Oligopolistic Price Leadership and Mergers: The United States Beer Industry*

Nathan H. Miller†
Georgetown University

Gloria Sheu‡
U.S. Department of Justice

Matthew C. Weinberg§
The Ohio State University

May 31, 2019

Abstract

We study an infinitely-repeated game of oligopolistic price leadership in which one firm, the leader, proposes a supermarkup over Bertrand prices to a coalition of rivals. We estimate the model with aggregate scanner data on the beer industry and find the supermarkup accounts for 6% of price. Price leadership increases profit by 8.9% relative to Bertrand competition, and decreases consumer surplus by nearly four times the change in profit. We use the model to simulate the ABI/Modelo merger. The merger relaxes incentive compatibility constraints and increases the equilibrium supermarkup. Merger efficiencies do not mitigate—and can amplify—this coordinated effect.

Keywords: price leadership, coordinated effects, mergers
JEL classification: K21; L13; L41; L66

*This material is based on work supported by the National Science Foundation under Grant Nos. 1824318 and 1824332. We thank seminar participants at the Federal Trade Commission, Harvard Business School, New York University, The Ohio State University, Pennsylvania State University, Princeton University, and Texas A&M. All estimates and analyses in this paper based on IRI data are by the authors and not by IRI. The views expressed herein are entirely those of the authors and should not be purported to reflect those of the U.S. Department of Justice.

†Georgetown University, McDonough School of Business, 37th and O Streets NW, Washington DC 20057. Email: nhm27@georgetown.edu.

‡U.S. Department of Justice, Antitrust Division, Economic Analysis Group, 450 5th St. NW, Washington DC 20530. Email: gloria.sheu@usdoj.gov.

§The Ohio State University, 410 Arps Hall, 1945 N. High Street, Columbus OH 43210. Email: weinberg.133@osu.edu.

1 Introduction

Firms in concentrated industries sometimes change their prices by similar magnitudes, with the changes initiated by a single firm. We follow Bain (1960) in referring to this pricing pattern as *oligopolistic price leadership*. The subject has a long history in the economics literature. Anecdotal examples are discussed in Scherer (1980) and an older series of articles (e.g., Stigler (1947); Markham (1951); Oxenfeldt (1952)). More recent studies utilizing extremely detailed data document leader/follower pricing in retail industries ranging from supermarkets, pharmacies, and gasoline (Clark and Houde (2013); Seaton and Waterson (2013); Chilet (2018); Lemus and Luco (2018); Byrne and de Roos (2019).[1] However, as these studies are largely descriptive, existing research does not examine the effectiveness of price leadership in supporting supracompetitive markups, explore implications for welfare, nor provide a framework for the analysis of counterfactuals.[2]

This paper presents an empirical model of oligopolistic price leadership that can be estimated with aggregate scanner data on prices and quantities. Our organizing premise is that price leadership may enable oligopolists to select among the many equilibria that exist in repeated pricing games (e.g., Friedman (1971); Abreu (1988)). The leader's price announcement provides a focal point that guides the prices of other firms. Although supracompetitive prices can result, information disseminates through normal market interactions, avoiding the explicit agreements frequently targeted by antitrust authorities. We apply the model to a setting for which there is documentary evidence of price leadership behavior—the United States beer industry. Once the model is estimated, we quantify the implications of price leadership for firms and consumers. We believe our research represents one of the first attempts to estimate a fully-specified structural model of price coordination.

One practical benefit of our approach is that it supports counterfactual analyses. This leads to our second main contribution, which is to provide a framework for evaluating the *coordinated effects* of mergers in markets characterized by price leadership. In our application, to the Anheuser-Busch InBev (ABI) acquisition of Grupo Modelo, we conceptualize coordinated effects as involving a movement from one supracompetitive equilibrium to another. Although antitrust authorities have long reviewed mergers for coordinated effects, the

[1]See also the discussions in Lanzillotti (2017) and Harrington and Harker (2017). In the popular press, see "Drugmakers Find Competition Doesn't Keep a Lid on Prices" by Jonathan D. Rockoff, Wall Street Journal, November 27, 2016 and "Your Chocolate Addiction is Only Going to Get More (and More, and More) Expensive" by Roberto A. Ferdman, Washington Post, July 18, 2014.

[2]The study by Clark and Houde (2013) is an exception in that it uses a repeated pricing game to study the efficacy of a strategy employed by a known cartel of gasoline retailers.

empirical industrial organization literature to date has provided little in the way of methodologies that could be used guide these efforts. Indeed, our research is among the first to formally model coordinated effects in real-world markets.[3]

We organize the paper as follows. We start with a description of U.S. brewing markets (Section 2). In scanner data spanning 2001-2011, four firms ultimately account for about 80% of retail revenue. We cite to legal documents filed by the Department of Justice (DOJ) alleging that ABI pre-announces its annual list price changes as a signal to competitors, and that its largest competitor, MillerCoors, tends to follow. We show an abrupt increase in the prices of ABI and MillerCoors shortly after the 2008 consummation of the Miller/Coors merger, both in absolute terms and relative to the prices of Modelo and Heineken, the other large brewers. The changes are difficult to rationalize with post-merger Bertrand competition (Miller and Weinberg (2017)) and play an important role in our identification strategy. We describe the differentiated-products model of consumer demand estimated in Miller and Weinberg (2017), which we take as given in this paper.

We then formalize the model of oligopolistic price leadership (Section 3). Firms compete in an infinitely repeated differentiated-products pricing game of perfect information. Each period has two stages. In the first, the leader announces a "supermarkup" above Bertrand prices. On the equilibrium path, a set of coalition firms, comprised of the leader and its followers, accept the supermarkup in a subsequent pricing stage. The leader selects the supermarkup to maximize its profit, subject to incentive compatibility (IC) constraints of the followers and, in order for the announcement to be credible, itself. The leader also accounts for the reaction of fringe firms, each of which prices to maximize current profits. We assume any deviation from the leader's supermarkup by a coalition firm is punished with infinite reversion to the Bertrand equilibrium. A perfect equilibrium exists under a sensible set of beliefs, and we label it the price leadership equilibrium (PLE).

We discuss identification and estimation in Section 4. Our main identification result is that the marginal costs that rationalize prices can be recovered for any candidate supermarkup. The connection flows through the Bertrand first order conditions (e.g., Rosse (1970)), although multiple numerical steps are required in implementation because Bertrand prices are unobserved. With this result in hand, a structural error term in the marginal cost function can be isolated, allowing for estimation with the method of moments. A final com-

[3]We refer readers to Baker (2001, 2010) and Harrington (2013) for a summary of the legal literature on coordinated effects. The theoretical literature includes Compte et al. (2002); Vasconcelso (2005); Ivaldi et al. (2007); Bos and Harrington (2010); and Loertscher and Marx (2019). Empirical models include Davis and Huse (2010) and Igami and Sugaya (2019).

plication is that the objects of interest in estimation (the supermarkups) are choice variables rather than structural parameters. Thus, a fully unrestricted model is under-identified as theory indicates that equilibrium supermarkups adjust with variation in valid instruments. In our application, we assume Bertrand competition prior to the Miller/Coors merger, which is sufficient for exact identification of the post-merger supermarkup.

We estimate supermarkups that range from \$0.60 to \$0.74 (Section 5), depending on the specific demand specification employed. For context, \$0.60 is about six percent of the average price of a 12 pack. Price leadership increases total industry profits by about ten percent relative to Bertrand competition. Consumer surplus decreases almost four times more than profit increases, as consumers pay more and may select less-preferred brands in response to higher prices. In counterfactual simulations, we find that higher supermarkups would increase ABI's profit. Thus, to rationalize pricing within the model, an IC constraint binds. This suggests that the economic consequences of price leadership may be sensitive to market structure—which affects the profit firms receive from coordination, deviation, and punishment. Indeed, as we develop shortly, this is the case.

To conduct counterfactuals that alter market structure, however, it is necessary to recover the parameters that enter the IC constraints. In our framework, these include the discount factor and an antitrust risk coefficient (which measures a disutility of coordination). As the results indicate an IC must bind, it must be that the present value of coordination and deviation are equal at the estimated supermarkup, for at least one firm. The implied equality constraint jointly identifies the parameters because the other inputs to the IC constraints—the profit of coordination, deviation, and punishment evaluated at the estimated supermarkup—are easily recovered using simple counterfactual simulations. Our analysis indicates that the IC of MillerCoors is the constraint on post-merger prices.

In Section 6, we use the model to examine the coordinated effects of ABI's acquisition of Modelo, approved in 2013 by the DOJ only after the Modelo brands were divested to a third party. We model the merger as it would have occurred without the divestiture. The DOJ Complaint characterizes Modelo as a maverick, defined in the Horizontal Merger Guidelines as "a firm that has often resisted otherwise prevailing industry norms to cooperate on price setting or other terms of competition." Mavericks are naturally incorporated as fringe firms in our framework. Our simulation results indicate that bringing Modelo into the coalition (as part of ABI) loosens the IC constraints of MillerCoors and allows ABI to support substantially higher supermarkups in equilibrium. Our most conservative simulation indicates the merger would increase the profit of ABI and Modelo by 5.95%, decrease consumer surplus by 2.64%, and decrease total surplus by 2.02%.

The coordinated effects of ABI/Modelo are not mitigated by marginal cost efficiencies. Because the IC constraint of MillerCoors binds, the marginal costs of ABI and Modelo affect the supermarkup only to the extent they influence MillerCoors' incentives. Indeed, our analysis shows that merger efficiencies cause a modest *increase* in the equilibrium supermarkup. The reason is that merger efficiencies reduce the profit that MillerCoors receives in the event of punishment (i.e., in Bertrand equilibrium) and this loosens the MillerCoors IC constraint. Thus, our analysis suggests the standard treatment of merger efficiencies as a countervailing influence may be more specific to static Nash-Bertrand and Nash-Cournot models than previously recognized.

We conclude in Section 7 with a short summary and a discussion of some of the more important modeling assumptions, with an eye toward informing future research efforts.

1.1 Literature Review

Our research connects to several literatures. We draw on a number of theoretical articles in building the empirical model. Most similar is the canonical Rotemberg and Saloner (1986) model of collusion, in which there is perfect information and collusive prices adjust to ensure that deviation does not occur along the equilibrium path. A repeated game in which oligopolistic price leadership emerges is provided in Rotemberg and Saloner (1990).[4] As their model incorporates asymmetric information, price announcements have informational and strategic content. Our model is simpler in that announcements have only strategic content, and can be interpreted as cheap talk (e.g., Farrell (1987); Farrell and Rabin (1996)) or as providing an endogenous focal point that selects among equilibria.[5] We take as given that price announcements shape firm beliefs about subsequent play.

A number of theoretical articles develop results on the organization of coalitions. Ishibashi (2008) and Mouraviev and Rey (2011) analyze repeated games in which (each period) the leader sets price in an initial stage and other firms set price in a subsequent stage; cartel profits are maximized by having the firm with the greatest incentive to deviate serve as the leader. Pastine and Pastine (2004) analyze a similar game in which a war of attrition determines the leader. Our model differs in that each period features an announce-

[4]In the earlier literature, Stigler (1947) emphasizes that price leadership may arise if one firm is better informed about the economic state, while Markham (1951) argues that its function may be to soften competition. See also Oxenfeldt (1952). These articles were motivated by a Supreme Court decision in which price leadership in the tobacco industry was determined to violate antitrust statutes (Nicholls (1949)).

[5]The notion that exogenous focal points may help firms coordinate in games with multiple equilibria dates at least to Schelling (1960); see also Knittel and Stango (2003) for an empirical analysis.

ment followed by simultaneous pricing, rather than sequential pricing.[6] Under the timing and informational assumptions we maintain, any coalition firm could serve as the leader, and thus we assume the leader is exogenously determined. In allowing for partial coalitions, we build on a literature that considers homogeneous-product quantity games (e.g., d'Aspremont et al. (1983), Donsimoni et al. (1986), and Bos and Harrington (2010)).

With respect to the empirical literature, our research is methodologically most similar to Igami and Sugaya (2019) on the vitamin C cartel of the 1990s.[7] The main result is that unexpected shocks to demand and fringe supply undermined incentive compatibility and led to the collapse of the cartel. As in our research, Igami and Sugaya estimate the structural parameters of a supergame in which trigger strategies sustain supracompetitive prices, and rely on counterfactual simulations to recover the profit terms that enter the IC constraints. There are also important differences. Igami and Sugaya assume all firms either engage in maximal collusion or revert to Cournot equilibrium. Thus, some interesting aspects of our model, such as partial coalitions and the leader's ability to adjust the supermarkup to satisfy incentive compatibility, are not present in their setup.

A number of empirical and theoretical articles have highlighted that mergers can make coordination more difficult to sustain by softening competition in punishment phases (e.g., Davidson and Deneckere (1984); Werden and Baumann (1986); Davis and Huse (2010)). Our counterfactual analyses of the ABI/Modelo merger incorporates this effect. However, by allowing for higher supermarkups, the merger also increases the gains to coordination, and we find this second effect dominates.

Our research relates to articles that seek to understand the equilibrium concept that governs competition in specific markets. Two of the more prominent focus on Bertrand equilibrium and joint profit maximization (e.g., Bresnahan (1987); Nevo (2001)), while others also explore Stackleberg leadership and other possibilities (e.g., Gasmi et al. (1992); Slade (2004); Rojas (2008)). The conduct parameter approach also can be used to test for changes in the equilibrium concept (e.g., Porter (1983); Ciliberto and Williams (2014); Igami (2015); Miller and Weinberg (2017); Michel and Weiergraeber (2018)). Closest to our research is Miller and Weinberg, as it uses the same data sample and demand model. The conduct parameter approach, however, abstracts from the underlying supergame and thus cannot support the counterfactual analyses conducted in the present research.

[6]As discussed above, Rotemberg and Saloner (1990) also model price leadership as involving non-binding announcements. See also Marshall et al. (2008) on price announcements in the vitamins cartels of the 1990s.

[7]Also similar is contemporaneous research of Eizenberg and Shilian (2019), which tests for Bertrand pricing in a number of Israeli food sectors. Marginal costs are recovered from first order conditions, and then the profit terms that enter IC constraints are obtained with counterfactual simulations.

2 The U.S Beer Market

2.1 Background

Most beer sold in the Unites States is produced by a handful of large brewers that compete across the country. These brewers compete in prices, product introduction, advertising, and periodic sales. The product offerings typically are characterized as differentiated along multiple dimensions, including taste, calories, brand image, and package size. The beer industry differs from typical retail consumer product industries in its vertical structure because of state laws regulating the sales and distribution of alcohol. Large brewers are prohibited from selling beer directly to retail outlets. Instead, they typically sell to state-licensed distributors, who, in turn, sell to retailers. Payments along the supply chain cannot include slotting fees, slotting allowances, or other fixed payments between firms.[8] While retail price maintenance is technically illegal in many states, in practice, distributors are often induced to sell at wholesale prices set by brewers (Asker (2016)).

Table 1 summarizes the revenue shares of the major brewers over 2001-2011. In the early years of the sample, Anheuser-Busch, SABMiller, and Molson Coors (domestic brewers) account for 61%-69% of revenue while Grupo Modelo and Heineken (importers) account for another 12%-16% of revenue.[9] Midway through the sample, in June 2008, SABMiller and Molson Coors consolidated their U.S. operations into the MillerCoors joint venture.[10]

There have been two major consolidating events since MillerCoors. First, ABI acquired Grupo Modelo in 2013. The DOJ sued to enjoin the acquisition and obtained a settlement under which the rights to the Grupo Modelo brands in the U.S. transferred to Constellation, at that time a major distributor of wine and liquor. The allegation of DOJ that Modelo constrained the coordinated pricing of ABI and MillerCoors is a focus of this study. Second, ABI acquired SABMiller in 2016. In order to obtain DOJ approval, SABMiller sold its stake in MillerCoors to Molson Coors. The remedy changed the ownership of the Miller and Coors brands, but did not change any product portfolios or production in the industry.

[8] The relevant statutes are the Alcoholic Beverage Control Act and the Federal Alcohol Administration Act, both of which are administered by the Bureau of Alcohol, Tobacco and Firearms (see their 2002 advisory at `https://www.abc.ca.gov/trade/Advisory-SlottingFees.htm`, last accessed November 4, 2014).

[9] We refer to the first three firms as "domestic" because their beer is brewed in the United States.

[10] The DOJ elected not to challenge on the basis that cost savings in distribution likely would offset any loss of competition. Subsequent academic research suggests that sizable costs savings were realized but were dominated by adverse competitive effects (Ashenfelter et al. (2015), Miller and Weinberg (2017)).

Table 1: Revenue-Based Market Shares

Year	ABI	MillerCoors	Miller	Coors	Modelo	Heineken	Total
2001	0.37	.	0.20	0.12	0.08	0.04	0.81
2003	0.39	.	0.19	0.11	0.08	0.05	0.82
2005	0.36	.	0.19	0.11	0.09	0.05	0.79
2007	0.35	.	0.18	0.11	0.10	0.06	0.80
2009	0.37	0.29	.	.	0.09	0.05	0.80
2011	0.35	0.28	.	.	0.09	0.07	0.79

Notes: The table provides revenue shares over 2001-2011. Firm-specific revenue shares are provided for ABI, Miller, Coors, Modelo, and Heineken. The total across these firms also is provided. The revenue shares incorporate changes in brand ownership during the sample period, including the merger of Anheuser-Busch (AB) and Inbev to form A-B Inbev (ABI), which closed in April 2009, and the acquisition by Heineken of the FEMSA brands in April 2010. All statistics are based on supermarket sales recorded in IRI scanner data.

2.2 Price Leadership in the Beer Industry

The industry appears to be a suitable match for the model. Legal documents filed by the DOJ to enjoin the ABI/Modelo acquisition allege price leadership behavior:

> ABI and MillerCoors typically announce annual price increases in late summer for execution in early fall. In most local markets, ABI is the market share leader and issues its price announcement first, purposely making its price increases transparent to the market so its competitors will get in line. In the past several years, MillerCoors has followed ABI's price increases to a significant degree.[11]

Leader/follower behavior during our sample period did not involve Modelo or Heineken. The legal filings state that Modelo adopted a "Momentum Plan" to "grow Modelo's market share by shrinking the price gaps."[12] Drennan et al. (2013), an article written by DOJ economists, notes that "[i]n internal strategy documents, ABI has repeatedly complained about pressure resulting from price competition with Modelo brands."[13]

In the model, the leader's price announcement serves as an equilibrium selection device, resolving the coordination problem that firms may face due to the folk theorem. The legal documents are helpful in ascertaining whether such a mechanism is consistent with the empirical setting. The following passage quotes from the business documents of ABI:

[11] Para 44 of the Complaint in *US v. Anheuser-Busch InBev SA/NV and Grupo Modelo S.A.B. de C.V.*

[12] Para 49 of the Complaint in *US v. Anheuser-Busch InBev SA/NV and Grupo Modelo S.A.B. de C.V.*

[13] Drennan et al. (2013), p., 295. The legal filings also speak to this. For example, the Competitive Impact Statement (p. 8) states that "[b]y compressing the price gap between high-end and premium brands, Modelo's actions have increasingly limited ABI's ability to lead beer prices higher." The legal filings do not address Heineken specifically, though their prices are similar to Modelo's in the data we examine.

> ABI's Conduct Plan emphasizes the importance of being "Transparent – so competitors can clearly see the plan;" "Simple – so competitors can understand the plan;" "Consistent – so competitors can predict the plan;" and "Targeted – consider competition's structure." By pursuing these goals, ABI seeks to "dictate consistent and transparent competitive response."[14]

Our interpretation of this passage is that the primary purpose of ABI's price announcements is to provide strategic clarity for MillerCoors. If this interpretation is correct then there is a tight connection between price announcements in the beer industry and in our model.

2.3 Prices

Figure 1 shows the time path of average retail prices over 2001-2011 for each firm's most popular 12 pack: Bud Light, Miller Lite, Coors Light, Corona Extra, and Heineken. The red vertical line at June 2008 marks the closing of the Miller/Coors merger. As shown, the prices of domestic beers increase starkly after the merger, while import prices continue on trend. Notably, the price increases of ABI are commensurate with those of MillerCoors. Miller and Weinberg (2017) estimates a post-merger conduct parameter and determines that the data are difficult to explain as a shift from one Bertrand equilibrium to another. We make progress in this paper by examining the data within the context of a fully-specified repeated game. As we develop, the data shown in the figure are entirely consistent with shift from a Bertrand equilibrium to a price leadership equilibrium with binding IC constraints. We test and reject the possibility that IC constraints are non-binding.

2.4 Data

We use retail scanner data from the IRI Academic Database (Bronnenberg et al. (2008)), which contains weekly revenue and unit sales by UPC code for a sample of stores over 2001-2011. We restrict attention to supermarkets, which account for 20% of off-premise beer sales (McClain (2012)).[15] We aggregate the data to the product-region-period-year level, where products are brand×size combinations. We consider alternative period definitions—months and quarters—to provide some robustness to sales and consumer stockpiling behavior. We focus on 13 flagship brands sold as six packs, 12 packs, 24 packs, and 30 packs. We measure quantities based on 144-ounce equivalent units, the size of a 12-pack, and measure price as

[14] Para 46 of the Complaint in *US v. Anheuser-Busch InBev SA/NV and Grupo Modelo S.A.B. de C.V.*

[15] The other major sources of off-premise beer sales are liquor stores (38%), convenience stores (26%), mass retailers (6%), and drugstores (3%). The price and quantity patterns that we observe for supermarkets also exist for drug stores, which are in the IRI Academic Database.

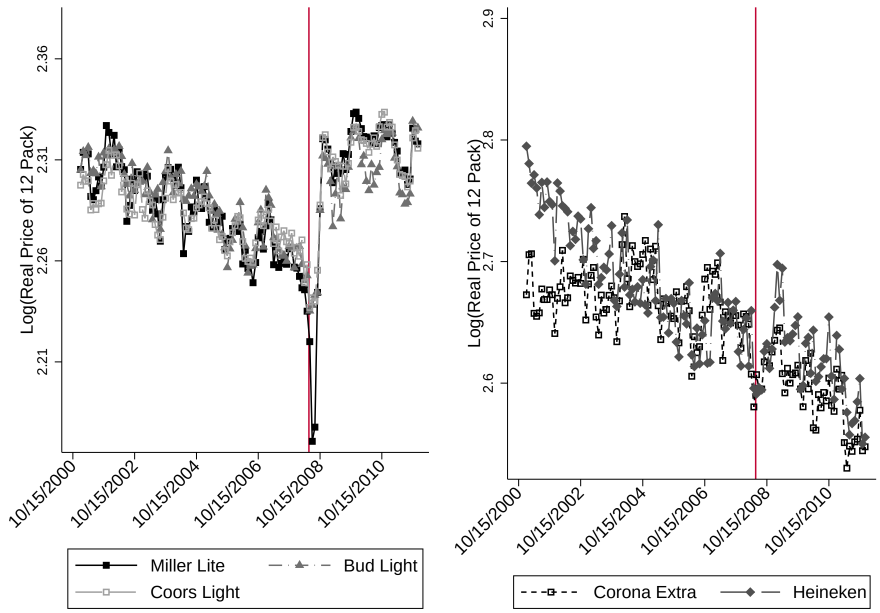

Figure 1: Average Retail Prices of Flagship Brand 12-Packs

Notes: The figure plots the national average price of a 12-pack over 2001-2011, separately for Bud Light, Miller Lite, Coors Light, Corona Extra and Heineken. The vertical axis is the natural log of the price in real 2010 dollars. The vertical bar drawn at June 2008 signifies the consummation of the Miller/Coors merger. Reproduced from Miller and Weinberg (2017).

the ratio of revenue to equivalent unit sales. Table 2 provides summary statistics. The final sample comports with that of Miller and Weinberg (2017).

2.5 Demand

We rely on the random coefficient nested logit (RCNL) model of Miller and Weinberg (2017) to characterize consumer demand. Details of the model are contained in Appendix B. Appendix Table D.1 presents results from the four main specifications. The first two (RCNL-1 and RCNL-2) allow income to affect the price parameter, thereby relaxing cross-price elasticities between more affordable domestic beers and the more expensive imported beers. The latter two (RCNL-3 and RCNL-4) allow income to affect tastes for imported beers directly. The coefficients are precisely estimated and intuitive. The median own price elasticities range from -4.45 to -6.10. The price elasticities of market demand are much smaller, ranging from -0.60 to -0.72, due to the magnitude of the nesting parameter. Most substitution occurs among the inside goods, rather than between the inside goods and the outside good. We provide additional summary statistics on product-level and firm-level elasticities

Table 2: Prices and Conditional Volume Shares in 2011

	6 Packs		12 Packs		24 Packs		All
Brand	Share	Price	Share	Price	Share	Price	Share
Bud Light	0.019	11.62	0.066	10.05	0.180	8.16	0.266
Budweiser	0.011	11.6	0.029	10.04	0.070	8.15	0.109
Coors	0.001	11.61	0.004	10.07	0.011	8.05	0.016
Coors Light	0.010	11.58	0.039	10.07	0.105	8.11	0.155
Corona Extra	0.010	15.82	0.043	13.01	0.024	12.43	0.077
Corona Light	0.006	15.67	0.020	13.05	0.003	12.42	0.028
Heineken	0.007	16.14	0.032	13.33	0.012	12.48	0.051
Heineken Light	0.002	16.21	0.008	13.38	0.001	11.91	0.011
Michelob	0.002	12.45	0.005	10.84	0.009	7.69	0.016
Michelob Light	0.007	12.55	0.023	10.87	0.020	8.68	0.050
Miller Gen. Draft	0.003	11.60	0.007	10.05	0.011	8.12	0.021
Miller High Life	0.004	9.12	0.020	7.91	0.026	6.71	0.050
Miller Lite	0.008	11.55	0.042	10.08	0.101	8.11	0.151

Notes: This table provides the conditional volume share and average price for each brand–size combination in the year 2011. The conditional volume shares sum to one. Prices are per 144 ounces (the size of a 12 pack).

in Appendix Tables D.2 and D.3.[16]

3 Model of Price Leadership

3.1 Primitives

We now develop the model of oligopoly price leadership. Let there be $f = 1, \dots, F$ firms and $j = 1, \dots, J$ differentiated products. Each firm f produces a subset $\mathbb{J}_f$ of all products. Without loss of generality, we assign firm 1 the role of "leader." In many markets, including the U.S. beer market, the pricing leader appears to be the largest firm, though some counter-examples exist (e.g., see Stigler (1947)). Here we take the identity of the leader as exogenously determined and focus on the subsequent price competition.

The game features $t = 0, \dots, \infty$ periods. At the beginning of the game, $t = 0$, the leader designates a set of firms, $\mathbb{C}$, as the coalition. The leader is always in the coalition. Other firms in the coalition are "followers," and firms outside the coalition are "fringe firms."

[16] The parameters are estimated with GMM. The general approach follows the standard nested fixed-point algorithm (Berry et al. (1995)), albeit with a slight modification to ensure a contraction mapping in the presence of the nested logit structure (Grigolon and Verboven (2014)). As demand estimation is not the primary focus of this paper, we refer readers to Miller and Weinberg (2017) for the details of implementation, a discussion of the identifying assumptions, specification tests, and a number of robustness analyses.

In each subsequent period, $t = 1, \ldots, \infty$, an economic state Ψ_t is realized and observed by all firms. Competition then plays out in two stages:

(i) The leader announces a non-binding supermarkup, $m_t \geq 0$, above Nash-Bertrand prices (to be defined), given history h_t (also to be defined).

(ii) All firms set prices simultaneously, given the announced supermarkup m_t and history h_t, and receive payoffs.

The timing of the game mimics a common practice in which one firm announces a price change before the new price becomes available to consumers.[17] However, the first stage is not a theoretical necessity. The price leadership equilibrium (defined later) can be obtained in a standard repeated pricing game with an assumption on equilibrium selection.

Payoffs are determined by continuous and differentiable profit functions and a fixed cost that coalition firms incur by adopting the supermarkup. The profit function of firm f in period $t = 1 \ldots, \infty$ is given by

$$\sum_{j \in \mathbb{J}_f} \pi_j(p_t, \Psi_t) = \sum_{j \in \mathbb{J}_f} (p_{jt} - mc_j(W_t)) q_j(p_t, X_t) \tag{1}$$

where $mc_j(W_t)$ and $q_j(p_t, X_t)$ are a constant marginal cost function and a demand function, respectively, with $(W_t, X_t) \in \Psi_t$ and p_t being a vector of all prices realized in the second stage. Any firm that maximizes its own profit in the second stage given competitors' prices solves the system of first order conditions

$$p_{ft} + \left(\frac{\partial q_f(p_t, X_t)}{\partial p_f}^T \right)^{-1} q_f(p_t, X_t) = mc_f(W_t) \tag{2}$$

where we apply the f subscript to refer to vectors of firm f's prices, quantities, and marginal costs. We assume the first order conditions generate a unique solution.[18] Coalition firms that adopt the supermarkup incur a fixed cost, $R(m_t)$, with $R(0) = 0$ and $R'(m) \geq 0$, which we motivate as arising from antitrust risk. We discuss micro-foundations in Section 5.3.

We assume the cost and demand functions are common knowledge and that all firms observe prices and quantities each period. Different assumptions regarding the evolution of economic states are possible. In this section, we rely on the assumption that Ψ_t is stochastic

[17]Not all leadership/follower behavior has this feature (e.g., Byrne and de Roos (2019)).
[18]The assumption can be verified under nested logit demand (Mizuno (2003)).

and iid across periods, yielding the history

$$h_t = \left((p_{k,\tau}, q_{k,\tau})_{k=1,\ldots,J,\tau=1,\ldots t}\,, (m_\tau)_{\tau=1}^{t-1}, (\Psi_\tau)_{\tau=1}^{t} \right).$$

This treatment of the economic states is theoretically appealing because it avoids certain scenarios in which price leadership unravels due to an adverse realization of Ψ_t.[19] As will be developed, deviation from the leader's proposed supermarkup does not occur on the equilibrium path because the leader adjusts the supermarkup to satisfy incentive compatibility constraints. Finally, we assume that firm actions do not affect the economic states.

3.2 Equilibrium

In this section we formally define the *price leadership equilibrium* (PLE), which is a subgame perfect equilibrium (SPE). Taking as given the coalition structure initially for notational simplicity, the leader's strategy is $\sigma_1 : \mathbb{H} \rightarrow \mathcal{M} \times \mathcal{R}^{J_1}$, where $\mathbb{H}$ is the set of histories, $\mathcal{M}$ is the set of possible supermarkups, and J_1 is the number of products controlled by the leader. The strategies of firms $f = 2, \ldots, F$ are $\sigma_f : \mathcal{M} \times \mathbb{H} \rightarrow \mathcal{R}^{J_f}$. We obtain the strategies that constitute the PLE, starting with the pricing stages, continuing with the announcement stages, and then finishing with the coalition selection at ($t = 0$). We then discuss the equilibrium and describe some of its characteristics.

Consider the pricing stage in some arbitrary period t. Each coalition firm $f \in \mathbb{C}$ "accepts" the leader's proposed supermarkup m_t if it prices according to $p_{ft}^{PL}(m_t; \Psi_t) = p_{ft}^{NB}(\Psi_t) + m_t$. Fringe firms accept simply by pricing on their best response functions. Thus, let $p_{ft}^{PL}(m_t; \Psi_t)$ for $f \notin \mathbb{C}$ solve the first order conditions of equation (2), taking as given the coalition prices and the prices of other fringe firms. Firms "reject" m_t if they select some other price. Given the beliefs to be enumerated below, two particular forms of rejection are relevant. First, let the vector $p_t^{D,f}(m_t; \Psi_t)$ collect the prices that arise if firm f solves equation (2) with the anticipation that other firms accept. Second, let the vector $p_t^{NB}(\Psi_t)$ collect the Bertrand prices that solve equation (2) for all firms. We refer to $p_t^{D,f}(\cdot)$ and $p_t^{NB}(\cdot)$ as deviation and Bertrand prices, respectively.

Let the *slack function* capture the present value of price leadership less the present value of deviation, under the assumption that deviation is punished in all future periods

[19] In the empirical implementation, we instead assume that firms know the entire sequence $(\Psi_\tau)_{\tau=1}^{\infty}$, which avoids having to specify a data generating process for the multi-dimensional economic state. This alternative assumption is plausible in the U.S. beer industry because demand and cost conditions are relatively stable.

with Bertrand prices. For a coalition firm, this difference can be expressed

$$
\begin{aligned}
g_{ft}(m_t;\Psi_t) \;=\; & \overbrace{\frac{\delta}{1-\delta}E_{\Psi}\left[\sum_{j\in\mathbb{J}_f}\pi_j^{PL}(\Psi) - R^*(\Psi) - \sum_{j\in\mathbb{J}_f}\pi_j^{NB}(\Psi)\right]}^{\text{Expected Future Net Benefit of Price Leadership}} \\
& - \underbrace{\left[\sum_{j\in\mathbb{J}_f}\pi_{jt}\left(p_t^{D,f}(m_t,\Psi_t);\Psi_t\right) - \sum_{j\in\mathbb{J}_f}\pi_{jt}\left(p_t^{PL}(m_t,\Psi_t);\Psi_t\right) + R(m_t)\right]}_{\text{Immediate Net Benefit of Deviation}}
\end{aligned}
\tag{3}
$$

where $\delta \in (0,1)$ denotes a common discount factor, $\pi^{NB}(\Psi) \equiv \pi(p^{NB}(\Psi);\Psi)$ is the profit from Bertrand, $\pi^{PL}(\Psi) \equiv \pi\left(p^{PL}(m^*(\Psi),\Psi);\Psi\right)$ is price leadership profit evaluated at $m^*(\Psi)$, defined below as the leader's optimal supermarkup, and $R^*(\Psi) \equiv R(m^*(\Psi))$. The slack functions of fringe firms do not include the antitrust risk terms but otherwise are identical. The slack functions can take positive or negative values for coalition firms, depending on m_t and Ψ_t, but are weakly positive for fringe firms by construction.

In the PLE, the inequalities $g_{ft}(m_t;\Psi_t) \geq 0$ play the role of the incentive compatibility (IC) constraints. As the history is common knowledge, so are the slack functions. We assume firms have the following beliefs: (i) other firms will accept m_t if $g_{ft}(m_t;\Psi_t) \geq 0$ for all f and if all firms have accepted in all previous periods; (ii) other firms will punish if $g_{ft}(m_t;\Psi_t) < 0$ for any f or if any firm has rejected in any previous period.

We can now state the strategies that constitute the equilibrium of the pricing subgame. In each period $t = 1,\ldots,\infty$, all firms price according to $p_t^{PL}(m_t;\Psi_t)$ if $g_{ft}(m_t;\Psi_t,\delta) \geq 0$ for all f and if there has been no previous rejection; otherwise firms price according to $p_t^{NB}(\Psi)$. It is easily verified that there is no profitable departure from these strategies given beliefs, and that beliefs are consistent with the strategies. Deviation prices are never realized in the equilibrium of the pricing subgame. The reason is that if any firm prefers deviation outcomes over price leadership outcomes, given the supermarkup m_t, then this is known by all firms and play shifts immediately to Bertrand prices.

Turning to the announcement stage of some period t, we assume the leader selects a supermarkup under the belief that firms play these equilibrium strategies of the price subgame. As actions do not affect the evolution of the economic state, the optimal supermarkup solves a constrained maximization problem:

$$\begin{aligned} m_t^*(\Psi_t) &= \arg\max_{m\geq 0} \sum_{j\in\mathbb{J}_1} \pi_{jt}\left(p_t^{PL}(m,\Psi_t);\Psi_t\right) - R(m) \\ &s.t. \quad g_{ft}(m;\Psi_t) \geq 0 \quad \forall f \in \mathbb{C} \end{aligned} \tag{4}$$

Our formulation of the leader's problem in the announcement stage accounts for the response of the fringe to the supermarkup because the vector $p_t^{PL}(m,\Psi_t)$ is defined as including best-response prices of fringe firms. A solution always exists because the slack functions equal zero at $m_t = 0$.[20] It follows that punishment never occurs on the equilibrium path because the leader can always find some supermarkup that satisfies IC of coalition firms, even if this implies Bertrand prices ($m_t = 0$) for some realizations of the economic state.

Finishing, in the coalition selection stage ($t = 0$), the leader selects the coalition that maximizes the present value of its payoffs, under the belief of equilibrium play in subsequent periods. In numerical experiments, we have confirmed that partial coalitions can be optimal for the leader. Typically this occurs if there is substantial heterogeneity in the slack functions, which can allow for higher supermarkups with a partial coalition as IC constraints are relaxed. However, heterogeneity is not necessary for partial coalitions generally (e.g., as in d'Aspremont et al. (1983), Donsimoni et al. (1986), and Bos and Harrington (2010)).

Positive supermarkups are not guaranteed. To help frame the empirical analysis, we provide a set of existence results:

Definition (Positive Profit Potential): *Coalition $\mathbb{C}$ has "positive profit potential" if, for all firms $f \in \mathbb{C}$, the following holds:*

$$E_\Psi\left[\sum_{j\in\mathbb{J}_f} \pi_j^{PL}(\Psi) - R^*(\Psi) - \sum_{j\in\mathbb{J}_f} \pi_j^{NB}(\Psi)\right] > 0$$

Proposition 1 (Incentive Compatibility): *Let the coalition $\mathbb{C}$ have positive profit potential. Consider an arbitrary $m_t > 0$. There exists some $\tilde{\delta}(m_t) \in (0,1)$ such that if $\delta > \tilde{\delta}(m_t)$ then $g_{ft}(m_t;\Psi_t) \geq 0$ for all $f \in \mathbb{C}$. Furthermore, for any $\delta \in (0,1)$, if antitrust risk is zero for all supermarkups, then there exists some $\tilde{m}(\delta) > 0$ such that $g_{ft}(\tilde{m}(\delta);\Psi_t) \geq 0$ for all $f \in \mathbb{C}$.*

[20] The solution is unique if the maximand is globally concave, which depends in part on second derivatives of the form $\left(\frac{\partial^2 \pi_j}{\partial p_j \partial p_k}\right)$ for $j \neq k$, as the leader takes into account that changing m affects all prices. To the extent multiple solutions exist, we assume a commonly-understood selection rule exists such that the slack functions can be evaluated. The empirical implementation does not require uniqueness.

Proof: See Appendix A.

The first part of the proposition is standard: if the coalition has future value (i.e., if it has positive profit potential) then any positive supermarkup satisfies IC in the pricing stage if firms are sufficiently patient. The second part states that, in the absence of antitrust risk, there exists a strictly positive supermarkup that satisfies IC. Thus, antitrust risk creates the theoretical possibility that some markets cannot support positive supermarkups. Our second proposition examines equilibrium supermarkups. The leader of a coalition with positive profit potential selects positive supermarkups for at least some realizations of the economic state, and for all realizations if there is no antitrust risk. Formally,

Proposition 2 (Positive Supermarkups): *Let the coalition $\mathbb{C}$ have positive profit potential. Then there exists some Ψ_t such that $m_t^*(\Psi_t) > 0$. If, in addition, antitrust risk is zero for all supermarkups, then $m_t^*(\Psi_t) > 0$ for every Ψ_t.*

Proof: See Appendix A.

3.3 Discussion

The price leadership model closely resembles the canonical Rotemberg and Saloner (1986) model of collusion. Because information is perfect and the supermarkup adjusts with the economic state, deviation does not occur along the equilibrium path. The main departure relates to equilibrium selection: the leader's price announcement selects an equilibrium because, by assumption, it determines firm beliefs. The conditions under which it is reasonable to assume cheap talk—such as the price announcement—affects beliefs have been debated in the literature (e.g., Aumann (1990), Farrell and Rabin (1996)).[21] In support of our approach, recent experimental evidence suggests price announcements can help facilitate coordination in repeated oligopoly games (Harrington et al. (2016)). Interestingly, the PLE is not generally Pareto optimal for the coalition firms because the leader acts in its own interest and side-payments are not incorporated.[22]

We develop a numerical example to provide graphical intuition. Consider a market with logit demand and three differentiated firms, all of which are in the coalition. The first

[21]In our model, the announcement is "self-committing" because the leader has no incentive to deviate from a perfect equilibrium. It is not "self-signaling" because the leader would prefer the followers to accept the supermarkup even if it plans to deviate. Farrell and Rabin (1996) state that "a message that is both self-signaling and self-committing seems highly credible" yet point to an experimental literature to support that cheap talk can be effective in shaping beliefs even if not self-signaling.

[22]See Asker (2010) and Asker et al. (2019) for two empirical examples of inefficient coordination.

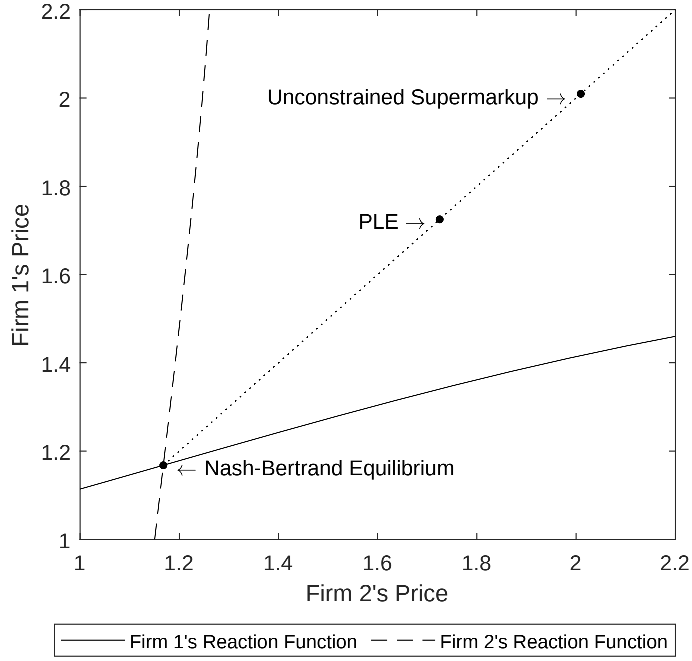

Figure 2: Illustration of the Price Leadership Equilibrium

and second firms have higher quality and lower marginal cost than the third firm. Figure 2 illustrates how price leadership can be interpreted as an equilibrium selection device. The Bertrand equilibrium is identifiable as the intersection of the firms' reaction functions. In selecting the supermarkup, leader considers symmetric price increases above Bertrand equilibrium, plotted as the 45-degree line extending upward from the Bertrand equilibrium. The supermarkup that maximizes the leader's profit (the "Unconstrained Supermarkup") violates IC, so the PLE features a smaller supermarkup of 0.56.[23]

Figure 3 plots the corresponding slack functions of the leader (Panel A) and the smaller follower (Panel B). The slack functions are positive for small enough supermarkups, and negative for larger supermarkups. The function for the smaller follower crosses zero at the PLE supermarkup of 0.56, marked in both panels by the vertical blue line. As the slack function for the other firms is positive at this point, it is the IC of the smaller follower that constrains equilibrium prices. The higher supermarkups preferred by the leader would not be accepted because the smaller follower would deviate.

[23]Demand is $q_i = \frac{\exp(\beta_i - \alpha p_i)}{1+\sum_{k=1}^{3}\exp(\beta_k - \alpha p_k)}$, for $i = 1, 2, 3$, with the parameterizations $\beta_1 = \beta_2 = 3$, $\beta_3 = 1$, and $\alpha = 1.5$. Marginal costs are $mc_1 = mc_2 = 0$ and $mc_3 = 1.25$, and the discount factor is $\delta = 0.4$. Firm 3's price is held fixed at the Bertrand level in constructing the reaction functions shown in Figure 2.

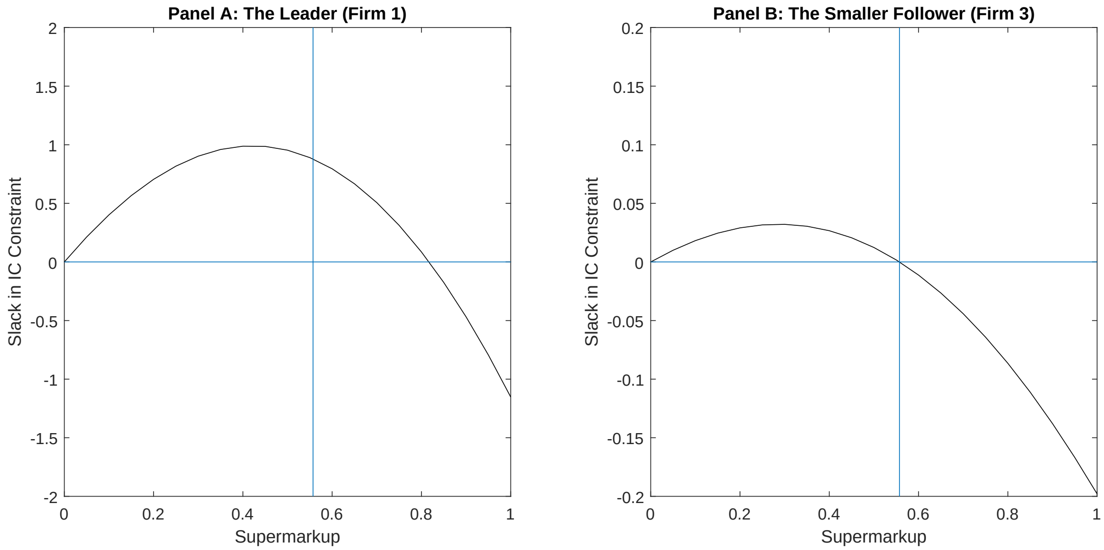

Figure 3: Slack Functions in the Numerical Illustration

Notes: The figure provides the slack functions for the leader (Panel A) and one of the followers (Panel B) with supermarkups $m \in [0, 1]$. IC is satisfied for supermarkup m if the slack functions are positive (i.e., above the horizontal blue line). The vertical blue line shows the equilibrium supermarkup of 0.56.

We have maintained certain timing assumptions that simplify the theoretical analysis. It is reasonable to wonder whether managers would implement grim trigger strategies in real-world settings. Relatedly, a period defines the length of time over which a firm could earn deviation profit before punishment ensues, and it might not be clear in practice whether this corresponds to a month, year, or some other interval. However, our model ends up being equivalent to alternatives with finite punishment or different durations of deviation profit, provided the discount factor is treated as a reduced-form parameter that summarizes both the patience of firms and the timing of the game (Appendix A.2).[24]

4 Empirical Implementation

In this section, we discuss the conditions under which the supermarkups can be estimated with data on prices and quantities. The estimation procedure tracks standard industrial organization methodologies: for any candidate set of supermarkups, one can recover marginal costs, isolate a residual from the cost function, and evaluate a loss function by interacting the residual with instruments taken from the demand-side of the model. Estimation does

[24]This equivalence is recognized in Rotemberg and Saloner (1986), which argues that infinite punishment with a low discount factor is isomorphic to finite punishment with a high discount factor.

not require an evaluation of IC. Nonetheless, with the supermarkups in hand, one can test whether IC binds. In the affirmative case, it also is possible to jointly identify the discount factor and the antitrust risk, a matter to which we return in Section 5.3.

4.1 Identification of Marginal Costs

The identification strategy is a variant on the standard methodology of inferring marginal costs from the Bertrand first order conditions, as introduced in Rosse (1970). To illustrate, we stack equation (2) for each firm and evaluate at Bertrand prices, which obtains the familiar solution that marginal revenue equals marginal cost:

$$mr_t(p_t^{NB}, X_t, \Omega_t) \equiv p_t^{NB} + \left[\Omega_t \circ \left(\left. \frac{\partial q_t(p_t, X_t)}{\partial p_t} \right|_{p=p_t^{NB}} \right)^T \right]^{-1} q_t(p_t^{NB}, X_t) = mc_t(W_t) \quad (5)$$

where the operation $\circ$ is element-by-element multiplication and $\Omega_t \in \Psi_t$ is a matrix that summarizes ownership structure; each of its (j,k) elements equal one if products j and k are produced by the same firm and zero otherwise.

In settings which feature Bertrand competition, equation (5) allows marginal costs to be recovered given knowledge of demand and data on prices. Our application is more complicated. As competition may not be Bertrand, observed prices (p_t) may not correspond with Bertrand prices (p_t^{NB}). It follows that equation (5) cannot be evaluated directly. Nonetheless, if the econometrician has knowledge of the supermarkup, then Bertrand prices and marginal costs can be recovered. We state this result as a proposition:

Proposition 3 (Identification). *Suppose the econometrician has knowledge of the demand system, the identities of the coalition firms (i.e., $\mathbb{C}$), and the supermarkup (m). Then Bertrand prices and marginal costs are identified.*

> **Proof:** The proof is constructive and proceeds in four steps, each of which is easily verified given the maintained assumptions. We enumerate the steps here as they are central to the estimation procedure. Suppressing region and period subscripts, the steps are:
>
> 1. Infer mc_j for each fringe firm $j \notin \mathbb{C}$ from the first order conditions of equation (2). This can be done with observed prices because fringe firms maximize per-period profit.

2. Obtain $p_k^{NB} = p_k - m$ for each coalition firm $k \in \mathbb{C}$.
3. Compute p_j^{NB} for each fringe firm $j \notin \mathbb{C}$ by simultaneously solving the first order conditions of equation (2), given the inferred marginal costs mc_j and holding the prices of coalition firms fixed at the Bertrand level (i.e., $p_k = p_k^{NB}$ for each $k \in \mathbb{C}$).
4. Infer mc_k for each coalition firm $k \in \mathbb{C}$ from the first order conditions of equation (2), evaluated at the already obtained Bertrand prices p^{NB}.

4.2 Specification of Marginal Costs

We parameterize the marginal cost function to complete the model. As we observe variation in the data at the product-region-period, we now introduce subscripts to denote the region. The marginal cost of product j in region r in period t is given by

$$mc_{jrt}(W_{rt}) = w_{jrt}\gamma + \sigma_j^S + \tau_t^S + \mu_r^S + \eta_{jrt} \tag{6}$$

where w_{jrt} includes the distance (miles $\times$ diesel index) between the region and brewery, and two indicators for Miller and Coors products in the post-merger periods, respectively. This specification allows the merger to affect costs through the rationalization of distribution and cost savings unrelated to distance. The unobserved portion of marginal costs depends on the product, period, and region-specific terms, σ_j^S, τ_t^S, and μ_r^S, for which we control using fixed effects, as well as residual costs η_{jrt}, which we leave as a structural error term.

4.3 Estimation

The objects of interest in estimation are $\theta_0 = (m_t, \gamma, \sigma_j^S, \tau_t^S, \mu_r^S)$. For each candidate $\widetilde{\theta}$, one can apply the four steps necessary to recover Bertrand prices and marginal costs (Proposition 3). The implied residuals then obtain:

$$\eta_{jrt}^*(\widetilde{\theta}; \Psi_t) = mr_{jrt}(p_{rt}^{NB}(\widetilde{m}_t; \Psi_t); X_t, \Omega_t) - w_{jrt}\widetilde{\gamma} - \widetilde{\sigma}_j^S - \widetilde{\tau}_t^S - \widetilde{\mu}_r^S \tag{7}$$

Marginal revenue is endogenous because residual costs enter implicitly through Bertrand prices. Valid instruments can be constructed from aspects of the economic state that enter demand (X_t) or ownership (Ω_t) and that satisfy the population moment condition $E[Z' \cdot \eta^*(\theta_0)] = 0$, where $\eta^*(\theta_0)$ is a stacked vector of residuals and Z is the matrix of instruments.

The corresponding generalized method-of-moments estimate is

$$\widehat{\theta} = \arg\min_{\theta} \eta^*(\theta; X, W, \Omega)' Z A Z' \eta^*(\theta; X, W, \Omega) \tag{8}$$

where A is some positive definite weighting matrix. We have exact identification in our application, given instruments that we define below, so A is an identity matrix. We concentrate the fixed effects and the marginal cost parameters out of the optimization problem using OLS to reduce the dimensionality of the nonlinear search.[25]

4.4 Instruments

An important departure from the literature is that the objects of interest in estimation include the supermarkup, which is not a structural parameter but a strategic choice variable that solves a constrained maximization problem. A simple example illustrates the ramifications for identification: Suppose that the econometrician attempts to use a single binary variable, Z_1, taken from the economic state, as the excluded instrument. The model is under-identified because variation in Z_1 implies the existence of two supermarkups that must be estimated. Adding a second instrument, Z_2, does not solve the under-identification problem because any additional variation provided by Z_2 implies the existence yet another supermarkup. Iterating, it follows that no set of instruments is sufficient for identification without additional restrictions on the model.

We make progress by assuming Bertrand pricing ($m_t = 0$) in periods predating the Miller/Coors merger, which resolves the otherwise intractable under-identification problem.[26] The reasonableness of this approach is supported by the available qualitative evidence and an *ex post* analysis of the merger (Appendix C.1). With the restriction in place, we rely on an instrument that equals one for ABI brands after the Miller/Coors merger and zero otherwise. Thus, identification exploits that different candidate supermarkups imply different Bertrand prices for ABI, and thus different post-merger marginal costs (see Appendix Figure D.1 for an illustration). Given the marginal cost specification, the instrument is valid

[25]The third step required to recover marginal costs and Bertrand prices requires that best response fringe prices be computed numerically. With many candidate parameter values, our equation solver does not find a solution for Boston (where the data coverage appears thin) and San Francisco. We therefore exclude these regions from the main regression samples. This does not appear to materially affect results.

[26]The under-identification problem connects to a debate about the identification of conduct parameters. In general, conduct may vary with demand conditions, so the under-identification problem extends. Indeed, it can be interpreted as a version of the famous Corts (1999) critique. A number of articles sidestep the problem by seeking to identify changes in conduct (e.g., Porter (1983); Ciliberto and Williams (2014); Igami (2015); Miller and Weinberg (2017)) using assumptions on conduct in some markets, similar to our approach.

if the average residual costs of ABI do not change contemporaneously with the Miller/Coors merger, relative to the average residual costs of the fringe firms.

The ABI post-merger instrument is sufficient to identify a single supermarkup, and indeed our main results are developed under the assumption that the coalition sets the same supermarkup in every post-merger period and region. Alternatively, it is possible to estimate region-specific or period-specific supermarkups by interacting the ABI post-merger instrument with region or period fixed effects, respectively, so as to maintain exact identification.[27] Doing so does not materially affect our conclusions, however, so we focus on the simpler model. Appendix C.2 provides results for a time-varying supermarkup.

5 Econometric Results

5.1 Estimates

Table 3 summarizes our supply-side estimates. Each column corresponds to one of the baseline demand specifications (see Appendix Table D.1). The marginal cost functions incorporate product, period, and region fixed effects in all cases. The estimates of the supermarkup range from \$0.596 to \$0.738. In our counterfactual analyses, we focus particularly on the RCNL-2 specification, which is somewhat computationally less demanding because periods are quarters, rather than months. The supermarkup we estimate with RCNL-2 is equivalent to about six percent of the average price of a 12 pack.

We estimate that the marginal cost intercepts of Miller and Coors decrease with the joint venture by \$0.53 and \$0.83, respectively, in the RCNL-2 specification. As the distance estimate is positive, a second source of efficiencies from Miller/Coors arises as production of Coors brands and, to a lesser extent Miller brands, is moved to breweries closer to retail locations. Miller and Weinberg (2017) estimate similar marginal cost parameters, and we refer reader to that article for a more in depth analysis of the merger efficiencies. See also Appendix C.3, where we provide an explicit comparison of results.

With the marginal cost estimates in hand, we use counterfactual simulations to recover the unconstrained supermarkups that would maximize the profit of ABI. That is, we solve the optimization problem of equation (4) under the assumption that slack functions do not bind. The solutions range from \$2.57 to \$3.25 across the four demand specifications.

[27] In principle, one could estimate a supermarkup for every region-period combination. The asymptotic properties of the estimator then are unclear, however, as Armstrong (2016) shows consistency may not obtain as the number of products grows large within a fixed set of markets.

Table 3: Baseline Supply Estimates

	Parameter	RCNL-1	RCNL-2	RCNL-3	RCNL-4
Estimation Results					
Supermarkup	m	0.643	0.596	0.738	0.709
		(0.025)	(0.027)	(0.034)	(0.033)
Miller×Post-Merger	γ_1	-0.540	-0.533	-0.583	-0.416
		(0.007)	(0.007)	(0.005)	(0.002)
Coors×Post-Merger	γ_2	-0.826	-0.831	-0.914	-0.666
		(0.009)	(0.009)	(0.006)	(0.004)
Distance	γ_3	0.168	0.164	0.172	0.153
		(0.001)	(0.001)	(0.001)	(0.001)
Supplementary Results					
Unconstrained Supermarkup		2.69	2.57	3.25	2.56
		[2.64,2.77]	[2.49, 2.66]	[3.18, 3.31]	[2.48,2.63]
Negative Marginal Costs		0.12%	0.09%	0.26%	0.03%
Welfare Effects of Price Leadership					
% Δ Profit		10.68	8.57	10.90	14.42
Δ Consumer Surplus / Δ Profit		3.73	3.93	3.90	3.88

Notes: The table shows the baseline supply results. Estimation is with the method-of-moments. There are 89,619 observations at the brand-size-region-month-year level (RCNL-1 and RCNL-3) and 30,078 observations at the brand-size-region-quarter-year level (RCNL-2 and RCNL-4). The samples excludes the months/quarters between June 2008 and May 2009. Regression includes product (brand×size), period (month or quarter), and region fixed effects. The unconstrained supermarkup is obtained using a post-estimation simulation. The welfare statistics are computed for the periods from June 2009 to December 2011. Standard errors are clustered by region and shown in parentheses. Bootstrapped 95% confidence intervals, shown in brackets, are provided for the unconstrained supermarkups.

Bootstrapped confidence intervals easily exclude the point estimates of the supermarkup. As the unconstrained supermarkups greatly exceed the estimated supermarkups, we interpret the results as indicating that at least one IC constraint binds in the PLE.

Finally, we report statistics on how price leadership affects firms and consumers, relative to counterfactual Bertrand prices, which we recover with counterfactual simulations. We find that price leadership increases profit by 8.57%–14.42% across the four specifications. The amount that consumer surplus decreases is almost four times greater than the amount that profit increases, as consumers pay more and may select less-preferred brands in response to higher prices.[28]

[28] Consumer surplus is the inclusive value of all consumer options, including the outside good. This value is identified up to a constant, which cancels out when considering a change in consumer surplus.

Table 4: Brewer Markups

Brand	6 Packs Pre	6 Packs Post	12 Packs Pre	12 Packs Post	24 Packs Pre	24 Packs Post
Bud Light	3.82	4.52	3.69	4.39	3.59	4.25
Budweiser	3.98	4.68	3.82	4.53	3.69	4.37
Coors	2.86	4.54	2.71	4.45	2.58	4.28
Coors Light	2.66	4.38	2.53	4.27	2.43	4.14
Corona Extra	3.59	3.43	3.28	3.11	3.18	3.18
Corona Light	3.33	3.14	3.00	2.88	3.09	3.01
Heineken	3.49	3.42	3.21	3.13	3.34	3.46
Heineken Light	3.21	3.10	2.88	2.75	3.00	2.94
Michelob	3.90	4.70	3.81	4.58	3.48	4.38
Michelob Light	3.83	4.55	3.71	4.40	3.60	4.15
Miller Gen. Draft	3.10	4.43	2.95	4.29	2.85	4.19
Miller High Life	3.09	4.38	2.95	4.29	2.87	4.21
Miller Lite	3.09	4.41	2.95	4.31	2.85	4.17

Notes: This table provides the average markups for each brand–size combination separately for the pre-merger and post-merger periods, based on the RCNL-2 demand specification.

Table 4 provides the average markup for each product in the data both before and after the Miller/Coors merger, based on the RCNL-2 specification. Across all 89,619 brand–size–month–region observations, the average markup is \$3.37 on an equivalent-unit basis, which accounts for 32% of the retail price. The average markups on ABI 12 packs tend to be about \$0.70 higher in the post-merger periods, which reflects the combination of higher Bertrand prices and the supermarkup. The markups on Miller 12 packs increase by about \$1.35 and the markups on Coors products increase by about \$1.75. Those changes reflect the combined impact of higher Bertrand prices, the supermarkup, and lower marginal costs. The markups on imported beers do not change much over the sample period.

5.2 Price Leadership and Deviation

The profit functions under price leadership and deviation, as well as the level of Bertrand profit, are essential inputs to our subsequent analyses. To build intuition, we use counterfactual simulations to examine a series of alternative supermarkups, $\tilde{m} = (0.00, 0.01, \ldots, 3.00)$. For each $\tilde{m}$ we obtain the profit that would be obtained by each firm, under price leadership and deviation. We compare to the profit that would be obtained under Bertrand.

Figure 4 provides results obtained with the RCNL-2 specification. Panel A focuses on ABI. The vertical axis is profit relative to Bertrand and the horizontal axis is the supermarkup. The profit functions take a value of one at $\tilde{m} = 0$ because price leadership is equivalent to Bertrand and there is no profitable deviation. From there, the profit under price leadership increases to its maximum at a supermarkup just over $2.50 (which accords with Table 3), and then decreases. This provides a graphical representation of the maximand in the leader's constrained optimization problem. By contrast, deviation profit increases monotonically in the supermarkup because higher supermarkups correspond to higher MillerCoors prices. If plotted over a much broader support, the deviation profit function would flatten in the supermarkup as the market share of MillerCoors shrinks.

Because the gap between the two profit function grows in the supermarkup, so too does the incentive to deviate. At our point estimate of the supermarkup, which we mark with the vertical blue line, ABI profit is about seven percent higher than Bertrand and deviation profit is about eight percent higher. Thus, deviation does not appear to increase profit much relative to price leadership. One may wonder whether this is a product of the logit-based demand system. To explore, we calibrate an alternative linear demand system that has the same elasticities at observed prices, and find a similar pattern (Appendix C.4).

In Panel B, we explore the price and share functions that contribute to profit functions. Under price leadership, these functions have slopes of quite similar magnitudes and of opposite sign. As the functions are indexed relative to Bertrand, this implies a coalition elasticity of demand around unity. At our point estimate of the supermarkup, ABI prices are about eight percent higher than Bertrand, and shares are about eight percent lower. The deviation price and share functions increase with the supermarkup. The prices of ABI and MillerCoors appear to be strategic complements across a wide support.

Panels C and D show that the statistics for MillerCoors are broadly similar, which reflects that ABI and MillerCoors have similar markups and firm elasticities in the post-merger periods (e.g., Table 4 and Appendix Table D.3).

5.3 Calibrating the Slack Functions

We make three modifications to the slack functions before bringing them to the data. First, we replace the assumption of a stochastic economic state with an assumption that the entire sequence $(\Psi_\tau)_{\tau=1}^{\infty}$ is common knowledge in every period. This raises the theoretical possibility that price leadership could unravel if positive supermarkups cannot be sustained beyond some future date, as in Igami and Sugaya (2019). However, unraveling does not occur in our application by construction, as we model the future using infinite repetitions of the year

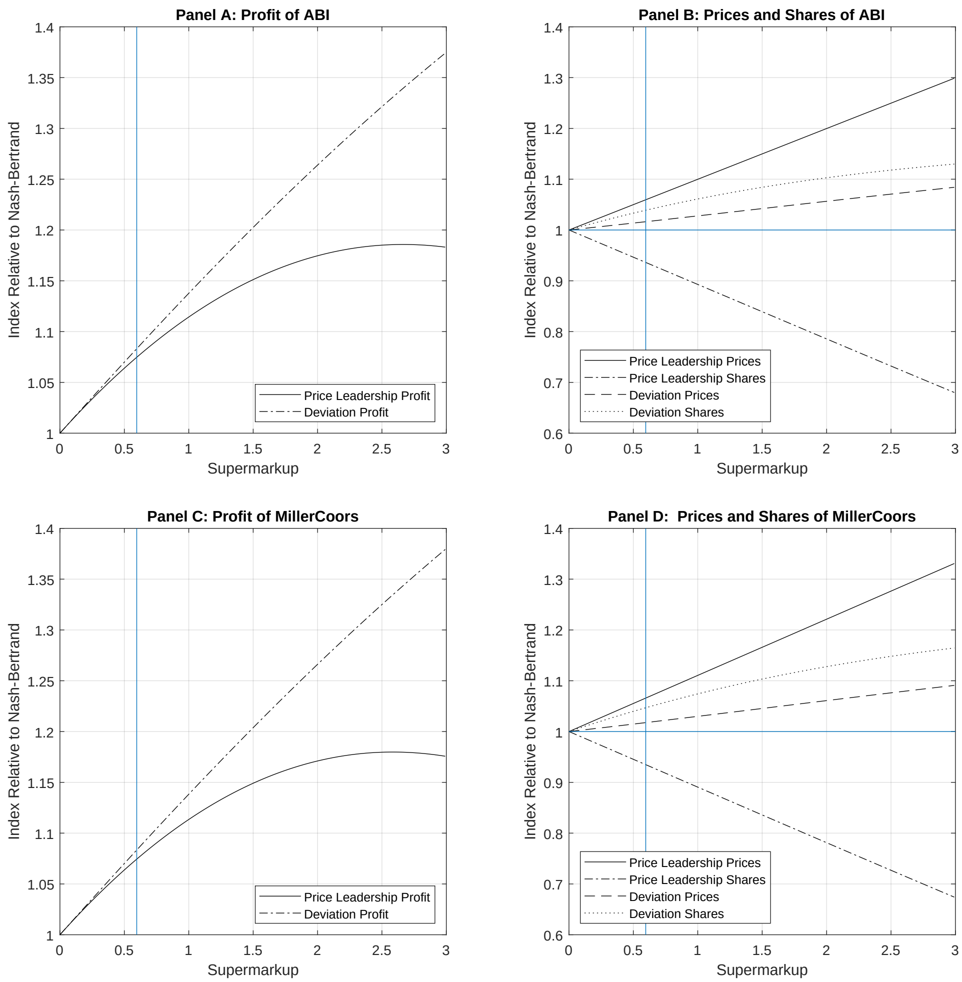

Figure 4: Profit, Prices and Shares with Price Leadership and Deviation

Notes: The figure provides the profit (Panel A and C) and average price and market share (Panels B and D) for ABI (Panels A and B) and MillerCoors (Panels C and D) in 2011:Q4 under price leadership and deviation. Statistics are computed for a range of supermarkups ($m \in [0, 3]$). All statistics are reported relative to their Bertrand analog. The vertical line marks the supermarkup estimated from the data. Results are based on the RCNL-2 demand specification.

2011.[29] Second, we assume that deviation profit is earned for a full calendar year before punishment ensues, which we motivate based on the observed practice of annual list price adjustments. We discuss timing assumptions below. Finally, we sum the functions across regions, creating a single IC constraint for each coalition firm.[30]

[29] Our approach accommodates constant percentage growth or decay in market size (Appendix A.2), provided that the discount factor is treated as a reduced-form statistic.

[30] Implicitly this assumes that a deviation in any regions triggers punishment in all regions. If regions are heterogeneous then pooling IC may loosen constraints (Bernheim and Whinston (1990)).

Among the objects in the slack functions, the profit terms are easily recovered via counterfactual simulations given knowledge of $(\Psi_\tau)_{\tau=1}^{\infty}$, leaving the discount factor and the antitrust risk as the only unknowns (see equation (3)). Antitrust risk plays an important role in the model because it creates the theoretical possibility that some market structures cannot support positive supermarkups. There are a variety of reasons that tacit coordination may impose explicit or implicit costs on firms, but one interpretation is legal risk. For instance, evidence of price leadership has been considered in a number of price-fixing lawsuits when courts have weighed whether discovery should be granted to the plaintiffs.[31] Further, historical evidence of pricing coordination sometimes is cited by antitrust authorities as contributing to a decision to challenge a merger.[32]

We apply a simple parameterization, $R(m_t; \phi) = \phi m_t$, that captures these influences in a simple reduced-form manner. We refer to ϕ as the risk coefficient. The econometric tests of Section 5 reject the null hypothesis that slack exists in both the ABI and MillerCoors IC constraints. Therefore we assume that least at one IC constraint binds. With one equation and two unknowns, the parameters (δ, ϕ) are jointly identified.

Figure 5 plots the values that balance the MillerCoors IC constraint in 2011:Q4. With $\phi = 0$, an annualized discount factor of 0.11 balances IC, and greater values of ϕ require higher discount factors. We attempt to remain agnostic about what constitutes an economically reasonable discount factor. The reason is that the IC constraints incorporate timing assumptions about deviation and punishment that are impossible to verify as they are off the equilibrium path (and therefore not observed in the data). Thus, recalling the discussion in Section 3.3, we interpret the discount factor as a reduced-form parameter that summarizes both the patience of firms and the timing of the game.[33]

Figure 6 plots the slack in IC of ABI (Panel A) and MillerCoors (Panel B) over the range of supermarkups $m \in [0, 0.8]$. Four alternative assumptions are used to calibrate the

[31] Examples include firms involved in flat glass (*Re: Flat Glass Antitrust Litig.*, 385 F.3d 350 (3rd Cir 2004)), text messaging (*Re: Text Messaging Antitrust Litig.*, 782 F.3d 867 (7th Cir 2015)), titanium dioxide (*Re: Titanium Dioxide Antitrust Litig.*, RDB-10-0318 (D. Md. 2013)), and chocolate (*Re: Chocolate Confectionary Antitrust Litig.*, 801 F.3d 383 (3rd Cir 2015)).

[32] Interestingly, a prime example is ABI's attempted acquisition of Modelo in 2012-2013, which the DOJ challenged in part due to a concern it would eliminate a constraint on coordinated price increases. We return to the economic effects of the proposed ABI/Modelo merger in Section 6. A second example is the Tronox/Cristal merger in the titanium dioxide industry (*Re: Fed. Trade Comm'n v. Tronox Ltd.*, Case No. 1:18-cv-01622 (TNM)(D.D.C. 2018)).

[33] In our application, with $\delta = 0.9$ and $\phi = 0$, about three months of punishment are sufficient to ensure incentive compatibility. That such a brief punishment period is required can be attributed to the results shown in Figure 4: the gap between price leadership and Bertrand per-period profit is much larger than the gap between deviation and price leadership per-period profit.

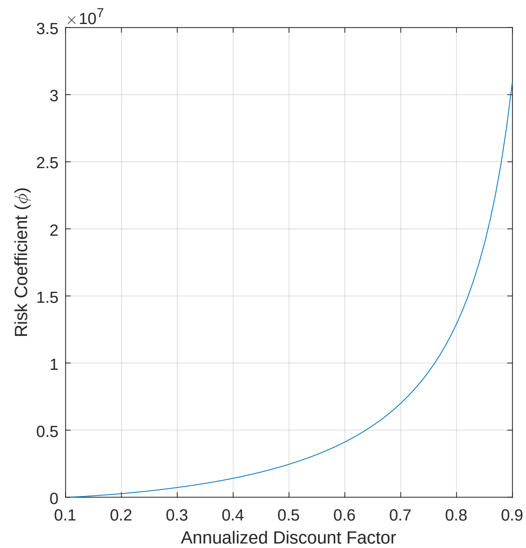

Figure 5: Joint Identification of Antitrust Risk and the Discount Factor

Notes: The figure shows the combinations risk coefficients (ϕ) and annualized discount factors (δ^*) for which the MillerCoors IC constraint binds in 2011:Q4, over the range $\delta^* \in [0.11, 0.90]$. Results are based on the RCNL-2 demand specification.

IC constraints: $\delta = 0.7$, $\delta = 0.5$, $\delta = 0.3$, and $\phi = 0$. In each case, we select the free parameter such that IC of MillerCoors binds at the estimated supermarkup of 0.596. We consider a number of candidate supermarkups, $m = 0.00, 0.01, 0.02, \ldots$, and for each we use counterfactual simulations to obtain profit with price leadership, deviation, and punishment. Pairing this with the calibrated (δ, ϕ) parameters, we recover firm-specific slack functions. The figure shows that slack exists in the IC constraints for any supermarkup less than 0.596. MillerCoors would prefer to deviate for any higher supermarkup. ABI, by contrast, still has slack in its IC constraint at $m = 0.596$. Thus we conclude that MillerCoors constrains coalition pricing in the observed equilibrium.[34]

[34]Readers may wonder why a higher discount factor is associated with less slack for some supermarkups, on the basis that increasing the discount factor unambiguously loosens IC constraints in the model, *ceterus parabis*. Here not all else is equal—a higher discount factor requires a greater risk coefficient to balance IC.

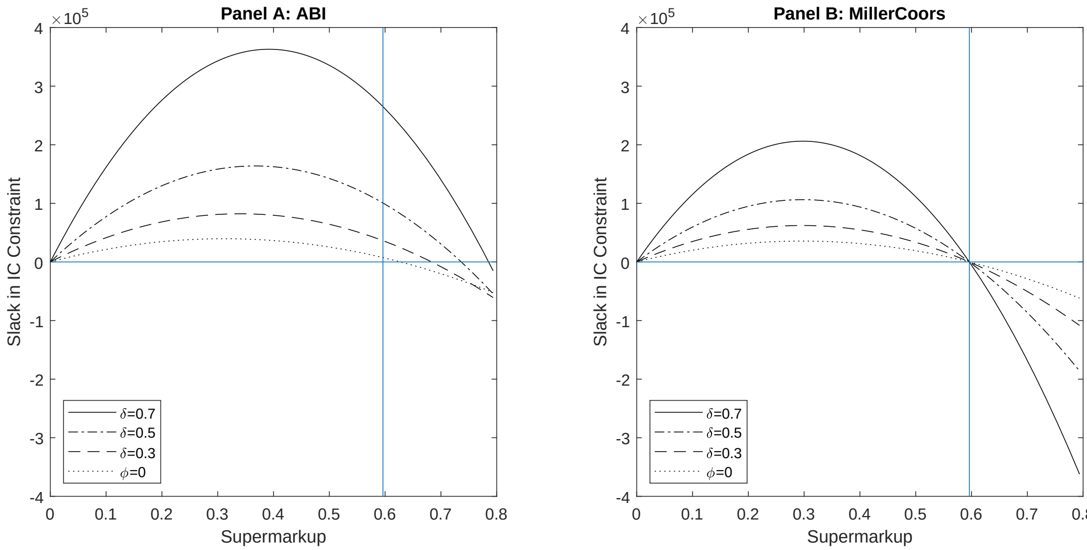

Figure 6: Slack Functions Given the Observed Market Structure

Notes: The figure provides the slack functions in 2011:Q4 for ABI (Panel A) and MillerCoors (Panel B) and with supermarkups $m \in [0, 0.8]$. IC is satisfied for supermarkup m if the slack functions are positive (i.e., above the horizontal blue line). The vertical line shows the estimated supermarkup of 0.596. We use four different balancing assumptions: $\delta = (0.7, 0.5, 0.3)$ and $\phi = 0$. The balancing assumptions ensure that the slack functions cross zero for one firm at the estimated supermarkup. Results are based on the RCNL-2 demand specification.

6 The ABI/Modelo Merger

6.1 Background

On June 28, 2012, ABI agreed to acquire Grupo Modelo for about $20 billion. The acquisition was reviewed by the DOJ, which sued in January 2013 to enjoin the acquisition.[35] Prior to trial the merging firms and the DOJ reached a settlement under which Modelo's entire U.S. business was divested to Constellation Brands, a major distributor of wine and liquor.[36] In its Complaint, the DOJ alleged that Modelo constrained the prices of ABI and MillerCoors:

> ABI and MillerCoors often find it more profitable to follow each other's prices than to compete aggressively.... In contrast, Modelo has resisted ABI-led price

[35] ABI held a 35% stake in Grupo Modelo prior to the acquisition. However, in an annual report, ABI stated that it did "not have voting or other effective control of... Grupo Modelo," consistent with the empirical and documentary evidence presented in Section 2.3. See Para 19 of the Complaint in *US v. Anheuser-Busch InBev SA/NV and Grupo Modelo S.A.B. de C.V.*

[36] The press release of the DOJ provides details on the settlement. See `https://www.justice.gov/opa/pr/justice-department-reaches-settlement-anheuser-busch-inbev-and-grupo-modelo-beer-case`, last accessed February 13, 2019.

> hikes.... If ABI were to acquire the remainder of Modelo, this competitive constraint on ABI's and MillerCoors' ability to raise their prices would be eliminated.[37]

We analyze the ABI/Modelo merger in this section using the price leadership model. We assume that the Modelo products would have been priced by ABI, that is, we model the merger as it would have occurred without the divestiture. We focus on the year 2011 because it is the period nearest to the acquisition date covered in our data.

6.2 Merger Simulation

Figure 7 graphs the new slack functions of ABI (Panel A) and MillerCoors (Panel B).[38] The vertical blue line marks $m = 0.596$, the supermarkup we estimate without the ABI/Modelo merger. Evaluated at that point, slack exists in all the IC constraints we consider. Thus, higher supermarkups can be sustained in the PLE after the ABI/Modelo merger. The new equilibrium supermarkup can be located visually as the crossing of the MillerCoors slack function with the horizontal blue line. We refer to the change in the supermarkup as the *coordinated effect* of the merger. Different calibrations of (δ, ϕ) produce coordinated effects of different magnitudes, though all are positive. Recalling that $p_t = p_t^{NB} + m$ for coalition firms, the total change in price also reflects a shift in the Bertrand equilibrium. We refer to the change in Bertrand prices as the *unilateral effect* of the merger.

Table 5 provides greater detail on the unilateral ("Δ Bertrand Price") and coordinated ("Δ Supermarkup") effects. Panel A shows that the Bertrand prices of ABI and Modelo brands increase by \$0.29 and \$1.76 on average, with the magnitude of the latter reflecting a strong incentive to steer customers toward higher-markup ABI brands. Prices also increase due to a higher supermarkup. For ABI and MillerCoors the magnitude of this change ranges from \$0.21 to \$1.01 across the calibrations selected for (δ, γ). For Modelo the change also reflects an adoption of the initial supermarkup of 0.596. The total changes in price ("Total Δ Price") equal the sum of these effects for the coalition firms. The average market share of Modelo brands decreases by more than 50% in all of the specifications we consider.[39] Overall,

[37] Paras 3-5 of the Complaint in *US v. Anheuser-Busch InBev SA/NV and Grupo Modelo S.A.B. de C.V.*

[38] We construct these slack functions numerically by evaluating candidate supermarkups $m = 0.00, 0.01, 0.02, \ldots$. For each, we obtain the profit of each firm with price leadership, deviation, and punishment. We then plug into the slack function for each of the calibrated (δ, ϕ) combinations.

[39] The results for Heineken are interesting. Its Bertrand prices increase by \$0.01, reflecting a small degree of strategic complementarity in prices. However, it responds to the (large) supermarkups in the post-merger PLE by *lowering* its price somewhat. Given the demand specification we employ, consumers that reduce purchases of ABI/Modelo in response to higher prices tend to be more price elastic. For some ranges of price this rotates Heineken's residual demand curve sufficiently to make its price a strategic substitute.

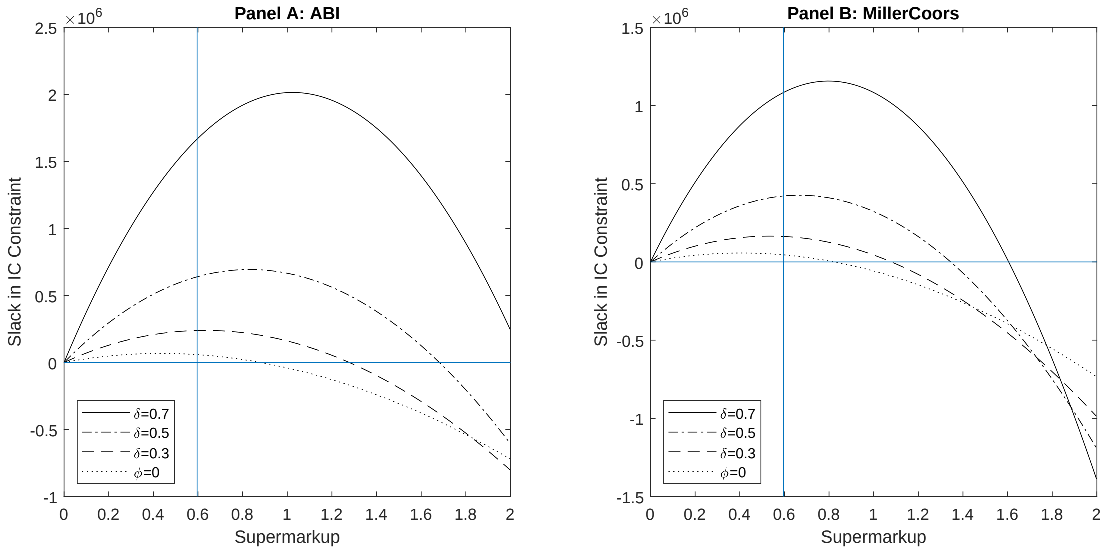

Figure 7: Slack Functions with an ABI/Modelo Merger

Notes: The figure provides the slack functions in 2011:Q4 IC constraint for ABI (Panel A) and MillerCoors (Panel B) and with supermarkups $m \in [0, 0.8]$. IC is satisfied for supermarkup m if the slack functions are positive (i.e., above the horizontal blue line). The vertical blue line shows the estimated supermarkup of 0.596. The slack functions are generated with four different balancing assumptions: $\delta = (0.7, 0.5, 0.3)$ and $\phi = 0$. Results are based on the RCNL-2 demand specification.

the results support the DOJ allegations that Modelo constrains coordinated pricing.

Panel B provides profit and welfare statistics. The increase in the joint profit of ABI and Modelo range from 13.53% to 5.95% across the selected calibrations of the slack function. The consumer surplus effects range from -5.38% to -2.64%. Recalling that price leadership reduces consumer surplus by 1.74% relative to Bertrand absent the ABI/Modelo merger (Table 3), our simulation results highlight that the economic consequences of price leadership can depend greatly on the ownership structure of the industry. Finally, the total surplus effects of the ABI/Modelo merger range from -4.14% to -2.02%.

6.3 Decomposition of the Slack Functions

We next decompose each firm's slack function into components governing the future benefit of continuing with price leadership and the immediate gain from deviating. We then build an understanding of how the ABI/Modelo merger changes incentives by computing how each component changes with the merger. With the modifications we make to the theoretical slack function in order to bring it to the data (Section 5.3), the empirical slack function of

Table 5: Economic Effects of the ABI/Modelo Merger

Panel A: Price and Quantity				
	ABI	MillerCoors	Modelo	Heineken
Δ Bertrand Prices	0.29	0.11	1.76	0.01
Δ Supermarkup				
$\delta = 0.7$	1.01	1.01	1.60	0.00
$\delta = 0.5$	0.73	0.73	1.33	0.00
$\delta = 0.3$	0.47	0.47	1.07	0.00
$\phi = 0.0$	0.21	0.21	0.81	0.00
Total Δ Price				
$\delta = 0.7$	1.30	1.12	3.36	-0.08
$\delta = 0.5$	1.02	0.85	3.09	-0.07
$\delta = 0.3$	0.77	0.59	2.83	-0.06
$\phi = 0.0$	0.51	0.33	2.58	-0.04
% Δ Market Share				
$\delta = 0.7$	-10.03	-4.17	-53.66	47.01
$\delta = 0.5$	-7.66	-1.59	-52.63	35.81
$\delta = 0.3$	-5.46	-0.82	-51.68	26.12
$\phi = 0.0$	-3.25	3.23	-50.73	17.08

Panel B: Profit and Welfare				
	$\delta = 0.7$	$\delta = 0.5$	$\delta = 0.3$	$\phi = 0.0$
% Δ ABI/Modelo Profit	13.53	11.12	8.65	5.95
% Δ Consumer Surplus	-5.38	-4.43	-3.54	-2.64
% Δ Total Surplus	-4.14	-3.40	-2.71	-2.02

Notes: The table shows unweighted averages for the prices and market shares and sums for profit. Based on the RCNL-2 demand specification.

firm f in period τ can be written as:

$$\begin{aligned} g_{f\tau}(m) \quad = \quad & \underbrace{\sum_{s\geq\tau+4}\sum_{r}\sum_{j\in\mathcal{J}_f}\delta^{s-1}\pi^{PL}_{jrs}(m)}_{\text{Price Leadership Continuation Value}} - \underbrace{\sum_{s\geq\tau+4}\sum_{r}\sum_{j\in\mathcal{J}_f}\delta^{s-1}\pi^{NB}_{jrs}}_{\text{Punishment Continuation Value}} \\ & - \underbrace{\sum_{s=\tau}^{\tau+3}\sum_{r}\sum_{j\in\mathcal{J}_f}\delta^{s-1}\pi_{jrs}(p^{D,f}_{rs}(m))}_{\text{Immediate Deviation Gain}} + \underbrace{\sum_{s=\tau}^{\tau+3}\sum_{r}\sum_{j\in\mathcal{J}_f}\delta^{s-1}\pi^{PL}_{jrs}(m)}_{\text{Immediate Price Leadership Gain}} - \underbrace{\frac{\phi m}{1-\delta}}_{\text{Antitrust Risk}} \end{aligned}$$

Table 6: Decomposition of Slack Function Before and After ABI/Modelo

	No Merger $m = 0.596$	Merger $m = 0.596$	Change	Merger $m = 1.60$	Change
	(i)	(ii)	(iii)=(ii)-(i)	(iv)	(v)=(iv)-(ii)
Price Leadership Continuation Value					
ABI	6.59	8.22	1.63	9.01	0.79
MillerCoors	6.24	6.77	0.53	7.48	0.71
Punishment Continuation Value					
ABI	6.13	7.61	1.48	7.61	0
MillerCoors	5.80	6.23	0.43	6.23	0
Immediate Deviation Gain					
ABI	2.85	3.54	0.69	4.00	0.46
MillerCoors	2.70	2.92	0.22	3.34	0.42
Immediate Price Leadership Gain					
ABI	2.83	3.52	0.69	3.86	0.34
MillerCoors	2.67	2.90	0.23	3.21	0.31

Notes: The table shows the results of the decomposition exercise. Units are millions of dollars. Slack functions are computed for 2011:Q4 using a discount factor of 0.7. The no merger scenario uses a supermarkup of 0.596. The merger scenario uses a supermarkup of 1.60 and no efficiencies. Based on the RCNL-2 specification.

The first and second terms are the continuation values of price leadership and punishment, respectively. Together, they yield the net continuation value of price leadership. The third and fourth terms are the immediate gains from deviation and price leadership.

Table 6 evaluates each component of the slack function and shows how they change with the counterfactual ABI/Modelo merger. Columns (i)-(iii) focus on the estimated supermarkup of 0.596 and can be used to understand the vertical shift in the slack functions plotted in Figure 7.[40] As shown, the ABI/Modelo merger increases the continuation value of punishment because competition is softer in Bertrand equilibrium. However, the continuation value of price leadership also increases due to the higher Modelo prices, and this second effect dominates. Thus, in our setting, the vertical upward shifts in the slack functions arise because the continuation value of coordination increases more than the punishment continuation value, evaluated at the pre-merger supermarkup. In contrast, the relative immediate gains from deviation versus coordination remain basically stable.

Columns (iv) and (v) focus on the post-merger equilibrium supermarkup. A key finding is that the ABI/Modelo merger would substantially increase the net continuation value

[40] The four terms shown do not combine to zero in either columns (i) or (ii) because the antitrust risk is not included. As these columns impose the same supermarkup, there is no change in antitrust risk.

of price leadership both for ABI and, importantly, for MillerCoors, the firm constraining the supermarkup. The increase in the supermarkup accounts for over half of the increase in the continuation value of price leadership profits for MillerCoors. This illustrates an important departure from previous attempts to model the coordinated effects of mergers, where firms set prices that maximize joint profits in the collusive state (Davidson and Deneckere (1984); Davis and Huse (2010)). In that case the merger would do nothing to the value of collusion across the merging firms, but would increase the value of deviating by raising static Nash profits, implying that mergers *reduce* incentives to coordinate. In contrast, our model allows the leader to adjust the supermarkup after a merger, taking into account the new IC constraints. This flexibility can increase the continuation value of coordination, creating the possibility of coordinated effects.

6.4 Incorporating Efficiencies

In our final analysis, we explore the economic effects of merger efficiencies under price leadership. To provide a comparison, we also obtain results under the assumption that competition is Bertrand in all periods, both before and after the ABI/Modelo merger. We consider a "minor" efficiency in which the marginal costs of Modelo decrease by \$0.50. We also consider a "major" efficiency in which the marginal costs of ABI and Modelo decrease such that the change in Bertrand prices due to the merger is exactly zero.[41]

Table 7 summarizes the results. Columns (i)-(iii) provide simulation results under Bertrand, and represent what might be obtained from a standard unilateral effects analysis of the merger. Without efficiencies, ABI and Modelo prices increase by \$0.34 and \$1.70, on average. Adding minor efficiencies, the Modelo price increase falls to \$1.15, and surplus loss is partially mitigated. With major efficiencies, there are no price changes because the cost reductions exactly offset the upward pricing pressure (Werden (1996); Farrell and Shapiro (2010); Jaffe and Weyl (2013)). Total surplus increases due to the lower marginal costs. Overall, the results in the first three columns are consistent with a tradeoff between upward pricing pressure and lower marginal cost that is standard in unilateral effects analysis.

Columns (iv)-(vi) show the results under price leadership ($\delta = 0.7$). The change in Bertrand prices are similar and reflect the established tradeoff.[42] More striking is that

[41]To implement the Bertrand simulation, we follow the standard procedure of imputing marginal costs from equation (5), and then finding post-merger prices that satisfy the first order conditions of equation (2). The major efficiency is a multi-product version of the compensating marginal cost reductions derived in Werden (1996). On average, we reduce ABI costs by \$0.51 and Modelo costs by \$1.72.

[42]The analysis in columns (i)-(iii) assumes that observed prices are generated by Bertrand competition,

Table 7: Efficiencies under Price Leadership and Bertrand

Equilibrium Assumption:	Bertrand			PLE with $\delta = 0.7$		
Efficiencies:	None	Minor	Major	None	Minor	Major
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Δ Bertrand Price						
ABI	0.34	0.36	0.00	0.29	0.31	-0.06
MillerCoors	0.13	0.12	0.00	0.11	0.10	-0.01
Modelo	1.70	1.15	0.00	1.76	1.21	0.06
Heineken	0.01	0.00	0.00	0.01	0.00	0.01
Δ Supermarkup	-	-	-	1.01	1.01	1.03
Δ Total Price						
ABI	0.34	0.36	0.00	1.30	1.31	0.96
MillerCoors	0.13	0.12	0.00	1.12	1.11	1.01
Modelo	1.70	1.15	0.00	3.36	2.81	1.68
Heineken	0.01	0.00	0.00	-0.08	-0.09	-0.09
% Δ Profit						
ABI	5.63	4.23	14.51	16.23	14.91	25.87
MillerCoors	8.56	7.55	0.00	20.01	19.27	12.70
Modelo	-0.53	13.76	46.58	0.46	14.79	45.79
Heineken	13.3	10.91	0.00	44.32	41.95	28.91
% Δ Consumer Surplus	-1.64	-1.36	0.00	-5.38	-5.12	-3.88
% Δ Total Surplus	-1.25	-0.99	0.52	-4.14	-3.88	-2.48

Notes: The table shows unweighted averages for the total prices, and percentage changes in firm profit (i.e., profit summed across products and regions). Based on the RCNL-2 specification.

the supermarkup increases by approximately the same amount (1.01, 1.01, 1.03) across the three efficiency scenarios. Efficiencies do not appear to offset coordinated effects. This occurs because the MillerCoors slack function constrains coalition pricing. Thus, the marginal costs of ABI and Modelo affect the supermarkup only through the MillerCoors slack function.

Implicit in these results is that cost pass-through in models of static Nash competition and models of constrained coordination are fundamentally different. To tease out some intuition, notice that the supermarkup actually *increases* slightly as the marginal costs of ABI/Modelo decrease. The effect of the ABI/Modelo cost reductions is through the

whereas the analysis in columns (iv)-(vi) assumes that observed prices are generated with price leadership. Thus, the changes in Bertrand prices are not identical.

MillerCoors slack function. With efficiencies, the profit that MillerCoors would receive in the event of punishment (i.e., Bertrand profit) is lower. This softens the binding IC constraint and allows for higher supermarkups to be supported in equilibrium.

The changes in total price reflect both the change in Bertrand price and the change in the supermarkup. In theory, then, sufficiently large efficiencies could decrease Bertrand prices by enough to offset the increase in the supermarkup. However, even the major efficiencies we consider do not come close to doing so. Across all the scenarios, the merger has greater adverse price and surplus effects under price leadership than under Bertrand.

7 Conclusion

There is a longstanding concern that horizontal mergers may facilitate or exacerbate tacit collusion. However, the empirical industrial organization literature to date has provided little in the way of methodologies to model coordinated effects in real-world markets. Two related obstacles in particular have hindered progress. First, the multiplicity of equilibria that often exist in repeated pricing games (e.g., Friedman (1971); Abreu (1988)) may frustrate predictions. Second, it can be difficult to understand firm strategies in repeated games, or more broadly to have confidence in the structure of the game itself.

We analyze a particular repeated pricing game—oligopolistic price leadership—in which these obstacles appear somewhat less daunting. Strategies along the equilibrium path are easily modeled as leader/follower interactions. Further, as the leader solves a simple constrained maximization problem, basic regularity conditions ensure a unique equilibrium. We show how the model can be estimated with aggregate scanner data and provide an empirical application to the beer industry. We use the merger of ABI and Modelo to illustrate that the framework can be used to model coordinated effects in real-world settings. We are also able to quantify the welfare effects of oligopolistic price leadership, which is of independent interest given the attention the pricing practice has received in the recent literature.

Despite the advantageous features of the price leadership model, some strong assumptions are necessary nonetheless. In our view, perhaps most vexing is that empirical inferences about the duration and severity of punishment are unavailable because deviation and punishment do not occur along the equilibrium path (a standard feature of collusion games with perfect information). Yet some inference about punishment is needed to conduct counterfactual analyses because any analysis of incentive compatibility depends on the full stream of profit obtained in the event of deviation and punishment.

Presented with this dilemma, we interpret the discount factor as a reduced-form statistic. This has the advantage of allowing us to remain agnostic about punishment duration. As a reduced-form statistic, the discount factor reflects both valuations of the future and the length of punishment. If one were to unpack these multiple interpretations and focus on punishment length more explicitly, the coalition may be able to relax incentive compatibility constraints with optimal punishments (Abreu (1986)). Further, in many repeated pricing games of imperfect information (e.g., Green and Porter (1984)), punishment is observed along the equilibrium path, potentially allowing for some of these assumptions to be supported with empirical evidence. However, incorporating imperfect information comes with its own set of challenges that we leave to future research.

A related set of questions pertain to whether the duration and severity of punishment responds endogenously to mergers or other market changes. We make the simplest assumption and hold punishment fixed (allowing for changes in static Nash payoffs). An alternative would be to assume optimal punishments, thereby allowing the model to generate an endogenous response. Absent some empirical support, it is unclear which approach better mimics the behavior of real-world firms. Thus, on this point our counterfactuals may be subject to a version of the Lucas (1976) critique. Nonetheless, we view empirical research on repeated pricing games as having great promise, and believe that exploring optimal punishment strategies will only add to the findings on price interactions obtained in this paper.

References

Abreu, D. (1988). On the theory of infinitely repeated games with discounting. *Econometrica 56*(2), 383–96.

Armstrong, T. B. (2016). Large market asymptotics for differentiated product demand estimators with economic models of supply. *Econometrica 85*(5), 1961–1980.

Ashenfelter, O., D. Hosken, and M. C. Weinberg (2015). Efficiencies Brewed: Pricing and Consolidation in U.S. Brewing. *RAND Journal of Economics 46*(2), 328–361.

Asker, J. (2010). A study of the internal organization of a bidding cartel. *The American Economic Review 100*(3), 724–762.

Asker, J. (2016). Diagnosing Foreclosure due to Exclusive Dealing. *Journal of Industrial Economics 64*(3), 375–410.

Asker, J., A. Collard-Wexler, and J. D. Loecker (2019). (Mis)Allocation, market power, and global oil extraction. *American Economic Review 109*(4), 1568–1615.

Aumann, R. (1990). Nash equilibrium are not self-enforcing. In Gabszewicz, J. J., J.-F. Richard, and L. A. Wolsey, eds., *Economic Decision-Making: Games, Econometrics and Optimisation*. Amsterdam: Elsevier, pp. 201-206.

Bain, J. S. (1960). Price leaders, barometers, and kinks. *Journal of Business 33*(3), 193–203.

Baker, J. B. (2001). Mavericks, mergers, and exclusion: Proving coordinated competitive effects under the antitrust laws. *New York University Law Review 77*, 135–203.

Baker, J. B. (2010). Market concentration in the antitrust analysis of horizontal mergers.

Bernheim, B. D. and M. D. Whinston (1990). Multimarket contact and collusive behavior. *RAND Journal of Economics 21*(1), 1–26.

Berry, S., J. Levinsohn, and A. Pakes (1995, July). Automobile prices in market equilibrium. *Econometrica 63*(4), 847–890.

Bos, I. and J. E. Harrington (2010). Endogenous Cartel Formation with Heterogeneous Firms. *RAND Journal of Economics 41*(1), 92–117.

Bresnahan, T. (1987). Competition and collusion in the american automobile oligopoly: The 1955 price war. *Journal of Industrial Economics 35*(4), 457–482.

Bronnenberg, B. J., M. W. Kruger, and C. F. Mela (2008). The IRI marketing data set. *Marketing Science 27*(4), 745–748.

Byrne, D. P. and N. de Roos (2019). Learning to Coordinate: A Study in Retail Gasoline. *American Economic Review 109*(2), 591–619.

Chilet, J. A. (2018). Gradually rebuilding a relationship: The emergence of collusion in retail pharmacies in Chile.

Ciliberto, F. and J. W. Williams (2014). Does Multimarket Contact Facilitate Tacit Collusion? Inference on Conduct Parameters in the Airline Industry. *RAND Journal of Economics 45*(4), 764–791.

Clark, R. and J.-F. Houde (2013, August). Collusion with asymmetric retailers: Evidence from a gasoline price-fixing case. *American Economic Journal: Microeconomics 5*(3), 97–123.

Compte, O., F. Jenny, and P. Rey (2002). Capacity constraints, mergers and collusion. *European Economic Review 46*(1), 1–29.

Corts, K. S. (1999). Conduct parameters and the measurement of market power. *Journal of Econometrics 88*, 227–225.

d'Aspremont, C., A. Jacquemin, J. J. Babszewicz, and J. A. Weymark (1983). On the Stability of Collusive Price Leadership. *Canadian Journal of Economics 26*, 17–25.

Davidson, C. and R. Deneckere (1984). Horizontal mergers and collusive behavior. *International Journal of Industrial Organization 2*(2), 117–132.

Davis, P. J. and C. Huse (2010). Estimating the "coordinated effects" of mergers. *Working Paper*.

Donsimoni, M., N. Economides, and H. Polemarchakis (1986). Stable Cartels. *International Economic Review 27*(2), pp. 317–327.

Drennan, R., M. Magura, and A. Nevo (2013). The year in review: Economics at the Antitrust Division 2012-2013. *Review of Industrial Organization 43*(4), 291–302.

Eizenberg, A. and D. Shilian (2019). Structure conduct, and contact: Competition in closely-related markets. *mimeo*.

Farrell, J. (1987). Cheap talk, coordination, and entry. *RAND Journal of Economics 18*(1), 34–39.

Farrell, J. and M. Rabin (1996). Cheap talk. *Journal of Economic Perspectives 10*(3), 103–118.

Farrell, J. and C. Shapiro (2010). Antitrust evaluation of horizontal mergers: An economic alternative to market definition. *B.E. Journal of Theoretical Economics: Policies and Perspectives 10*(1).

Friedman, J. W. (1971). A non-cooperative equilibrium for supergames. *Review of Economic Studies 38*(1), 1–12.

Gasmi, F., J.-J. Laffont, and Q. Vuong (1992). Econometric analysis of collusive behavior in a soft-drink market. *Journal of Economics & Management Strategy 1*(2), 277–311.

Grigolon, L. and F. Verboven (2014). Nested logit or random coefficient logit? A comparison of alternative discrete choice models of product differentiation. *Review of Economics and Statistics 96*(5), 916–935.

Harrington, J. E. (2013). Evaluating mergers for coordinated effects and the role of "parallel accommodating conduct". *Antitrust Law Journal* (78), 651–668.

Harrington, J. E., R. H. Gonzalez, and P. Kujal (2016). The relative efficacy of price announcements and express communication for collusion: Experimental findings. *Journal of Economic Behavior & Organization 128*, 251–264.

Harrington, J. E. and P. T. Harker (2017). Developing competition law for collusion by autonomous price-setting agents.

Igami, M. (2015). Market Power in International Commodity Trade: The case of coffee. *Journal of Industrial Economics 63*(2), 225–248.

Igami, M. and T. Sugaya (2019). Measuring the incentive to collude: The vitamin cartels, 1990-1999. Available at SSRN: `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2889837`.

Ishibashi, I. (2008). Collusive price leadership with capacity constraints. *International Journal of Industrial Organization 26*(3), 704–715.

Ivaldi, M., B. Jullien, P. Rey, P. Seabright, and J. Tirole (2007). The economics of tacit collusion: Implications for merger control. In *The Political Economy of Antitrust,* Vivek Ghosal, and Johan Stennek (eds.), Elsevier Science, 217-240.

Jaffe, S. and E. G. Weyl (2013). The first order approach to merger analysis. *American Economic Journal: Microeconomics 5*(4), 188–213.

Knittel, C. R. and V. Stango (2003). Price ceilings as focal points for tacit collusion: Evidence from the credit cart market. *American Economic Review 93*(5), 1703–1729.

Lanzillotti, R. F. (2017). Collusion/Competition: A New Learning? *Antitrust Bulletin 62*(3), 591–602.

Lemus, J. and F. Luco (2018). Pricing dynamics and leadership: Evidence from the retail gasoline industry. Available at SSRN: `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3186144`.

Loertscher, S. and L. M. Marx (2019). Coordinated effects. *mimeo*.

Markham, J. W. (1951). The nature and significance of price leadership. *Americal Economic Review 41*(5), 891–905.

Marshall, R. C., L. M. Marx, and M. E. Raiff (2008). Cartel price announcements: The vitamins industry. *International Journal of Industrial Organization*.

McClain, J. (2012). 2011-2012 annual report. Technical report, Beer Institute, Chicago Illinois.

McFadden, D. (1978). Modelling of choice of residential location. In *Spatial Interaction Theory and Residential Location*, F. Snickers, A. Karlquist, L. Lundquist, J. Weibull (eds), Amsterdam: North Holland.

Michel, C. and S. Weiergraeber (2018). Estimating industry conduct in differentiated products markets: The evolution of pricing behavior in the RTE cereal industry. *mimeo*.

Miller, N. H., M. Remer, C. Ryan, and G. Sheu (2016). Pass-through and the prediction of merger price effects. *Journal of Industrial Economics 64*, 684–709.

Miller, N. H. and M. C. Weinberg (2017). Understanding the Price Effects of the MillerCoors Joint Venture. *Econometrica 85*(6), 1763–1791.

Mizuno, T. (2003). On the existence of a unique price equilibrium for models of product differentiation. *Internation Journal of Industrial Organization 21*(6), 761–793.

Mouraviev, I. and P. Rey (2011). Collusion and leadership. *International Journal of Industrial Organization 29*, 705–717.

Nevo, A. (2001). Measuring market power in the ready-to-eat cereal industry. *Econometrica 69*(2), pp. 307–342.

Nicholls, W. H. (1949). The tobacco case of 1946. *Americal Economic Review 39*(3), 284–296.

Oxenfeldt, A. R. (1952). Professor Markham on price leaderhip: Some unanswered questions. *American Economic Review 42*(3), 380–384.

Pastine, I. and T. Pastine (2004). Cost of delay and endogenous price leadership. *International Journal of Industrial Organization 22*(1), 135–145.

Porter, R. (1983). A Study of Cartel Stability: The Joint Executive Committee, 1880-1886. *Bell Journal of Economics 14*(2), 301–314.

Rojas, C. (2008). Price competition in U.S. brewing. *Journal of Industrial Economics 66*(1), 1–31.

Rosse, J. N. (1970). Estimating cost function parameters without using cost data: Illustrated methodology. *Econometrica 38*(2), 256–275.

Rotemberg, J. J. and G. Saloner (1986). A supergame-theoretical model of price wars during booms. *American Economic Review 76*(3), 390–407.

Rotemberg, J. J. and G. Saloner (1990). Collusive price leadership. *Journal of Industrial Economics 39*(1), 93–111.

Schelling, T. (1960). *The Strategy of Conflict*. Boston, MA: Harvard Business School Press.

Scherer, F. (1980). *Industrial Market Structure and Economic Performance*. Houghton Mifflin.

Seaton, J. S. and M. Waterson (2013). Identifying and characterising price leadership in British supermarkets. *International Journal of Industrial Organization 31*, 392–403.

Slade, M. E. (2004). Market power and joint dominance in U.K. brewing. *Journal of Industrial Economics 52*(1), 133–163.

Stigler, G. J. (1947). The kinky demand curve and rigid prices. *Journal of Political Economy 55*(5).

Vasconcelso, H. (2005). Tacit collusion, cost asymmetries, and mergers. *RAND Journal of Economics 36*(1), 39–62.

Werden, G. J. (1996). A robust test for consumer welfare enhancing mergers among sellers of differentiated products. *The Journal of Industrial Economics 44*(4), pp. 409–413.

Werden, G. J. and M. G. Baumann (1986). A simple model of imperfect competition in which four are few but three are not. *Journal of Industrial Economics 34*(3), 331–335.

Appendix for Online Publication

A Theoretical Details

A.1 Proofs

Proof of Proposition 1

The proof of the first part of Proposition 1 is standard. With positive profit potential, the slack function of (3) is strictly increasing in δ for any given $m_t > 0$. By inspection, we have $\lim_{\delta \to 1^-} g_{ft}(m_t) = +\infty$, because the term labeled "Expected Future Net Benefit of Price Leadership" converges to infinity as δ approaches 1 from below, while the term labeled "Immediate Net Benefit of Deviation" is unaffected by the discount factor. Also by inspection, we have $\lim_{\delta \to 0^+} g_{ft}(m_t) < 0$. Thus, for each coalition firm, as δ increases towards 1, there is a threshold $\tilde{\delta}_f(m_t)$ at which point the slack function becomes positive. The maximum of these thresholds gives $\tilde{\delta}(m_t)$.

For the second claim in Proposition 1, let $t_{ft}(m_t) = \left[\sum_{j \in \mathbb{J}_f} \pi^D_{jt}(m_t) - \sum_{j \in \mathbb{J}_f} \pi^{PL}_{jt}(m_t)\right]$. That is, $t_{ft}(m_t)$ is the immediate net benefit of deviation for firm f, under the assumption that antitrust risk is zero. Because the coalition has positive profit potential, we need only show that there exists some $m_t > 0$ such that $t_{ft}(m_t)$ does not outweigh this future value. We have $t_{ft}(0) = 0$ because there is no profitable deviation if the supermarkup is zero. Furthermore, $t_{ft}(\cdot)$ is continuous because the firm profit functions are continuous. Thus, by choosing an appropriate supermarkup $m_t > 0$ in the neighborhood of zero, we can ensure that $t_{ft}(m_t)$ is arbitrarily close to zero. Such a $t_{ft}(m_t)$ does not outweigh the expected future benefits to price leadership for any coalition firm. QED.

Proof of Proposition 2

For the first statement, we employ a simple proof by contradiction. Suppose $m^*_t(\Psi_t) = 0$ for all Ψ_t. Then, regardless of the state, we have $\pi^{PL} = \pi^{NB}$ and $R = 0$, and the coalition does not have positive profit potential. As this is a contradiction, $m^*(\Psi_t) > 0$ for some Ψ_t. For the second statement, we start with the result (Proposition 1) that there exists an arbitrarily small $\tilde{m}(\delta) > 0$ that satisfies incentive compatibility for any $\delta \in (0, 1)$. Thus, it is sufficient to show that the leader's profit at $\tilde{m}(\delta)$ exceeds its profit at $m = 0$. If this is the case then the leader's constrained maximization problem is guaranteed to produce an $m^*_t(\Psi_t) > 0$ for any Ψ_t. We focus on single-product firms without loss of generality. Let the leader be firm

j. We have:

$$\left.\frac{\partial \pi_j(p)}{\partial m}\right|_{p=p^{NB}} = \left.\frac{\partial \pi_j(p)}{\partial p_j}\right|_{p=p^{NB}} + \sum_{k\neq j} \left.\frac{\partial \pi_j(p)}{\partial p_k}\right|_{p=p^{NB}}$$

the first term on the right is zero by the envelop theorem, and the second term is positive because products are substitutes. Thus, a shift from $m = 0$ to an arbitrarily small $\tilde{m}(\delta)$ increases the leader's profit. QED.

A.2 The Discount Factor as a Reduced-Form Parameter

There are at least three reasons that the discount factor as it appears in the empirical slack functions might summarize more than firm patience. First, punishment may (in actuality) be limited in duration. Second, deviation might be detected and punished in less than one year. Third, changes in market size over time are not captured by infinite repetitions of the year 2011. In this appendix, we show that none of these misspecifications are consequential so long as the discount factor is interpreted as a reduced-form parameter.

A.2.1 Punishment Length

We formalize the argument of Rotemberg and Saloner (1986) that an infinite punishment period with a low discount factor is equivalent to a finite punishment period with a high discount factor. For the sake of discussion, assume that coalition, deviation, and punishment profits are constant over time. With grim trigger strategies, the IC constraint takes the form

$$\frac{1}{1-\delta}\pi^{PL} \geq \pi^D + \frac{\delta}{1-\delta}\pi^{NB}, \tag{A.1}$$

with a discount factor of δ. If instead punishment occurs for only n periods, the IC constraint takes the form

$$\sum_{t=0}^{\infty} \eta^t \pi^{PL} \geq \pi^D + \sum_{t=1}^{n} \eta^t \pi^{NB} + \sum_{t=n+1}^{\infty} \eta^t \pi^{PL}, \tag{A.2}$$

with a discount factor of η. Rearranging equation (A.2) and applying rules for geometric series yields

$$\frac{1-\eta^{n+1}}{1-\eta}\pi^{PL} \geq \pi^D + \frac{\eta(1-\eta^n)}{1-\eta}\pi^{NB}. \tag{A.3}$$

By inspection, equations (A.1) and (A.3) are equivalent if and only if

$$\frac{1}{1-\delta}=\frac{1-\eta^{n+1}}{1-\eta} \quad \text{and} \quad \frac{\delta}{1-\delta}=\frac{\eta(1-\eta^{n})}{1-\eta}.$$

These conditions are satisfied for

$$\delta=\frac{\eta(1-\eta^{n})}{1-\eta^{n+1}}. \tag{A.4}$$

Punishment for n periods at a discount factor of η is equivalent to grim trigger punishment at a discount factor of δ, provided equation (A.4) holds. Further, by inspection, for a given η, decreasing n will decrease δ. Thus, a model with a low discount factor and lengthy punishment is equivalent to a model with a high discount factor and short punishment. To provide a sense of magnitudes, in Table A.1 we provide the reduced-form discount factor under infinite Nash reversion that is economically equivalent to a discount factor of 0.90 with finite punishment (of varying lengths).

Table A.1: The Discount Factor and Punishment Length

	Years of Punishment with Discount Factor of 0.90							
	1	2	3	4	5	10	15	20
Equivalent Reduced-Form Discount Factor under Grim Trigger Strategies	0.474	0.631	0.709	0.756	0.787	0.854	0.877	0.888

A.2.2 Speed of Detection and Punishment

Suppose that deviation profits are earned for n periods instead of one. This gives an IC constraint of

$$\sum_{t=0}^{\infty}\eta^{t}\pi^{PL}\geq\sum_{t=0}^{n}\eta^{t}\pi^{D}+\sum_{t=n+1}^{\infty}\eta^{t}\pi^{NB}$$

Applying the rules of geometric series, this reduces to

$$\frac{1}{1-\eta^{n+1}}\pi^{PL}\geq\pi^{D}+\frac{\eta^{n+1}}{1-\eta^{n+1}}\pi^{NB},$$

which means that if $\delta=\eta^{n+1}$, we return to our original expression with one period of deviation profits earned (equation (A.1)). Therefore, if we calibrate a discount factor assuming that deviation profits are earned for one period, but in actuality these profits are earned for $n+1$ periods, the resulting estimate is equal to the true discount factor raised to $n+1$.

Similarly, if we calibrate a discount factor assuming that deviation profits are earned for $n+1$ periods, but in actuality these profits are earned for one period, the resulting estimate is equal to the true discount factor raised to $1/(n+1)$. A higher discount factor and more periods of earning deviation profits can be equivalent to a lower discount factor and fewer periods of earning deviation profits.

A.2.3 Growth and Decay in Market Size

In logit-based demand systems, including the RCNL we employ, the quantity demanded of any good j is determined by a multiplicative product, $q_{jt} = s_{jt}M_t$, where s_{jt} is the good's market share within the market and M_t is the market size. Further, with constant marginal cost, changes in market size do not affect profit-maximizing prices. Thus, the good's contribution to profit takes the form $\pi_{jt} = (p_{jt} - c_{jt})s_{jt}M_t$. Suppose that market size undergoes constant percentage growth or decay. Letting profit at $t = 0$ be given by π^{PL}, π^D, and π^{NB}, the IC constraint takes the form

$$\sum_{t=0}^{\infty} \eta^t (1+r)^t \pi^{PL} \geq \pi^D + \sum_{t=1}^{\infty} \eta^t (1+r)^t \pi^{NB}$$

for growth/decay rate $r \in (-1, 1)$ and discount factor η. Then if we set $\delta = \eta(1+r)$, and provided that the normalcy condition $\eta(1+r) < 1$ holds, we obtain the original IC constraint provided in equation (A.1). Thus, our empirical approach accommodates constant growth or decay in market size, as the reduced-form discount factor scales appropriately.

B The Demand System

Here we sketch the Miller and Weinberg (2017) random coefficients nested logit (RCNL) model of demand. Suppose we observe $r = 1, \ldots, R$ regions over $t = 1, \ldots, T$ time periods. Each consumer i purchases one of the observed products ($j = 1, \ldots, J_{rt}$) or selects the outside option ($j = 0$). The conditional indirect utility that consumer i receives from the inside good j in region r and period t is

$$u_{ijrt} \quad = \quad x_j \beta_i^* - \alpha_i^* p_{jrt} + \sigma_j^D + \tau_t^D + \xi_{jrt} + \bar{\epsilon}_{ijrt} \tag{B.1}$$

where x_j is a vector of observable product characteristics, p_{jrt} is the retail price, σ_j^D is the mean valuation of unobserved product characteristics, τ_t^D is the period-specific mean valuation of unobservables that is common among all inside goods, ξ_{jrt} is a region-period deviation from these means, and $\bar{\epsilon}_{ijrt}$ is a mean-zero stochastic term.

The observable product characteristics include a constant (which equals one for the inside goods), calories, package size, and an indicator for whether the product is imported. The consumer-specific coefficients are $[\alpha_i^*, \beta_i^*]' = [\alpha, \beta]' + \Pi D_i$ where D_i is consumer income. Define two groups, $g = 0, 1$, such that group 1 includes the inside goods and group 0 is the outside good. Then the stochastic term is decomposed according to

$$\bar{\epsilon}_{ijrt} = \zeta_{igrt} + (1-\rho)\epsilon_{ijrt} \tag{B.2}$$

where ϵ_{ijrt} is i.i.d extreme value, ζ_{igrt} has the unique distribution such that $\bar{\epsilon}_{ijrt}$ is extreme value, and ρ is a nesting parameter ($0 \leq \rho < 1$). Larger values of ρ correspond to less substitution between the inside and outside goods. The quantity sold of good j in region r and period t is

$$q_{jrt} = \frac{1}{N_{rt}} \sum_{i=1}^{N_{rt}} \frac{\exp((\delta_{jrt} + \mu_{ijrt})/(1-\rho))}{\exp(I_{igrt}/(1-\rho))} \frac{\exp(I_{igrt})}{\exp(I_{irt})} M_r \tag{B.3}$$

where I_{igrt} and I_{irt} are the McFadden (1978) inclusive values, M_r is the market size of the region, $\delta_{jrt} = x_j\beta + \alpha p_{jrt} + \sigma_j^D + \tau_t^D + \xi_{jrt}$, and $\mu_{ijrt} = [p_{jrt}, x_j]' * \Pi D_i$. The normalization on the mean indirect utility of the outside good yields $I_{i0rt} = 0$. The inclusive value of the inside goods is $I_{i1rt} = (1-\rho)\log\left(\sum_{j=1}^{J_{rt}} \exp((\delta_{jrt} + \mu_{ijrt})/(1-\rho))\right)$ and the inclusive value of all goods is $I_{irt} = \log(1 + \exp(I_{i1rt}))$. We assume market sizes 50% greater than the maximum observed unit sales within each region. Expressions for the price derivatives of demand are supplied in Grigolon and Verboven (2014).

C Additional Analyses

C.1 Nash-Bertrand Competition Before Miller/Coors

To obtain identification, we assume $m_t = 0$ before the Miller/Coors merger (Section 4.4). Our interpretation of the available qualitative evidence is that it supports the reasonableness of this identifying assumption—if price leadership existed prior to the Miller/Coors merger then it probably did not elevate prices much above the Nash-Bertrand equilibrium. Here we discuss the qualitative evidence and then show the Miller/Coors merger plausibly was pivotal in allowing for coordination to be sustained in equilibrium.

Qualitative Evidence

The annual reports of the companies point to intense price competition in the years before the Miller/Coors merger. For example, the 2005 SABMiller annual report describes "intensified competition" and an "extremely competitive environment." The 2005 Anheuser-Busch report states that the company was "collapsing the price umbrella by reducing our price premium relative to major domestic competitors." SABMiller characterizes price competition as "intense" in its 2006 and 2007 reports.[43] A contemporaneous article in the *New York Times* (2009) supports the language of the annual reports and provides context for why price competition may have intense:

> After South African Breweries [SAB] bought Miller in 2002, it set out to take market share from Bud. Its bigger rival responded by slashing prices. The others were then forced to match. This competition fostered a better outcome for consumers—indeed, the summer of 2005 was a beer drinkers' dream.[44]

Lending veracity to the annual reports and the 2009 *New York Times* article, both sources describe the softening of competition after the Miller/Coors merger for which we find econometric support. In its 2009 report, SABMiller attributes increasing earnings before interest, taxes, and amortization expenses to "robust pricing" and "reduced promotions and discounts." In its 2010 and 2011 reports, it references "sustained price increases" and "disciplined revenue management with selected price increases."[45] The *New York Times* article reports "That's all changed. SABMiller and Molson Coors kicked off a joint venture last year that combines the market powers of the second- and third-largest players."

Unwinding MillerCoors

Given that we impose $m_t = 0$ in periods predating the Miller/Coors merger, it would be comforting if our results indicate the merger actually is pivotal in supporting positive supermarkups. This is *not* guaranteed by our identifying assumption. The reason is that Bertrand equilibrium is always an SPE of a repeated pricing game. Thus, $m = 0$ could arise either because (1) positive supermarkups would lead to deviation, or (2) positive supermarkups are sustainable but, for one reason or another, beliefs lead firms to price according to the

[43]See SABMiller's Annual Report of 2005 (p. 13), 2006 (p. 5), 2007 (pp. 4 and 8), and Anheuser-Busch's Annual Report in 2005 (p. 5).

[44]See "Rising Beer Prices Hint at Oligopoly" by Alize Rosenbaum, Rob Cox and Pierre Briancon, *New York Times*, August 27, 2009.

[45]See SABMiller's Annual Report of 2009 (pp. 9 and 24), 2010 (pp. 29), and 2011 (p. 28). ABI's annual reports in the post-merger years are more opaque.

Bertrand equilibrium. The latter possibility has less theoretical appeal because it involves a change in equilibrium selection that occurs for reasons outside the model. Thus, the reasonableness of identifying assumption would be bolstered if results indicate the Miller/Coors merger is necessary for sustainable coordination.

We conduct counterfactuals to explore this question. In particular, we unwind the joint venture by assigning the Miller and Coors brands to separate firms and applying the pre-merger cost structure.[46] Figure C.1 plots the results for the calibrations that use $\delta = 0.7$ (Panel A), $\delta = 0.5$ (Panel B), $\delta = 0.3$ (Panel C), and $\gamma = 0$ (Panel D). In the first three panels, any ABI/Miller/Coors coalition is unsustainable. In Panels A and B, the IC of both Miller and Coors is violated for any positive supermarkup. In Panel C, the Coors IC is violated for any positive supermarkup. In Panel D, by contrast, all IC constraints are satisfied for $m \leq 0.48$ and coordination is sustainable at that level.

Considered together, the results we obtain indicate the Miller/Coors merger is indeed pivotal for coordination, for most parameterizations of (δ, ϕ). The transition from Bertrand to price leadership can be explained without invoking forces outside the model, bolstering the identifying assumption. The caveat is that a parameterization with zero antitrust risk produces sustainable coordination without the Miller/Coors merger—as is guaranteed to arise theoretically (Proposition 2). That outcome is something of an edge case, however, because the results show that even small amounts of risk are sufficient to undermine coordination.

C.2 Time-Varying Supermarkups

Our baseline results use a version of the price leadership model with a single supermarkup in every region and period post-dating the Miller/Coors merger. It is possible to relax that restriction and allow for region-varying or time-varying supermarkups. We do the latter here. Our identifying assumption is unchanged: the residual costs of ABI do not change, on average, relative to those of the fringe firms. To implement, we construct instruments by interacting the ABI-post merger indicator variable with indicators for (sets of) post-merger periods. Exact identification is maintained.

The qualitative evidence of Section 2.2 suggests that ABI issues its price announcement in August, to take effect in early Fall. Thus, we assume that the same supermarkup applies

[46] We focus on the year 2011, which isolates the effects of the joint venture as other demand and cost factors are unchanged. The marginal cost specification allows the merger to affect marginal costs by reducing shipping distances and via separate vertical shifts for Miller and Coors (e.g., see the discussion under equation (6)). To conduct the counterfactual, we recalculate distribution costs for the year 2011 using pre-merger brewery ownership and 2011 gasoline prices. We also eliminate the estimated vertical shifts in marginal cost.

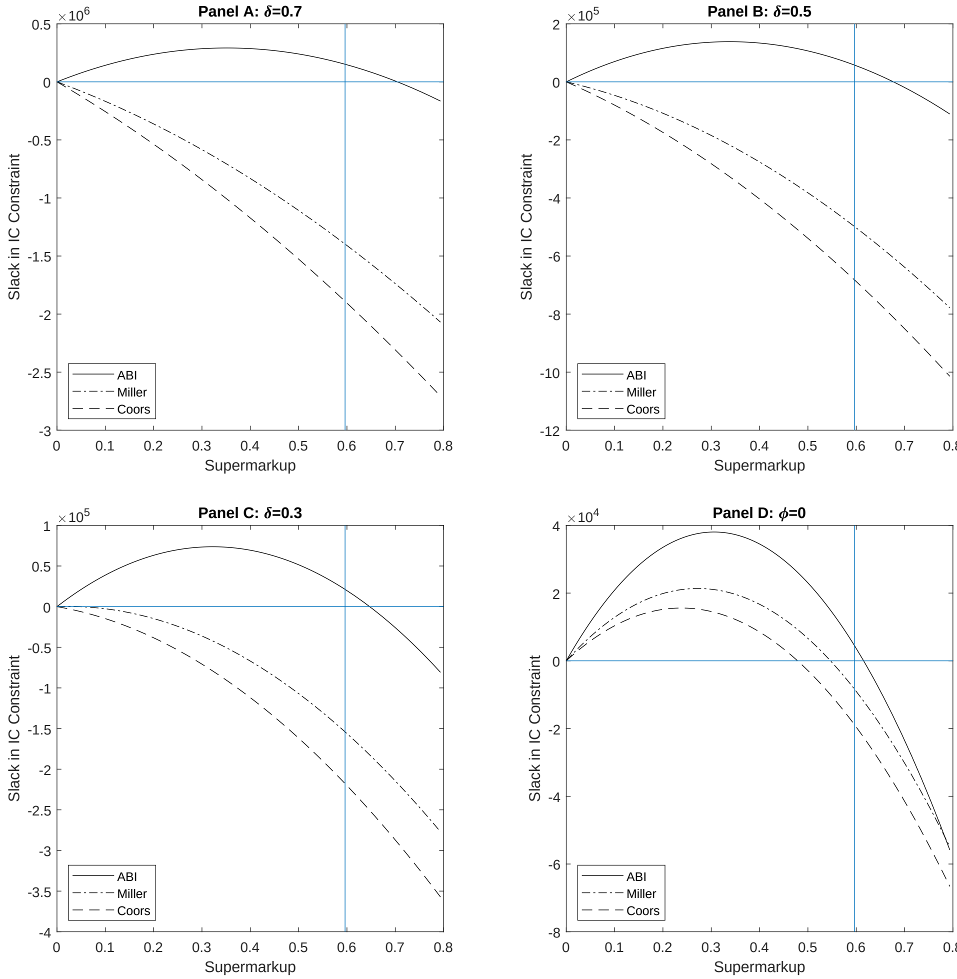

Figure C.1: Slack Functions with an ABI/Miller/Coors Coalition

Notes: The figure provides the slack functions in 2011:Q4 under a counterfactual in which Miller and Coors are independent firms and the coalition includes ABI, Miller, and Coors. IC is satisfied for supermarkup m if the slack functions are positive (i.e., above the horizontal blue line). The vertical blue line shows the estimated supermarkup of 0.596. Four different balancing assumptions are employed: $\delta = 0.7$ (Panel A), $\delta = 0.5$ (Panel B), $\delta = 0.3$ (Panel C), and $\phi = 0$ (Panel D). Results are based on the RCNL-2 demand specification.

to all periods within "fiscal years," which we define as beginning in October and ending in the following September. Our regression sample includes periods for fiscal years 2009 (June-September), 2010 and 2011 (full coverage), and 2012 (October-December). There are four supermarkups to be estimated and four instruments.

Table C.1 provides the estimation results. For each demand specification, we find that supermarkups increase somewhat over time. From a statistical standpoint, this reflects that ABI prices increase relative to fringe prices during the post-merger periods (Figure 1). In general, higher supermarkups imply larger discount factors for a given risk coefficient because the profitability of deviation is greater. However, the time-varying supermarkups are close enough to the constant supermarkups that the slack functions do not change much.

Table C.1: Estimation with Time-Varying Super-Markups

	Parameter	RCNL-1	RCNL-2	RCNL-3	RCNL-4
Super-Markup	m_{2009}	0.386	0.333	0.425	0.393
		(0.064)	(0.067)	(0.069)	(0.068)
Super-Markup	m_{2010}	0.571	0.513	0.659	0.619
		(0.066)	(0.0684)	(0.070)	(0.067)
Super-Markup	m_{2011}	0.737	0.683	0.849	0.809
		(0.093)	(0.083)	(0.078)	(0.076)
Super-Markup	m_{2012}	0.925	0.871	1.064	1.020
		(0.080)	(0.087)	(0.079)	(0.079)

Notes: The table shows the baseline supply results. Estimation is with the method-of-moments. There are 94,656 observations at the brand-size-region-month-year level. The samples excludes the months/quarters between June 2008 and May 2009. Regression includes the marginal shifters, product (brand×size), period (month or quarter), and region fixed effects. Standard errors clustered by region and shown in parentheses.

C.3 Comparison to a Conduct Parameters Approach

Miller and Weinberg (2017) analyze the MillerCoors joint venture using a conduct parameter model. Specifically, brewers are assumed to set prices to satisfy

$$p_t = mc_t - \left[\Omega_t(\kappa) \circ \left(\frac{\partial q_t(p_t)}{\partial p_t}\right)^T\right]^{-1} q_t(p_t)$$

where Ω_t is an ownership matrix, κ is a conduct parameter, and the operation $\circ$ is element-by-element matrix multiplication. The $(j,\, k)$ element of the ownership matrix equals one if products j and k are produced by the same firm, κ if they are sold by ABI and MillerCoors and the period postdates the merger, and zero otherwise. The model nests post-merger Bertrand ($\kappa = 0$) and joint profit maximization for ABI/MillerCoors ($\kappa = 1$).

The identifying assumption—that ABI residual costs do not change relative to fringe firms—is identical to what we employ in this paper. Thus, it is interesting to compare the results generated from the conduct parameter model to those of the more structural price leadership model, as any differences are due solely to how the models interpret the data. As the main parameters of interest—the conduct parameter and the supermarkup—have

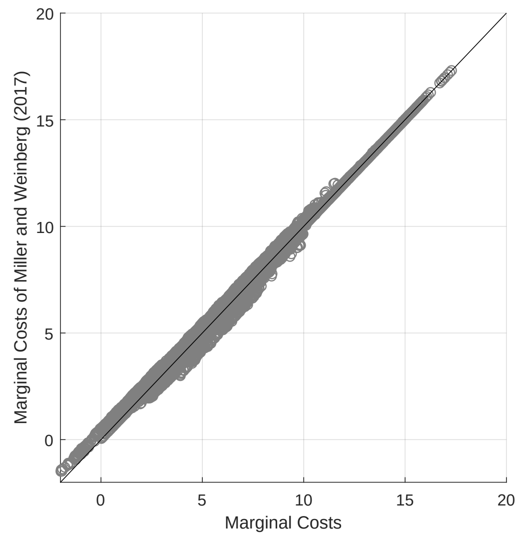

Figure C.2: Empirical Distribution of Marginal Costs

Notes: The figure plots the marginal costs obtained from the price leadership model (horizontal axis) against the marginal costs obtained from the conduct parameter approach of Miller and Weinberg (2017) (vertical axis). Results are based on the RCNL-1 demand specification.

different economic interpretations, we view the vector of implied marginal costs as providing the cleanest comparison. Figure C.2 plots the marginal costs of the two models. The dots, each representing a product-region-year observation, fall along the 45-degree line, indicating that the models have similar implications for costs.

C.4 Linear Demands

One potentially surprising result from the price leadership model is that deviation does not increase profit much, relative to the price leadership equilibrium, at the estimated super-markups (recall Figure 4). We obtain the result by numerically simulating the best response of each coalition firm to the price leadership prices. One may wonder, then, how important is the curvature of the RCNL demand system in generating the result. As one simple check, we calibrate the linear demand system

$$q_j = a_j + \sum_k b_{jk} p_k$$

such that the elasticities exactly match those of the RCNL when evaluated at the average prices and quantities in 2011.[47] This allows us to repeat the numerical simulations using the same initial elasticities, but with different curvature assumptions. The results are shown in Appendix Figure C.3. As is the case with RCNL demand, the benefit of deviation increases with the supermarkup, but is still relatively small at estimated supermarkup. Thus, we conclude that our findings are not overly dependent on the logit assumption.

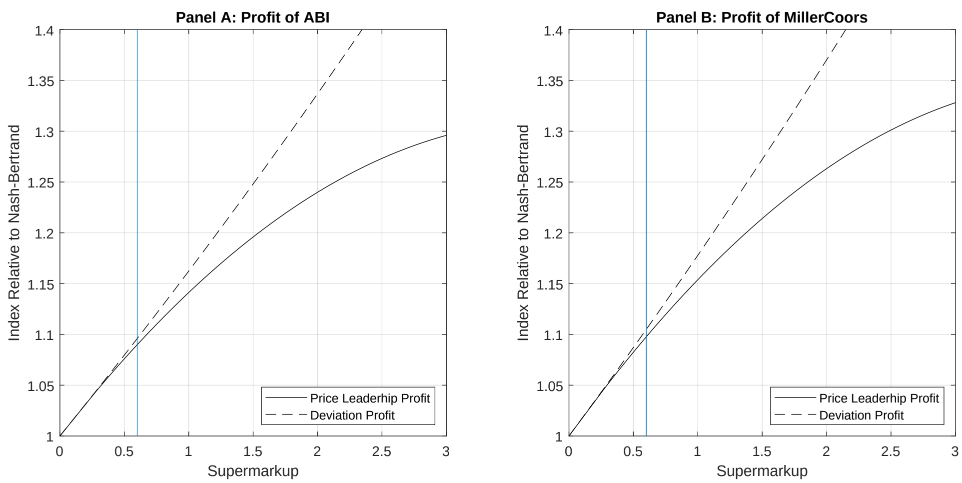

Figure C.3: Profit with Price Leadership and Deviation under a Linear Demand System

Notes: The figure provides the profit of ABI (Panel A) and MillerCoors (Panel B) in 2011:Q4 under price leadership and deviation. Results are generated with simulations that employ a linear demand system that is calibrated to RCNL-2 derivatives evaluated at observed prices. Statistics are computed for a range of supermarkups ($m \in [0, 3]$). All statistics are reported relative to their Bertrand analog. The vertical line marks the supermarkup estimated from the data.

[47] See Miller et al. (2016) for mathematical details on linear demand calibration.

D Additional Figures and Tables

Table D.1: Demand Estimates

Demand Model:		RCNL-1	RCNL-2	RCNL-3	RCNL-4
Data Frequency:		Monthly	Quarterly	Monthly	Quarterly
Variable	Parameter	(i)	(ii)	(iii)	(iv)
Price	α	-0.0887	-0.1087	-0.0798	-0.0944
		(0.0141)	(0.0163)	(0.0147)	(0.0146)
Nesting Parameter	ρ	0.8299	0.7779	0.8079	0.8344
		(0.0402)	(0.0479)	(0.0602)	(0.0519)
Demographic Interactions					
Income×Price	Π_1	0.0007	0.0009		
		(0.0002)	(0.0003)		
Income×Constant	Π_2	0.0143	0.0125	0.0228	0.0241
		(0.0051)	(0.0055)	(0.0042)	(0.0042)
Income×Calories	Π_3	0.0043	0.0045	0.0038	0.0031
		(0.0016)	(0.0017)	(0.0018)	(0.0015)
Income×Import	Π_4			0.0039	0.0031
				(0.0019)	(0.0016)
Income×Package Size	Π_5			-0.0013	-0.0017
				(0.0007)	(0.006)
Other Statistics					
Median Own Price Elasticity		-4.74	-4.33	-4.45	-6.10
Median Market Price Elasticity		-0.60	-0.72	-0.60	-0.69

Notes: This table shows the baseline demand results. There are 94,656 observations at the brand–size–region–month–year level in columns (i) and (iii), and 31,784 observations at the brand–size–region–year–quarter level in columns (ii) and (iv). The samples exclude the months/quarters between June 2008 and May 2009. All regressions include product (brand×size) and period (month or quarter) fixed effects. The elasticity numbers represent medians among all the brand–size–region–month/quarter–year observations. Standard errors are clustered by region and shown in parentheses. Reproduced from Miller and Weinberg (2017).

Table D.2: Product-Specific Elasticities for 12 Packs

	Brand/Category	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
								Product-Specific Own and Cross-Elasticities						
(1)	Bud Light	-4.389	0.160	0.019	0.182	0.235	0.101	0.146	0.047	0.040	0.130	0.046	0.072	0.196
(2)	Budweiser	0.323	-4.272	0.019	0.166	0.258	0.103	0.166	0.047	0.039	0.121	0.043	0.068	0.183
(3)	Coors	0.316	0.154	-4.371	0.163	0.259	0.102	0.167	0.046	0.038	0.119	0.042	0.066	0.180
(4)	Coors Light	0.351	0.160	0.019	-4.628	0.230	0.100	0.142	0.047	0.041	0.132	0.047	0.073	0.199
(5)	Corona Extra	0.279	0.147	0.018	0.137	-5.178	0.108	0.203	0.047	0.035	0.104	0.035	0.061	0.158
(6)	Corona Light	0.302	0.151	0.018	0.153	0.279	-5.795	0.183	0.048	0.037	0.113	0.039	0.065	0.171
(7)	Heineken	0.269	0.145	0.018	0.131	0.311	0.108	-5.147	0.047	0.035	0.101	0.034	0.059	0.153
(8)	Heineken Light	0.240	0.112	0.014	0.124	0.210	0.086	0.138	-5.900	0.026	0.089	0.028	0.051	0.135
(9)	Michelob	0.301	0.140	0.015	0.146	0.208	0.089	0.135	0.042	-4.970	0.116	0.036	0.061	0.175
(10)	Michelob Light	0.345	0.159	0.019	0.181	0.235	0.101	0.146	0.047	0.041	-5.071	0.046	0.072	0.196
(11)	Miller Gen. Draft	0.346	0.159	0.019	0.182	0.235	0.101	0.146	0.047	0.040	0.130	-4.696	0.072	0.196
(12)	Miller High Life	0.338	0.159	0.019	0.177	0.242	0.102	0.153	0.047	0.040	0.127	0.045	-3.495	0.191
(13)	Miller Lite	0.344	0.159	0.019	0.180	0.237	0.101	0.148	0.047	0.040	0.129	0.046	0.071	-4.517
(14)	Outside Good	0.016	0.007	0.001	0.009	0.011	0.005	0.006	0.002	0.002	0.006	0.002	0.003	0.009
								Cross Elasticities by Category						
	6 Packs	0.307	0.152	0.018	0.155	0.275	0.104	0.180	0.047	0.038	0.115	0.039	0.065	0.174
	12 Packs	0.320	0.154	0.019	0.163	0.250	0.102	0.161	0.047	0.039	0.121	0.042	0.068	0.183
	24 Packs	0.356	0.160	0.019	0.189	0.222	0.099	0.136	0.047	0.041	0.134	0.048	0.073	0.201
	Domestic	0.349	0.160	0.019	0.184	0.229	0.100	0.142	0.047	0.040	0.131	0.047	0.072	0.197
	Imported	0.279	0.147	0.018	0.138	0.301	0.108	0.200	0.047	0.035	0.104	0.035	0.061	0.158

Notes: This table provides the mean elasticities of demand for 12 packs based on the RCNL-1 specification (column (i) of Table D.1). The cell in row i and column j is the percentage change in the quantity of product i with respect to the price of product j. Means are calculated across year–month–region combinations. The category cross-elasticities are the percentage change in the combined quantities of products in the category due to a 1 percent change in the price of the product in question. Letting the category be defined by the set B, we calculate $\left(\sum_{j\in B, j\neq k} \frac{\partial q_j(p)}{\partial p_k}\right) \frac{p_k}{\sum_{j\in B, j\neq k} q_j(p)}$. The categories exclude the product in question. Thus, for example, the table shows that a 1 percent change in the price of a Bud Light 12 pack increases the sales of other 12 packs by 0.320 percent. Reproduced from Miller and Weinberg (2017).

Table D.3: Firm-Specific Elasticities

Panel A: Mean Elasticities in 2007

Brand/Category		(1)	(2)	(3)	(4)	(5)
(1)	ABI	-2.92	1.00	0.63	0.48	0.25
(2)	Miller	2.02	-3.30	0.65	0.47	0.24
(3)	Coors	2.05	1.04	-4.08	0.46	0.23
(4)	Modelo	1.55	0.75	0.44	-5.26	0.34
(5)	Heineken	1.51	0.73	0.42	0.65	-5.44

Panel B: Mean Elasticities in 2011

Brand/Category		(1)	(2)	(3)	(4)
(1)	ABI	-2.97	1.68	0.41	0.23
(2)	MillerCoors	2.01	-2.86	0.40	0.23
(3)	Modelo	1.67	1.36	-5.24	0.29
(4)	Heineken	1.61	1.30	0.49	-5.42

Notes: This table provides the mean firm-specific elasticities of demand in 2007 and 2011 based on the RCNL-1 specification (column (i) of Table D.1). The cell in row i and column j is the percentage change in the quantity of firm i with respect to the prices of firm j. The elasticity of demand for products in set A with respect to prices of products in set B is defined as: $\left(\sum_{n\in A}\sum_{k\in B}\frac{\partial q_n}{\partial p_k}\right)\frac{\overline{p}_B}{\sum_{n\in A} q_n}$. Means are calculated across month–region combinations.

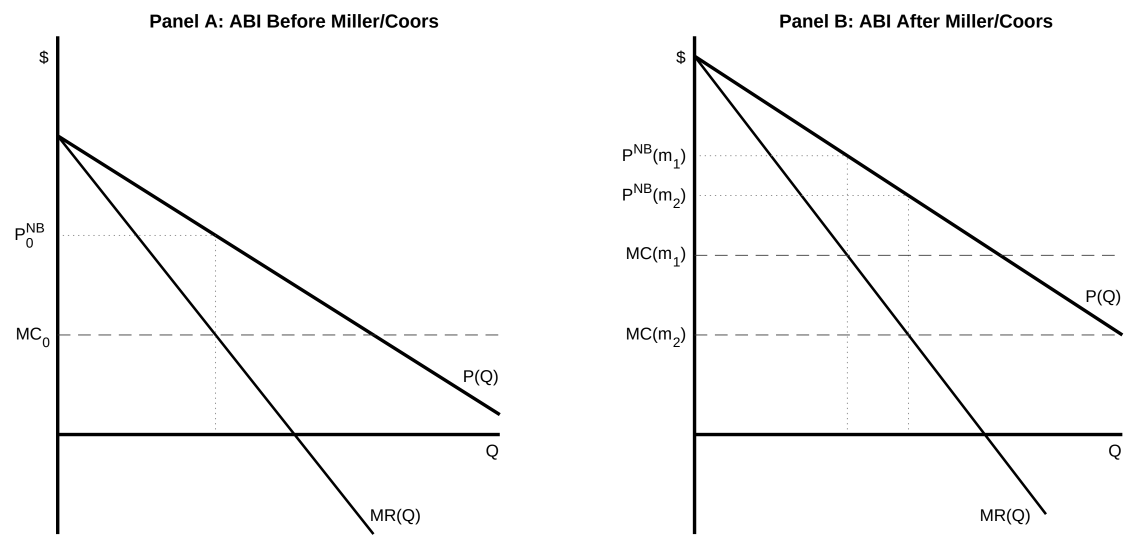

Figure D.1: Illustration of the Identification Strategy

Notes: Panel A considers ABI before the Miller/Coors merger. The residual demand function ($P(Q)$) and marginal revenue function ($MR(Q)$) are known from demand estimates. ABI's Nash-Bertrand prices (P_0^{NB}) are data. Thus, marginal costs can be recovered (MC_0). Panel B considers ABI after the Miller/Coors merger. The residual demand and marginal revenue functions shift out in the Nash-Bertrand equilibrium because Miller and Coors prices are higher. Each candidate super-markup (m_1 and m_2) corresponds to a different implied Nash-Bertrand price of ABI, and thus a different implied marginal cost ($MC(m_1)$ and $MC(m_2)$). Thus, a restriction on the differences in marginal costs across panels can identify the supermarkup. In this illustrative example, the restriction $MC_0 = MC(m)$ implies the supermarkup is $m = m_2$.