find that the expected regrets for most of the retailers are close to zero and the estimated markups of retailers are compatible with traditional retail markets.

2.8 Appendix

2.8.1 Figures

2.8.2 Tables

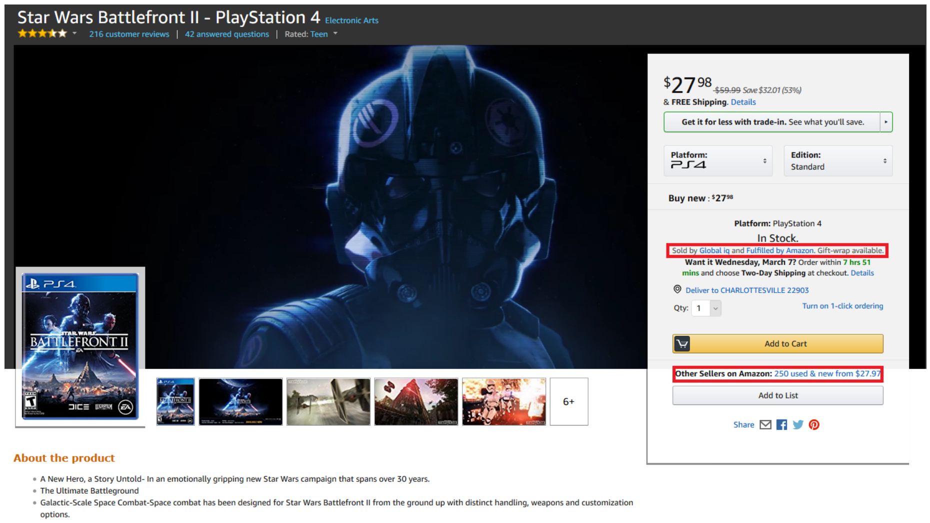

Figure 2.1: Sample product page on Amazon

Notes: The graph shows a sample product page for video game "Star Wars Battlefront II - PlayStation 4". We mark the seller winning the "Buy Box" in the upper red box and mark the link to other sellers in the lower red box.

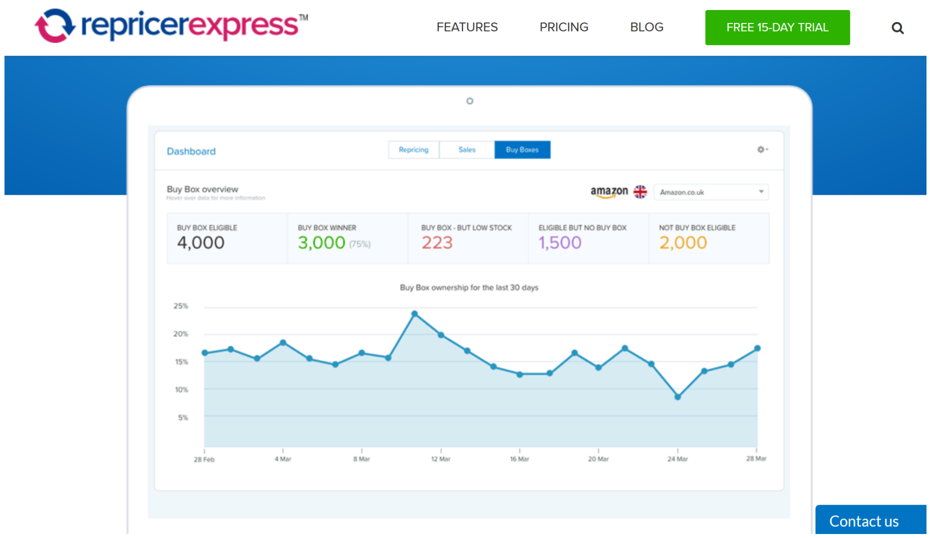

Figure 2.2: Sample Repricing Tool Webpage

Notes: The graph shows a webpage for a third-party repricing tool named "RepricerExpress".

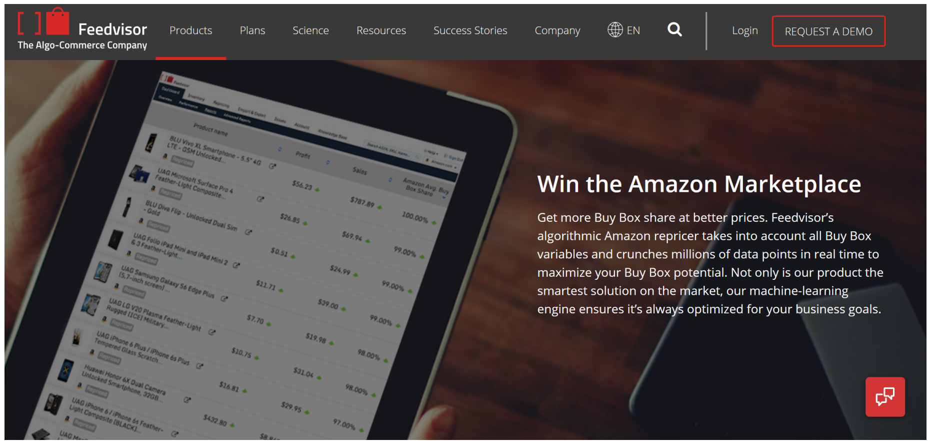

Figure 2.3: Sample Repricing Tool Webpage

Notes: The graph shows a webpage for a third-party repricing tool named "Feedvisor". On the webpage, the repricing tool facilitates sellers to win the "Buy Box".

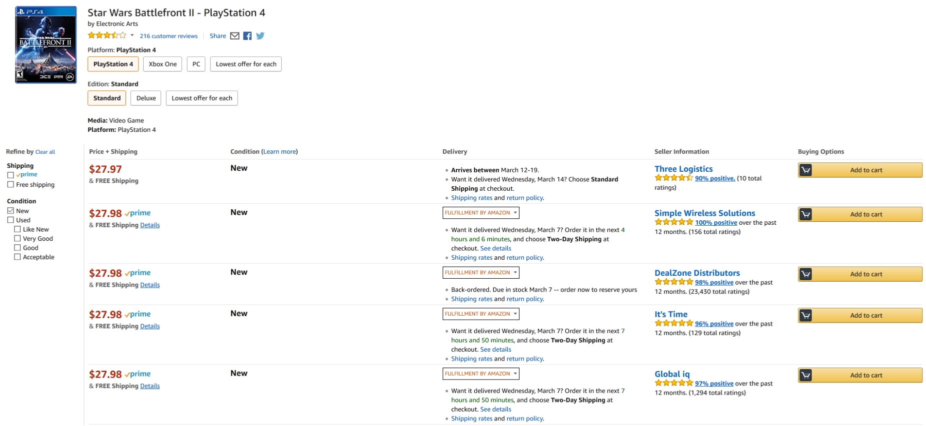

Figure 2.4: Sample list of other sellers on Amazon

Notes: The graph presents a sample list of other sellers for product "Star Wars Battlefront II - PlayStation 4". Sellers are listed in an ascending order of the sum of price and shipping.

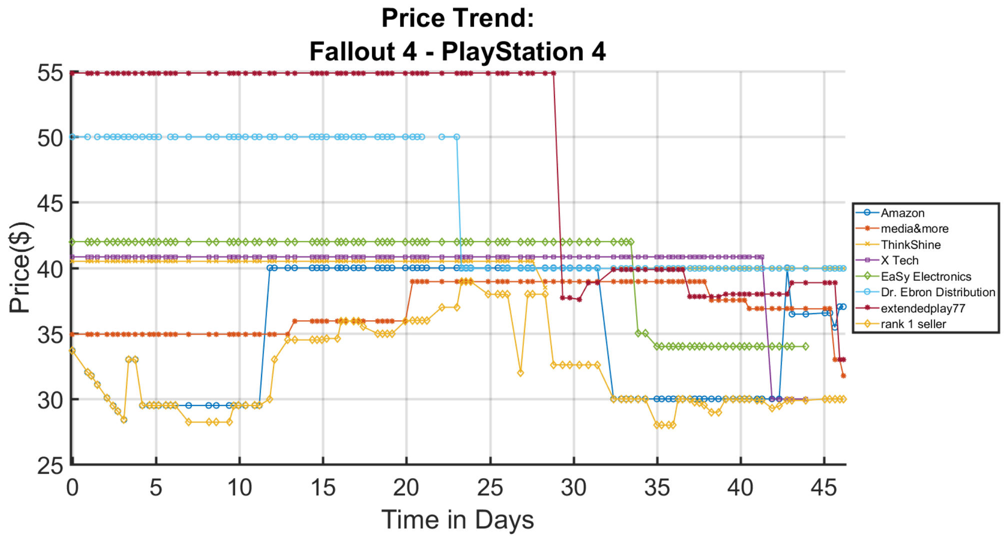

Figure 2.5: Price Trajectory: video game

Notes: The graph presents the price trajectories of selected active sellers and price sequence in the top rank for the product “Fallout 4-PlayStation 4”. Each point corresponds to a landed price (price + shipping) in a day from a seller.

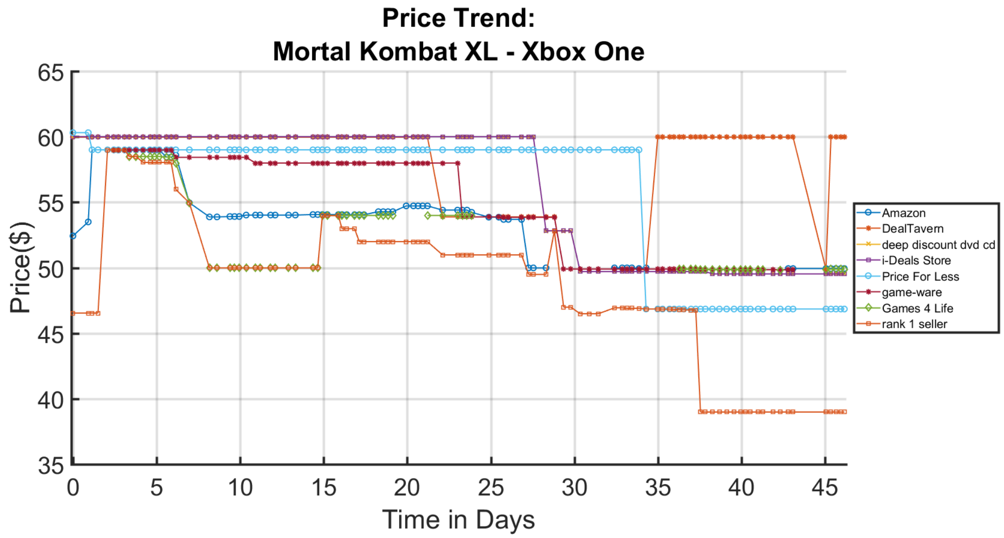

Figure 2.6: Price Trajectory: video game

Notes: The graph presents the price trajectories of selected active sellers and price sequence in the top rank for the product "Mortal Kombat XL-Xbox One". Each point corresponds to a landed price (price + shipping) in a day from a seller.

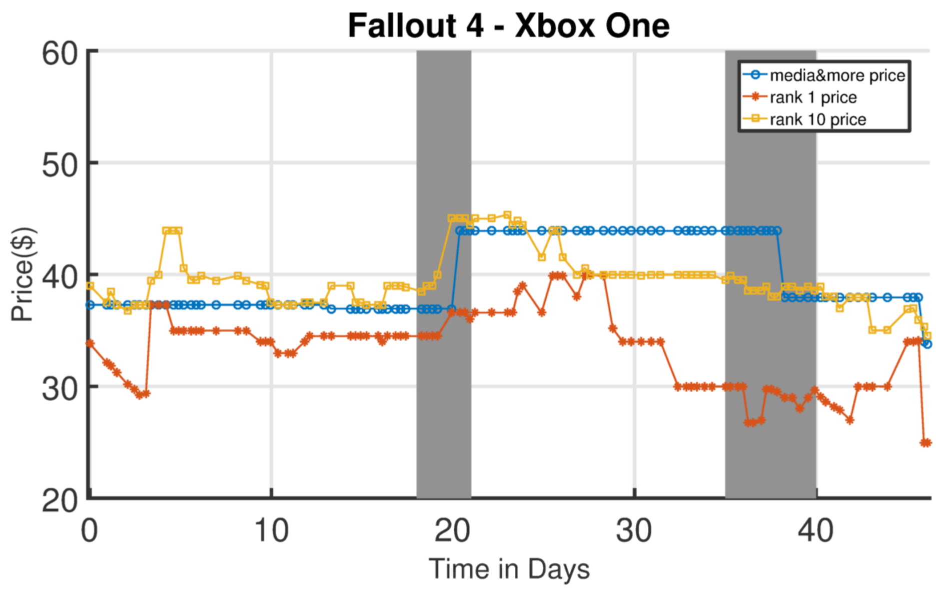

Figure 2.7: Price trajectory: third-party seller vs top 1 and top 10 prices

Notes: The graph presents the price trajectory of seller "media&more" and price sequences in rank 1 and rank 10 for the product "Fallout 4-Xbox One". Each point corresponds to a landed price (price + shipping) in a day from a seller. The grey areas mark the areas of interest.

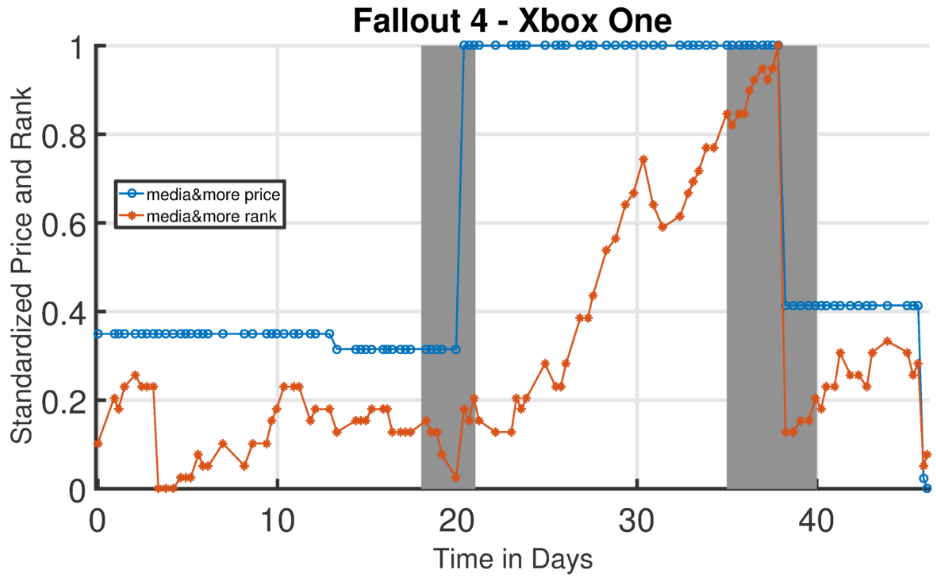

Figure 2.8: Price trajectory: third-party seller vs top 1 and top 10 prices

Notes: The graph presents the standardized price and rank of seller “media&more” for the product “Fallout 4-Xbox One”. We standardize the price sequence and rank sequence based on the maximum and minimum of price and rank in the history. Each point corresponds to a standardized price and standardized rank in a day from a seller. The grey areas mark the areas of interest.

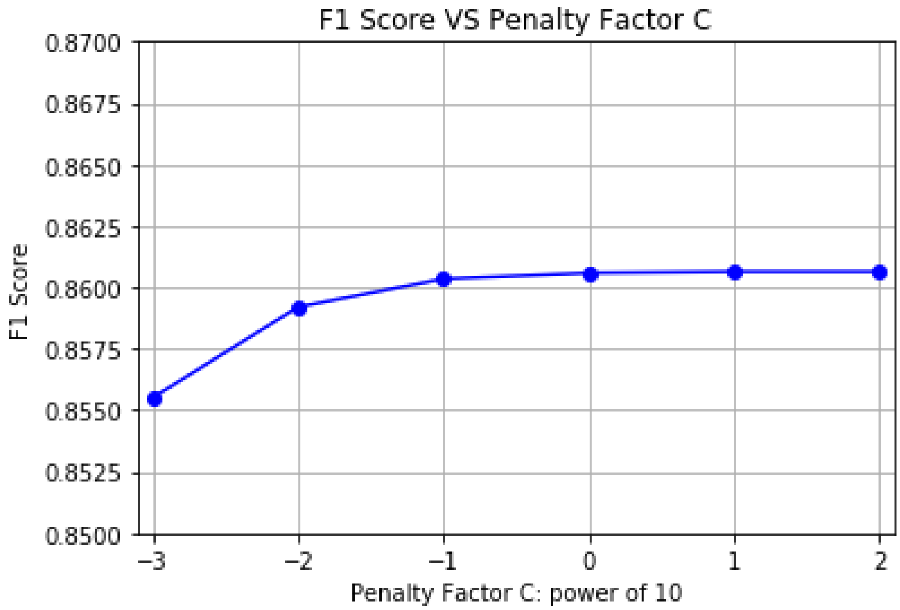

Figure 2.9: F1 Score vs Penalty Factor

Notes: The figure displays the relationship between penalty factor C and F1 score. When penalty factor is larger than 1, the increase of penalty factor only contributes to a marginal increase in the F1 score. Thus, we choose the penalty factor as 1.

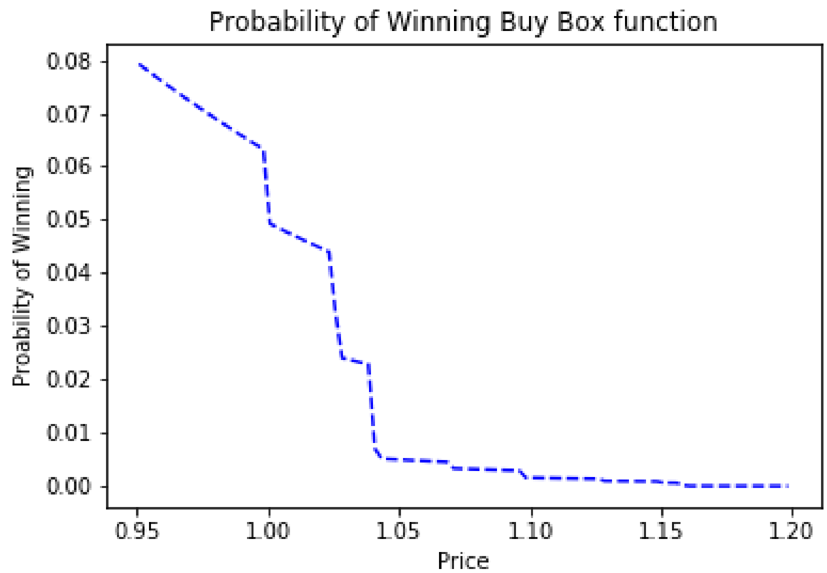

Figure 2.10: Sample Pr function vs Price

Notes: The figure displays a sample Pr function of seller "media&more" in market "Fallout 4 - PlayStation 4". The horizontal axis is the standardized price which is defined by product landed price divided by the lowest landed price in the market at a certain time. The vertical axis is the probability of winning the Buy Box for this seller named "media&more". The Pr function includes seller's price, rating, shipping dummies, ranks, and characteristics describing the competition environment.

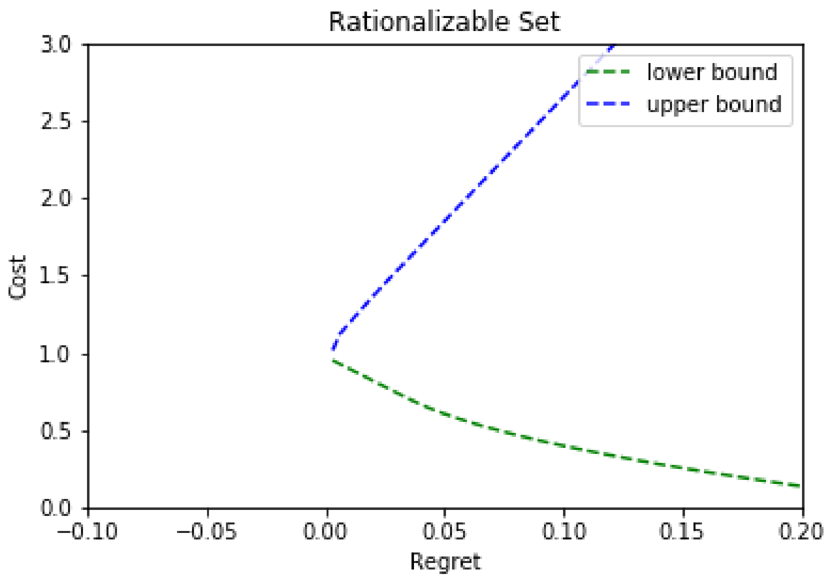

Figure 2.11: The rationalizable set for a high-frequency price changing seller

Notes: The graph presents the rationazliable set for a high-frequency price changing seller. The blue dash line is the upper bound on cost, and the green dash line is the lower bound on cost.

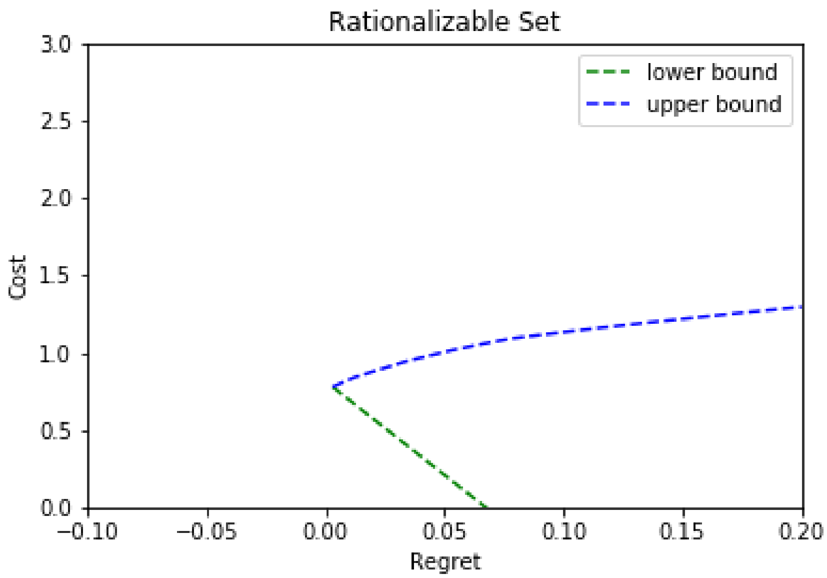

Figure 2.12: The rationalizable set for a high-frequency price changing seller

Notes: The graph presents the rationazliable set for a high-frequency price changing seller. The blue dash line is the upper bound on cost, and the green dash line is the lower bound on cost.

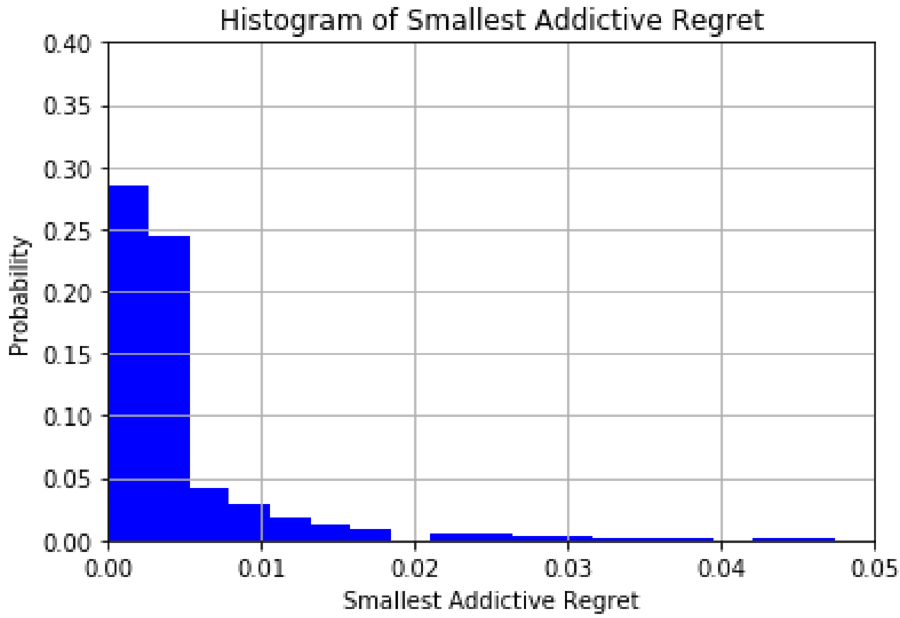

Figure 2.13: Histogram of the corresponding average regret across sellers and markets

Notes: The graph presents the histgram of smallest addictive regret across sellers and markets. For each seller-market combination, we compute the smallest addictive regret and the corresponding cost.

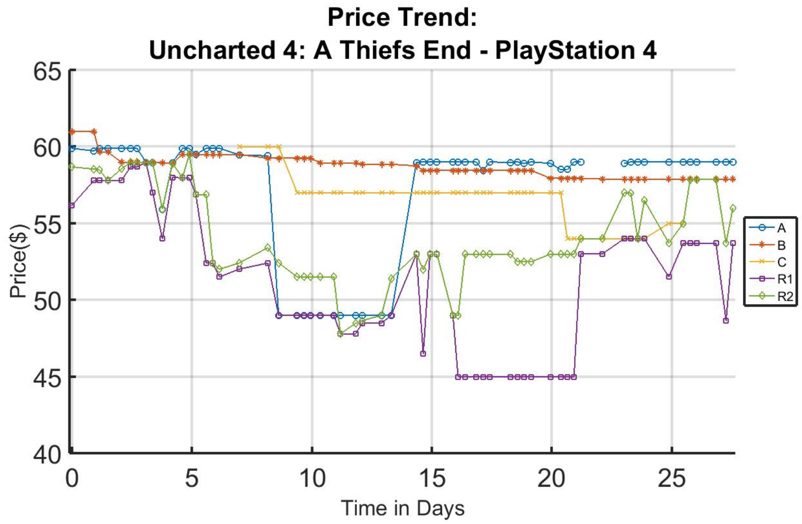

Figure 2.14: Price trend: active sellers and price sequences in rank 1 and 2

Notes: The graph presents the price trajectories of selected sellers (A, B, and C) and price sequences in rank 1 and 2.

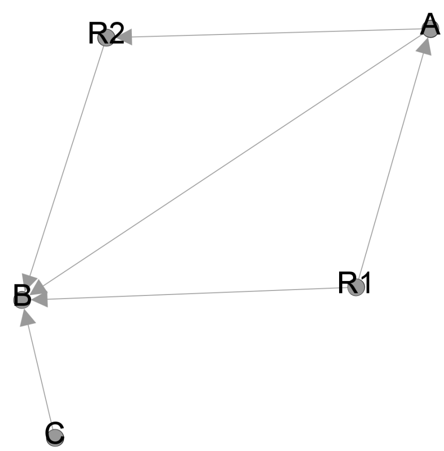

Figure 2.15: Granger Causality Result: active sellers and price sequences in rank 1 and 2

Notes: The graph visualizes the granger causality test results for product "Uncharted 4: Thiefs End - PlayStation 4". An arrowed line from A to B means that seller A granger causes seller B to change prices. Seller A, B, and C are selected active sellers, and seller R1, and R2 are price sequences in rank 1 and 2.

Table 2.1: Summary Statistics

Sample (2016 - 2017)						
Variable	Obs	Mean	Std	25%	50%	75%
own_price	323780	1.29	0.32	1.08	1.21	1.40
FBA	323780	0.38	0.49	0.00	0.00	1.00
rate	323780	89.40	23.11	92.00	96.00	99.00
Amazon	323780	0.04	0.20	0.00	0.00	0.00
FBA_price	323780	1.10	0.22	1.00	1.03	1.13
FBA_rate	323780	97.10	10.79	97.00	99.00	100.00
FBA_price_hr	323780	1.13	0.24	1.00	1.06	1.17
FBA_rate_hr	323780	99.92	0.92	100.00	100.00	100.00
top1_rate	323780	85.09	29.91	92.00	97.00	100.00
Amazon_price	323780	0.87	0.51	1.00	1.02	1.12
No_amazon	323780	0.23	0.42	0.00	0.00	0.00
Sample (2017 - 2018)						
Variable	Obs	Mean	Std	25%	50%	75%
own_price	12971	1.12	0.10	1.03	1.10	1.17
FBA	12971	0.60	0.49	0.00	1.00	1.00
rate	12971	93.05	16.96	94.00	97.00	99.00
Amazon	12971	0.02	0.13	0.00	0.00	0.00
FBA_price	12971	1.06	0.07	1.00	1.02	1.10
FBA_rate	12971	96.72	7.39	96.00	99.00	99.00
FBA_price_hr	12971	1.09	0.08	1.02	1.07	1.14
FBA_rate_hr	12971	99.98	0.14	100.00	100.00	100.00
top1_rate	12971	88.65	23.92	92.00	96.00	99.00
Amazon_price	12971	0.38	0.52	0.00	0.00	1.02
No_amazon	12971	0.65	0.48	0.00	1.00	1.00

Notes: The table presents summary statistics. Each observation corresponds to a landed price, shipping, rate, and competition variables by a seller in a market. "own_price" is a seller's own price normalized by the lowest price in the market. "FBA" is a dummy variable for shipping method. "rate" is a seller's rating, and "Amazon" is a dummy variable if the seller is Amazon. "FBA_price" is the lowest price of FBA seller, and "FBA_rate" is the rating of the lowest FBA seller. "FBA_price_hr" is the price of the highest-rated FBA seller, and "FBA_rate_hr" is the rate of the highest-rated FBA seller."top1_rate" is the rate of the lowest-priced seller. "Amazon_price" is the price of Amazon. "No_amazon" is a dummy variable if Amazon is not present in the market.

Table 2.2: Markup Ratios by Market

Xbox One

Market	25%	50%	75%
72	2.91%	5.59%	12.62%
4	8.05%	12.75%	22.74%
71	17.19%	29.49%	68.63%

PlayStation 4

Market	25%	50%	75%
67	2.30%	4.04%	11.02%
15	8.22%	12.78%	21.35%
60	21.11%	26.51%	36.19%

Notes: The table presents markup ratio quantiles of sample markets. A market is defined as a combination of video game and game console. We have over 16000 seller-market combinations. The market-level markup ratio is calculated by an average of markup ratios of sellers in the market. Markup ratio is defined as the ratio of price-cost difference over cost. Sample markets are selected based on the median markup ratios. We select the markets with the minimum, maximum, and median 50% markup ratio quantiles.

Table 2.3: Markup Ratios by Seller Life Length Group

Category	25%	50%	75%
1	5.53%	11.65%	24.67%
2	8.79%	13.99%	25.41%
3	10.55%	16.39%	29.54%
4	11.02%	16.11%	26.53%
5	11.28%	18.61%	35.98%

Notes: The table presents markup ratio quantiles within each life length group across sellers and markets. A market-seller combination is defined as a listing in the market. We have over 16000 seller-market combinations. For each seller-market combination, we find the minimal addictive regret and derive the corresponding cost. We use the length of periods a seller is in the market as a criteria and group sellers into 5 categories. Category 1 sellers stay in the market for less than 2.6% of the sample period. Category 2 sellers stay in the market from 2.6% to 25% of the sample period. Category 3 sellers stay in the market from 25% to 50% of the sample period, and category 4 sellers stay in the market from 50% to 75% of the sample period. Category 5 sellers stay in the market for more than 75% of the sample period.

Table 2.4: Markup Ratios by Rating Group

Category	25%	50%	75%
1	8.04%	14.35%	27.05%
2	6.87%	13.78%	27.87%
3	6.33%	11.66%	21.84%
4	5.60%	11.78%	24.85%

Notes: The table presents markup ratio quantiles within each rating group across sellers and markets. A market-seller combination is defined as a listing in the market. We have over 16000 seller-market combinations. For each seller-market combination, we find the minimal addictive regret and derive the corresponding cost. We group sellers into 4 rating categories. Category 1 sellers include sellers with 0 rating and sellers just launched on the platform. Category 2 sellers have ratings from 1 to 80, which captures the sellers with a medium rating. Category 3 sellers have ratings from 80 to 90, and category 4 sellers have ratings from 90 to 100.

Table 2.5: Markup Ratios By FBA Group

FBA	25%	50%	75%
0	7.29%	13.05%	25.23%
1	4.25%	10.31%	24.54%

Notes: The table presents markup ratio quantiles with/without FBA shipping across sellers and markets. A market-seller combination is defined as a listing in the market. We have over 16000 seller-market combinations. For each seller-market combination, we find the minimal addictive regret and derive the corresponding cost. The markup ratio is defined as the ratio of price-cost difference over cost. We find the 25%, 50%, and 75% markup ratios for combinations with/without FBA shipping.

Table 2.6: Tracking Ranks by Seller

Tracking Ranks	Share
1&2&10	40.23%
1&2 only	10.40%
2&10 only	3.44%
1&10 only	9.73%
1 only	12.09%
2 only	8.94%
10 only	15.16%

Notes: The table presents the summary of types of responding sellers. Each seller is defined as a listing in a market. 1&2&10 type means a seller responds to price sequences in rank 1, 2, and 10.

Table 2.7: Markup Quantiles by Market

Market	25%	50%	75%
0	6.24%	11.98%	19.35%
1	8.42%	16.39%	26.93%
2	8.21%	14.92%	25.28%
3	9.67%	16.49%	23.82%
4	8.05%	12.75%	22.74%
5	11.34%	21.07%	40.90%
6	6.90%	13.37%	34.53%
7	9.79%	16.60%	28.94%
8	5.42%	8.89%	33.21%
9	15.10%	26.91%	46.43%
10	4.84%	12.24%	27.15%
11	5.33%	11.43%	21.47%
12	6.72%	14.28%	24.38%
13	8.90%	14.05%	21.39%
14	11.98%	24.60%	32.29%
15	8.22%	12.78%	21.35%
16	7.73%	14.70%	32.97%
17	5.80%	7.83%	33.86%
18	7.20%	15.68%	29.84%
19	8.62%	20.96%	44.90%
20	9.31%	17.49%	31.54%
21	7.02%	15.78%	27.07%
22	8.33%	19.39%	41.59%
23	8.00%	13.48%	28.67%
24	10.19%	19.72%	38.41%
25	7.15%	14.17%	26.25%
26	8.55%	14.96%	25.73%
27	13.41%	26.21%	101.39%
28	7.12%	13.60%	30.55%
29	6.90%	11.74%	19.40%
30	5.42%	9.16%	15.76%
31	5.94%	11.81%	25.82%
32	10.27%	18.88%	31.02%
33	5.37%	10.40%	20.00%
34	10.26%	15.94%	27.19%
35	9.87%	17.52%	28.26%
36	6.79%	14.38%	25.93%
37	6.01%	12.24%	21.05%

Notes: The table presents markup ratio quantiles across sellers for all markets. A market is defined as a combination of video game and game console. We have over 16000 seller-market combinations. The market-level markup ratio is calculated by an average of markup ratios of sellers in the market. Markup ratio is defined as the ratio of price-cost difference over cost.

Table 2.8: Markup Quantiles by Market (cont.)

Market	25%	50%	75%
38	8.91%	14.13%	24.81%
39	11.45%	19.62%	45.85%
40	8.71%	13.94%	28.31%
41	9.00%	15.40%	31.48%
42	7.42%	12.64%	23.31%
43	7.25%	13.70%	26.31%
44	3.95%	7.72%	16.00%
45	3.41%	6.82%	13.98%
46	7.43%	15.03%	33.55%
47	5.40%	10.21%	22.25%
48	2.75%	7.43%	32.46%
49	3.24%	7.29%	27.21%
50	2.26%	4.32%	10.88%
51	10.10%	12.40%	20.52%
52	7.65%	11.67%	24.90%
53	7.97%	15.33%	24.41%
54	10.06%	12.11%	20.30%
55	8.81%	11.90%	30.07%
56	7.46%	11.99%	17.42%
57	11.68%	18.87%	29.86%
58	4.81%	10.19%	18.88%
59	5.18%	9.17%	22.82%
60	21.11%	26.51%	36.19%
61	4.02%	7.51%	20.69%
62	5.98%	10.84%	19.89%
63	9.65%	16.06%	27.50%
64	5.67%	12.94%	23.92%
65	2.42%	5.73%	15.27%
66	2.98%	6.00%	14.75%
67	2.30%	4.04%	11.02%
68	1.93%	3.91%	8.09%
69	5.98%	12.63%	38.51%
70	3.68%	6.68%	11.84%
71	17.19%	29.49%	68.63%
72	2.91%	5.59%	12.62%
73	5.34%	9.67%	15.97%
74	6.04%	11.53%	20.25%

Notes: The table presents markup ratio quantiles across sellers for all markets. A market is defined as a combination of video game and game console. We have over 16000 seller-market combinations. The market-level markup ratio is calculated by an average of markup ratios of sellers in the market. Markup ratio is defined as the ratio of price-cost difference over cost.

Table 2.9: Relative Revenue Quantiles by Market

Market	25%	50%	75%
0	1.06	1.12	1.19
1	1.08	1.16	1.27
2	1.08	1.15	1.25
3	1.10	1.16	1.24
4	1.08	1.13	1.23
5	1.11	1.21	1.41
6	1.07	1.13	1.35
7	1.10	1.17	1.29
8	1.05	1.09	1.33
9	1.15	1.27	1.46
10	1.05	1.12	1.27
11	1.05	1.11	1.21
12	1.07	1.14	1.24
13	1.09	1.14	1.21
14	1.12	1.25	1.32
15	1.08	1.13	1.21
16	1.08	1.15	1.33
17	1.06	1.08	1.34
18	1.07	1.16	1.30
19	1.09	1.21	1.45
20	1.09	1.17	1.32
21	1.07	1.16	1.27
22	1.08	1.19	1.42
23	1.08	1.13	1.29
24	1.10	1.20	1.38
25	1.07	1.14	1.26
26	1.09	1.15	1.26
27	1.13	1.26	2.01
28	1.07	1.14	1.31
29	1.07	1.12	1.19
30	1.05	1.09	1.16
31	1.06	1.12	1.26
32	1.10	1.19	1.31
33	1.05	1.10	1.20
34	1.10	1.16	1.27
35	1.10	1.18	1.28
36	1.07	1.14	1.26
37	1.06	1.12	1.21

Notes: The table presents relative revenue quantiles across sellers for all markets. A market is defined as a combination of video game and game console. We have over 16000 seller-market combinations. The market-level relative revenue ratio is calculated by an average of relative revenue ratios of sellers in the market. Relative revenue ratio is defined as the ratio of price over cost.

Table 2.10: Relative Revenue Quantiles by Market (cont.)

Market	25%	50%	75%
38	1.09	1.14	1.25
39	1.11	1.20	1.46
40	1.09	1.14	1.28
41	1.09	1.15	1.31
42	1.07	1.13	1.23
43	1.07	1.14	1.26
44	1.04	1.08	1.16
45	1.03	1.07	1.14
46	1.07	1.15	1.34
47	1.05	1.10	1.22
48	1.03	1.07	1.32
49	1.03	1.07	1.27
50	1.02	1.04	1.11
51	1.10	1.12	1.21
52	1.08	1.12	1.25
53	1.08	1.15	1.24
54	1.10	1.12	1.20
55	1.09	1.12	1.30
56	1.07	1.12	1.17
57	1.12	1.19	1.30
58	1.05	1.10	1.19
59	1.05	1.09	1.23
60	1.21	1.27	1.36
61	1.04	1.08	1.21
62	1.06	1.11	1.20
63	1.10	1.16	1.28
64	1.06	1.13	1.24
65	1.02	1.06	1.15
66	1.03	1.06	1.15
67	1.02	1.04	1.11
68	1.02	1.04	1.08
69	1.06	1.13	1.39
70	1.04	1.07	1.12
71	1.17	1.29	1.69
72	1.03	1.06	1.13
73	1.05	1.10	1.16
74	1.06	1.12	1.20

Notes: The table presents relative revenue quantiles across sellers for all markets. A market is defined as a combination of video game and game console. We have over 16000 seller-market combinations. The market-level relative revenue ratio is calculated by an average of relative revenue ratios of sellers in the market. Relative revenue ratio is defined as the ratio of price over cost.

Table 2.11: Granger Causality Test by Market

Market	# of Active Sellers	# of Responding Sellers(Rank 1)
0	105	2
2	53	9
3	47	2
4	79	60
5	102	39
6	96	3
7	67	15
9	21	0
10	120	10
11	137	102
12	66	7
13	67	50
14	28	6
15	67	23
16	71	51
17	11	2
18	84	28
19	80	16
20	76	72
21	66	51
22	98	96
23	28	15
24	27	24
25	56	35
26	36	10
27	13	3
28	60	53
29	56	2
30	134	4
31	117	73
32	78	64
33	122	12
34	51	21
35	50	40
36	92	82
37	83	9

Notes: The table presents the number of sellers who respond to the price sequence in rank 1 in each market. Each seller is defined as a listing in a market. A seller is defined as a responding seller if the price sequence in certain rank can be used to forecast the seller's price sequence. A seller is defined as an active seller if the seller had prices at least 10 percent of its price history and stayed in top 10 for at least once.

Table 2.12: Granger Causality Test by Market (cont.)

Market	# of Active Sellers	# of Responding Sellers(Rank 1)
38	55	6
39	53	9
40	35	5
41	30	2
42	34	0
43	38	6
44	75	3
45	74	2
46	62	0
47	73	15
48	168	6
49	147	2
50	123	5
51	37	0
52	56	7
53	58	40
54	48	2
55	37	13
56	53	26
57	22	5
58	91	0
59	99	11
61	53	36
62	51	38
63	66	41
64	54	3
65	124	109
66	132	70
67	126	65
68	175	168
69	74	4
70	79	65
71	26	5
72	97	5
73	61	47
74	48	30

Notes: The table presents the number of sellers who respond to the price sequence in rank 1 in each market. Each seller is defined as a listing in a market. A seller is defined as a responding seller if the price sequence in certain rank can be used to forecast the seller's price sequence. A seller is defined as an active seller if the seller had prices at least 10 percent of its price history and stayed in top 10 for at least once.

Table 2.13: Granger Causality Test by Market

Market	# of Active Sellers	# of Responding Sellers(Rank 2)
0	105	5
2	53	44
3	47	2
4	79	16
5	102	35
6	96	21
7	67	11
9	21	0
10	120	8
11	137	96
12	66	6
13	67	24
14	28	6
15	67	15
16	71	15
17	11	2
18	84	4
19	80	48
20	76	73
21	66	6
22	98	96
23	28	10
24	27	25
25	56	20
26	36	24
27	13	3
28	60	51
29	56	2
30	134	5
31	117	103
32	78	56
33	122	11
34	51	35
35	50	40
36	92	49
37	83	15

Notes: The table presents the number of sellers who respond to the price sequence in rank 2 in each market. Each seller is defined as a listing in a market. A seller is defined as a responding seller if the price sequence in certain rank can be used to forecast the seller's price sequence. A seller is defined as an active seller if the seller had prices at least 10 percent of its price history and stayed in top 10 for at least once.

Table 2.14: Granger Causality Test by Market (cont.)

Market	# of Active Sellers	# of Responding Sellers(Rank 2)
38	55	7
39	53	4
40	35	2
41	30	0
42	34	2
43	38	11
44	75	6
45	74	2
46	62	0
47	73	2
48	168	5
49	147	2
50	123	14
51	37	0
52	56	7
53	58	37
54	48	2
55	37	2
56	53	26
57	22	2
58	91	23
59	99	29
61	53	41
62	51	38
63	66	12
64	54	2
65	124	45
66	132	58
67	126	4
68	175	166
69	74	2
70	79	66
71	26	7
72	97	5
73	61	53
74	48	27

Notes: The table presents the number of sellers who respond to the price sequence in rank 2 in each market. Each seller is defined as a listing in a market. A seller is defined as a responding seller if the price sequence in certain rank can be used to forecast the seller's price sequence. A seller is defined as an active seller if the seller had prices at least 10 percent of its price history and stayed in top 10 for at least once.

Table 2.15: Granger Causality Test By Market

Market	# of Active Sellers	# of Responding Sellers(Rank 10)
0	105	8
2	53	12
3	47	6
4	79	44
5	102	45
6	96	4
7	67	22
9	21	2
10	120	24
11	137	97
12	66	13
13	67	12
14	28	20
15	67	51
16	71	16
17	11	2
18	84	16
19	80	20
20	76	64
21	66	25
22	98	96
23	28	11
24	27	10
25	56	32
26	36	6
27	13	5
28	60	51
29	56	2
30	134	12
31	117	106
32	78	46
33	122	6
34	51	11
35	50	38
36	92	34
37	83	21

Notes: The table presents the number of sellers who respond to the price sequence in rank 10 in each market. Each seller is defined as a listing in a market. A seller is defined as a responding seller if the price sequence in certain rank can be used to forecast the seller's price sequence. A seller is defined as an active seller if the seller had prices at least 10 percent of its price history and stayed in top 10 for at least once.

Table 2.16: Granger Causality Test By Market (cont.)

Market	# of Active Sellers	# of Responding Sellers(Rank 10)
38	55	8
39	53	0
40	35	5
41	30	2
42	34	16
43	38	2
44	75	0
45	74	4
46	62	2
47	73	3
48	168	4
49	147	0
50	123	9
51	37	2
52	56	8
53	58	53
54	48	2
55	37	2
56	53	30
57	22	4
58	91	2
59	99	12
61	53	36
62	51	41
63	66	24
64	54	4
65	124	118
66	132	47
67	126	64
68	175	170
69	74	19
70	79	61
71	26	8
72	97	15
73	61	47
74	48	24

Notes: The table presents the number of sellers who respond to the price sequence in rank 10 in each market. Each seller is defined as a listing in a market. A seller is defined as a responding seller if the price sequence in certain rank can be used to forecast the seller's price sequence. A seller is defined as an active seller if the seller had prices at least 10 percent of its price history and stayed in top 10 for at least once.

Bibliography

Aida Ali, Siti Mariyam Shamsuddin, and Anca L. Ralescu. Classification with class imbalance problem: A review. *International Journal of Advances in Soft Computing and its Applications*, 7(3), 2015.

Jason Allen, Robert Clark, and Jean-Francois Houde. The effect of mergers in search markets: Evidence from the canadian mortgage industry. *American Economic Review*, 104(10):3365–96, 2014.

James J. Anton and Gopal Das Varma. Storability, market structure, and demand-shift incentives. *RAND Journal of Economics*, 36(3):520–543, 2005.

Orley C. Ashenfelter, Daniel S. Hosken, and Matthew C. Weinberg. Efficiencies brewed: pricing and consolidation in the us beer industry. *RAND Journal of Economics*, 46(2):328–361, 2015.

John Asker. Diagnosing foreclosure due to exclusive dealing. *The Journal of Industrial Economics*, LXIV(3):375–410, 2016.

Scott R. Baker, Stephanie Johnson, and Lorenz Kueng. Shopping for lower sales tax rates. *NBER Working Paper No. 23665*, 2017.

Anantaram Balakrishnan, Shankar Sundaresan, and Bo Zhang. Browse-and-switch: Retail-online competition under value uncertainty. *Production and Operations Management*, 23(7):1129–1145, 2014.

David R. Bell, Jeongwen Chiang, and V. Padmanabhan. The decomposition of promotional response: An empirical generalization. *Marketing Science*, 18(4):504–526, 1999.

David R. Bell, Ganesh Iyer, and V. Padmanabhan. Price competition under stockpiling and flexible consumption. *Journal of Marketing Research*, 39(3):292–303, 2002.

Omar Besbes and Ilan Lobel. Intertemporal price discrimination: Structure and computation of optimal policies. *Management Science*, 61(1):92–110, 2015.

Omar Besbes and Assaf Zeevi. Dynamic pricing without knowing the demand function: Risk bounds and near-optimal algorithms. *Operations Research*, 57(6):1407–1420, 2009. doi: 10.1287/opre.1080.0640.

Robert C. Blattberg, Gary D. Eppen, and Joshua Lieberman. A theoretical and empirical evaluation of price deals for consumer nondurables. *Journal of Marketing*, 45(1):116–129, 1981.

Robert C. Blattberg, Richard Briesch, and Edward J. Fox. How promotions work. *Marketing Science*, 14(3):122–132, 1995.

H. Onur Bodur, Noreen M. Klein, and Neeraj Arora. Online price search: Impact of price comparison sites on offline price evaluations. *Journal of Retailing*, 91(1):125 –

139, 2015. ISSN 0022-4359. doi: https://doi.org/10.1016/j.jretai.2014.09.003. URL `http://www.sciencedirect.com/science/article/pii/S0022435914000645`.

Abhishek Borah and Gerard J. Tellis. Halo (spillover) effects in social media: Do product recalls of one brand hurt or help rival brands? *Journal of Marketing Research*, 2015.

Severin Borenstein. Airline mergers, airport dominance, and market power. *American Economic Review*, 80(2):400–404, 1990.

Randolph E. Bucklin, Sunil Gupta, and S. Siddarth. Determining segmentation in sales response across consumer purchase behaviors. *Journal of Marketing Research*, 35(2):189–197, 1998.

Grard P. Cachon and Pnina Feldman. Price commitments with strategic consumers: Why it can be optimal to discount more frequently than optimal. *Manufacturing & Service Operations Management*, 17(3):399–410, 2015. doi: 10.1287/msom.2015.0527. URL `https://doi.org/10.1287/msom.2015.0527`.

Grard P. Cachon and Robert Swinney. Purchasing, pricing, and quick response in the presence of strategic consumers. *Management Science*, 55(3):497–511, 2009. doi: 10.1287/mnsc.1080.0948. URL `https://doi.org/10.1287/mnsc.1080.0948`.

Alberto Cavallo. Are online and offline prices similar? evidence from large multi-channel retailers. *American Economic Review*, 107(1):283–303, January 2017. doi: 10.1257/aer.20160542. URL `http://www.aeaweb.org/articles?id=10.1257/aer.20160542`.

Tat Chan, Chakravarthi Narasimhan, and Qin Zhang. Decomposing promotional effects with a dynamic structural model of flexible consumption. *Journal of Marketing Research*, 45(4):487–498, 2008.

Ambarish Chandra and Matthew C. Weinberg. How does advertising depend on competition? evidence from u.s. brewing. *Management Science*, 2018.

Le Chen, Alan Mislove, and Christo Wilson. An empirical analysis of algorithmic pricing on amazon marketplace. In *Proceedings of the 25th International Conference on World Wide Web*, WWW '16, pages 1339–1349, Republic and Canton of Geneva, Switzerland, 2016a. International World Wide Web Conferences Steering Committee. ISBN 978-1-4503-4143-1. doi: 10.1145/2872427.2883089. URL `https://doi.org/10.1145/2872427.2883089`.

Zhiyuan Chen, Xiaoying Liang, and Lei Xie. Inter-temporal price discrimination and satiety-driven repeat purchases. *European Journal of Operational Research*, 251:225–236, 2016b.

Jeongwen Chiang. A simultaneous approach to the whether, what and how much to buy questions. *Marketing Science*, 10(4):297–315, 1991.

Andrew T. Ching and Matthew Osborne. Identification and estimation of forward-looking behavior: The case of consumer stockpiling. *Rotman School of Management Working Paper No. 2594032*, 2017.

Pradeep K. Chintagunta. Investigating purchase incidence, brand choice and purchase quantity decisions of households. *Marketing Science*, 12(2):184–209, 1993.

John Conlisk, Eitan Gerstner, and Joel Sobel. Cyclic pricing by a durable goods monopolist. *The Quarterly Journal of Economics*, 99(3):489–505, 1984.

Lemore Dafny, Mark Duggan, and Subramaniam Ramanarayanan. Paying a premium on your premium? consolidation in the us health insurance industry. *American Economic Review*, 102(2):1161–85, 2012.

Marnik G. Dekimpe and Dominique M. Hanssens. The persistence of marketing effects on sales. *Marketing Science*, 14(1):1–21, 1995.

Marnik G. Dekimpe and Dominique M. Hanssens. Time-series models in marketing: Past, present and future. *International Journal of Research in Marketing*, 17(2): 183–193, 2000.

Peter Doyle and John Saunders. Multiproduct advertising budgeting. *Marketing Science*, 9(2):97–113, 1990.

Kenneth G. Elzinga. The u.s. beer industry: Concentration, fragmentation, and a nexus with wine. *Journal of Wine Economics*, 6(2):217–230, 2011.

Walter Enders. In *Applied Econometric Time Series*. John Wiley, New York, 2004.

Ying Fan. Ownership consolidation and product characteristics: A study of the us daily newspaper market. *American Economic Review*, 103(5):1598–1628, 2013.

Kris Johnson Ferreira, Bin Hong Alex Lee, and David Simchi-Levi. Analytics for an online retailer: Demand forecasting and price optimization. *Manufacturing & Service Operations Management*, 18(1):69–88, 2016. doi: 10.1287/msom.2015.0561. URL `https://doi.org/10.1287/msom.2015.0561`.

Marshall Fisher, Santiago Gallino, and Jun Li. Competition-based dynamic pricing in online retailing: A methodology validated with field experiments. *Management Science*, 0(0), 2017. doi: 10.1287/mnsc.2017.2753. URL https://doi.org/10.1287/mnsc.2017.2753.

Philip Hans Franses. Primary demand for beer in the netherlands: An application of armax model specification. *Journal of Marketing Research*, 28(2):240–245, 1991.

Guillermo Gallego and Garrett van Ryzin. Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Management Science*, 40(8):999–1020, 1994.

Manish Gangwar, Nanda Kumar, and Ram C. Rao. Consumer stockpiling and competitive promotional strategies. *Marketing Science*, 33(1):94–113, 2014.

Vishal Gaur and Young-Hoon Park. Asymmetric consumer learning and inventory competition. *Management Science*, 53(2):227–240, 2007. doi: 10.1287/mnsc.1060.0615. URL https://doi.org/10.1287/mnsc.1060.0615.

Liang Guo and J. Miguel Villas-Boas. Consumer stockpiling and price competition in differentiated markets. *Journal of Economics & Management Strategy*, 16(4):827–858, 2007.

Sunil Gupta. Impact of sales promotions on when, what, and how much to buy. *Journal of Marketing Research*, 25(4):342–355, 1988.

Harald J. Van Heerde, Sachin Gupta, and Dick R. Wittink. Is 75% of the sales promotion bump due to brand switching? no, only 33% is. *Marketing Science*, 40(4):481–491, 2003.

Paul Heidhues and Botond Koszegi. Regular prices and sales. *Theoretical Economics*, 9:217 – 251, 2014.

Richard M. Helmer and Johny K. Johansson. An exposition of the box-jenkins transfer function analysis with an application to the advertising-sales relationship. *Journal of Marketing Research*, 14(2):227–239, 1977.

Igal Hendel and Aviv Nevo. Measuring the implications of sales and consumer inventory behavior. *Econometrica*, 74(6):1637–1673, 2006a.

Igal Hendel and Aviv Nevo. Sales and consumer inventory. *RAND Journal of Economics*, 37(3):543–561, 2006b.

Igal Hendel and Aviv Nevo. Intertemporal price discrimination in storable goods markets. *American Economic Review*, 103(7):2722–51, 2013.

Ken Heyer, Carl Shapiro, and Jeffery Wilder. The year in review: Economics at the antitrust division, 2008-2009. *Review of Industrial Organization*, 35:349–367, 2008.

Marit Hinnosaar. Time inconsistency and alcohol sales restrictions. *European Economic Review*, 87:108–131, 2016.

Pilky Hong, R.Preston McAfee, and Ashish Nayyar. Equilibrium price dispersion with consumer inventories. *Journal of Economic Theory*, 105:503–517, 2002.

Horizontal Merger Guidelines. *Horizontal Merger Guidelines*. U.S. Department of Justice and the Federal Trade Commission, 2010.

Pooya Jalaly, Denis Nekipelov, and Eva Tardos. Learning and trust in auction markets. *ArXiv e-prints*, March 2017.

Amit Joshi and Dominique M. Hanssens. The direct and indirect effects of advertising spending on firm value. *Journal of Marketing*, 74(1):20–33, January 2010.

Daniel Kahneman and Amos Tversky. Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–291, 1979.

E.Han Kim and Vijay Singal. Mergers and market power: Evidence from the airline industry. *American Economic Review*, 83(3):549–569, 1993.

Rajiv Lal. Manufacturer trade deals and retail price promotions. *Journal of Marketing Research*, 27(4):428–444, 1990.

Rajiv Lal and Carmen Matutes. Retail pricing and advertising strategies. *The Journal of Business*, 67(3):345–370, 1994.

Eyal Lanxner. The amazon buy box: How it works for sellers, and why its so important, 2018. URL `https://www.bigcommerce.com/blog/win-amazon-buy-box/`.

Robert P. Leone. Modeling sales-advertising relationships: An integrated time series-econometric approach. *Journal of Marketing Research*, 20(3):291–295, 1983.

Yuri Levin, Jeff McGill, and Mikhail Nediak. Dynamic pricing in the presence of strategic consumers and oligopolistic competition. *Management Science*, 55(1): 32–46, 2009. doi: 10.1287/mnsc.1080.0936. URL `https://pubsonline.informs.org/doi/abs/10.1287/mnsc.1080.0936`.

Jun Li, Nelson Granados, and Serguei Netessine. Are consumers strategic? structural estimation from the air-travel industry. *Management Science*, 60(9):2114–2137, 2014. doi: 10.1287/mnsc.2013.1860. URL `https://doi.org/10.1287/mnsc.2013.1860`.

Jun Li, Serguei Netessine, and Sergei Koulayev. Price to compete with many: How to identify price competition in high-dimensional space. *Management Science*, 0 (0), 2017. doi: 10.1287/mnsc.2017.2820. URL https://doi.org/10.1287/mnsc.2017.2820.

Rushi Longadge, Snehlata S. Dongre, and Latesh Malik. Class imbalance problem in data mining: Review. *International Journal of Computer Science and Network (IJCSN)*, 2(1), 2013.

Carl F. Mela, Kamel Jedidi, and Douglas Bownman. The long-term impact of promotions on consumer stockpiling behavior. *Journal of Marketing Research*, 35(2): 250–262, 1998.

Nathan H. Miller and Matthew C. Weinberg. Understanding the price effects of the miller/coors joint venture. 2016.

Ken Moon, Kostas Bimpikis, and Haim Mendelson. Randomized markdowns and online monitoring. *Management Science*, 2017. doi: 10.1287/mnsc.2016.2661. URL https://doi.org/10.1287/mnsc.2016.2661.

Denis Nekipelov, Vasilis Syrgkanis, and Eva Tardos. Econometrics for learning agents. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation*, EC '15, pages 1–18, New York, NY, USA, 2015. ACM. ISBN 978-1-4503-3410-5. doi: 10.1145/2764468.2764522. URL http://doi.acm.org/10.1145/2764468.2764522.

Whitney K. Newey. Generalized method of moments specification testing. *Journal of Econometrics*, 29(3):229–256, 1985.

Vincent Nijs, Kanishka Misra, Eric T. Anderson, Karsten Hansen, and Lakshman Krishnamurthi. Channel pass-through of trade promotions. *Marketing Science*, 29(2):250–267, 2009.

Vincent R. Nijs, Marnik G. Dekimpe, Jan-Benedict E.M. Steenkamps, and Dominique M. Hanssens. The category-demand effects of price promotions. *Marketing Science*, 20(1):1–22, 2001.

Vincent R. Nijs, Shuba Srinivasan, and Koen Pauwels. Retail-price drivers and retailer profits. *Marketing Science*, 26(4):473–487, 2007.

Koen Pauwels, Dominique M. Hanssens, and S. Siddarth. The long-term effects of price promotions on category incidence, brand choice, and purchase quantity. *Journal of Marketing Research*, 39(4):421–439, 2002.

Helena Perrone. Demand for nondurable goods: a shortcut to estimating long-run price elasticities. *The RAND Journal of Economics*, 48(3):856–873, 2017.

Abel P.Jeuland and Chakravarthi Narasimhan. Dealing-temporary price cuts-by seller as a buyer discrimination mechanism. *The Journal of Business*, 58(3):295–308, 1985.

RepricerExpress. How to win the amazon buy box in 2018, 2018. URL `https://www.repricerexpress.com/win-amazon-buy-box/`.

S. Salop and J. E. Stiglitz. The theory of sales: A simple model of equilibrium price dispersion with identical agents. *The American Economic Review*, 72(5):1121–1130, 1982.

Paola Sapienza. The effects of banking mergers on loan contracts. *The Journal of Finance*, 57(1):329–367, 2002.

Joel Sobel. The timing of sales. *The Review of Economic Studies*, 51(3):353–368, 1984.

Shuba Srinivasan, Peter T.L. Popkowski Leszczyc, and Frank M. Bass. Market share response and competitive interaction: The impact of temporary, evolving and structural changes in prices. *International Journal of Research in Marketing*, 17(4): 281–305, 2000.

Shuba Srinivasan, Koen Pauwels, Dominique M. Hanssens, and Marnik G. Dekimpe. Do promotions benefit manufacturers, retailers, or both. *Management Science*, 00(0):1–13, 2004.

Statista. Amazon statistics and facts, 2017a. URL `https://www.statista.com/topics/846/amazon/`.

Statista. Global net revenue of amazon.com from 2014 to 2016, by segment (in billion u.s. dollars), 2017b. URL `https://www.statista.com/statistics/672747/amazons-consolidated-net-revenue-by-segment/`.

Statista. Percentage of paid units sold by third-party sellers on amazon platform as of 4th quarter 2017, 2017c. URL `https://www.statista.com/statistics/259782/third-party-seller-share-of-amazon-platform/`.

Thomas J. Steenburgh. Measuring consumer and competitive impact with elasticity decompositions. *Journal of Marketing Research*, 44(4):636–646, 2007.

Xuanming Su. Intertemporal pricing with strategic customer behavior. *Management Science*, 53(5):726–741, 2007. doi: 10.1287/mnsc.1060.0667. URL https://doi.org/10.1287/mnsc.1060.0667.

Xuanming Su. Intertemporal pricing and consumer stockpiling. *Operations Research*, 58(4):1133–1147, 2010.

Xuanming Su and Fuqiang Zhang. Strategic customer behavior, commitment, and supply chain performance. *Management Science*, 54(10):1759–1773, 2008. doi: 10.1287/mnsc.1080.0886. URL https://pubsonline.informs.org/doi/abs/10.1287/mnsc.1080.0886.

Andrew Sweeting. The effects of mergers on product positioning: evidence from the music radio industry. *RAND Journal of Economics*, 41(2):372–397, 2010.

Andrew Sweeting and Xuezhen Tao. Dynamic oligopoly pricing with asymmetric information: Implications for mergers. 2016.

Victor J. Tremblay and Carol Horton Tremblay. In *The US. Brewing Industry: Data and Economic Analysis*. MIT Press, Cambridge, MA, 2005.

Harald J. van Heerde, Peter S. H. Leeflang, and Dick R. Wittink. Decomposing the sales promotion bump with store data. *Marketing Science*, 23(3):317–334, 2004.

Hal R. Varian. A model of sales. *The American Economic Review*, 70(4):651–659, 1980.

Sofia Berto Villas-Boas and J. Miguel Villas-Boas. Learning, forgetting, and sales. *Management Science*, 54(11):1951 – 1960, 2008.

Richard Volpe, Corey Risch, and Michael Boland. The determinants of price adjustments in retail supermarkets. *Managerial and decision Economics*, 38:37–52, 2017.

Emily Wang, Christian Rojas, and Francesca Colantuoni. Heterogeneous behavior, obesity, and storability in the demand for soft drinks. *American Journal of Agricultural Economics*, 99(1):18–33, 2017.

Emily Yucai Wang. The impact of soda taxes on consumer welfare: implications of storability and taste heterogeneity. *The RAND Journal of Economics*, 46(2): 409–441, 2015.

WebRetailer. 10 statistics from the online marketplace seller survey, 2014. URL `https://www.webretailer.com/lean-commerce/statistics-marketplace-seller-survey/`.

Kanghyun Yoon and Thanh V. Tran. Revisiting the relationship between consumer loyalty and price sensitivity: The moderating role of deal-proneness. *Journal of Marketing Theory and Practice*, 19(3):293–306, 2011.

Dennis Zhang, Hengchen Dai, Lingxiu Dong, Fangfang Qi, Nannan Zhang, Xiaofei Liu, and Zhongyi Liu. How does dynamic pricing affect customer behavior on retailing platforms? evidence from a large randomized experiment on alibaba. vailable at SSRN: https://ssrn.com/abstract=3029707, 2017.