

Oligopolistic Price Leadership and Mergers: The United States Beer Industry[†]

By NATHAN H. MILLER, GLORIA SHEU, AND MATTHEW C. WEINBERG*

We study a repeated game of price leadership in which a firm proposes supermarkups over Bertrand prices to a coalition of rivals. Supermarkups and marginal costs are recoverable from data on prices and quantities using the model's structure. In an application to the beer industry, we find that price leadership increases profit relative to Bertrand competition by 17 percent in fiscal years 2006 and 2007, and by 22 percent in 2010 and 2011, with the change mostly due to consolidation. We simulate two mergers, which relax binding incentive compatibility constraints and increase supermarkups. These coordinated effects arise even with efficiencies that offset price increases under Bertrand competition. (JEL G34, K21, L13, L14, L41, L66)

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 economics. 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 follow-the-leader pricing in retail industries ranging from supermarkets, pharmacies, and gasoline (Clark and Houde 2013, Seaton and Waterson 2013, Chilet 2018, Lemus and Luco 2021, 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

* Miller: McDonough School of Business, Georgetown University (email: nhm27@georgetown.edu); Sheu: Board of Governors of the Federal Reserve System (email: gloria.sheu@frb.gov); Weinberg: Ohio State University (email: weinberg.133@osu.edu). Liran Einav was the coeditor for this article. This material is based on work supported by the National Science Foundation under Grants 1824318 (Miller) and 1824332 (Weinberg). We thank many seminar and conference participants for insightful comments and suggestions. All estimates and analyses in this article based on IRI data are by the authors and not by IRI. Much of the research was conducted while Gloria Sheu was a staff economist at the US Department of Justice. The analysis and conclusions set forth are those of the authors and do not indicate concurrence by other members of the Board research staff or the Board of Governors. Furthermore, the views expressed here should not be purported to reflect those of the US Department of Justice.

[†] Go to <https://doi.org/10.1257/aer.20190913> to visit the article page for additional materials and author disclosure statements.

¹ See also the discussions in Lanzillotti (2017) and Harrington (2018). 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.

supracompetitive markups, explore implications for welfare, nor provide a framework for the analysis of counterfactuals.²

This article presents an empirical model of oligopolistic price leadership that can be evaluated with market level data on prices and quantities. In each period of an infinitely repeated game, the leader makes a nonbinding price announcement and then all firms set prices simultaneously. The price announcement is cheap talk that shapes firm beliefs and facilitates supracompetitive pricing. We apply the model to a setting that exhibits such price leadership behavior—the beer industry of the United States. We recover the marginal costs and markups of each product from the data using first order conditions for profit maximization and the demand estimates of Miller and Weinberg (2017). A comparison to Bertrand equilibrium, obtained with a counterfactual simulation, allows us to quantify the implications of price leadership for firms and consumers.

Our modeling approach provides a framework for evaluating the coordinated effects of mergers in markets characterized by price leadership. To illustrate, we examine two mergers in particular, and show that they relax incentive compatibility (IC) constraints and increase prices. These results obtain even in the presence of marginal costs efficiencies sufficient to offset unilateral effects, which are the price changes under the common assumption of static price competition before and after the merger. Previously, the empirical industrial organization literature has provided little in the way of methodologies that could be used to guide coordinated effects analysis. Indeed, our research is among the first to formally model coordinated effects in real-world markets.³

We organize the article as follows. We start in Section I with a description of US brewing markets. In scanner data spanning 2001–2011, a handful of brewers account for the bulk of retail revenue. In the earlier years of the sample, these firms are Anheuser-Busch (ABI), SABMiller, Molson Coors, Grupo Modelo, and Heineken. In the later years, SABMiller and Molson Coors are replaced with their joint venture, MillerCoors; we often refer to this consolidating event as the Miller-Coors merger. We summarize the qualitative evidence of price leadership behavior, citing in particular to legal documents filed by the Department of Justice (DOJ) alleging that ABI preannounces its annual list price changes as a signal to competitors, and that MillerCoors tends to follow. We also discuss data sources, provide additional summary statistics and stylized facts, and describe the demand model of Miller and Weinberg (2017), which we take as given in this article.

We then formalize the model of oligopolistic price leadership in Section II. Firms compete across multiple geographic regions in an infinitely repeated game of perfect information. Each period has two stages. In the first, the leader announces nonbinding and region-specific *supermarkups* above Bertrand prices. On the equilibrium path, a set of coalition firms, comprised of the leader and its followers, adopts the supermarkups in a subsequent pricing stage. The leader

²The 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.

³We refer readers to Baker (2002, 2010) and Harrington (2013) for a summary of the legal literature on coordinated effects. The theoretical literature includes Compte, Jenny, and Rey (2002); Vasconcelso (2005); Ivaldi et al. (2007); Bos and Harrington (2010); and Loertscher and Marx (2021). Empirical models include Davis and Huse (2010) and Igami and Sugaya (forthcoming).

selects the supermarkups to maximize its profit, subject to the 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, which price to maximize static profit functions. Deviation, which occurs only off the equilibrium path, is punished with reversion to Bertrand pricing in all regions. These strategies constitute a subgame perfect equilibrium (SPE).

We discuss the empirical implementation in Section III. Our approach relies on a pair of identification results: (i) marginal costs can be recovered, given any supermarkups, using first order conditions for static profit maximization; and (ii) the supermarkups can be recovered using another set of first order conditions that arise from the leader's constrained maximization problem. The latter result requires an *ex ante* assumption on a reduced-form timing parameter, which summarizes the patience of firms and various (unknown) aspects of the game, such as the duration of punishment. Larger timing parameters imply higher supermarkups and thus lower marginal costs for coalition firms, so with some prior knowledge of costs it is possible to evaluate the timing assumptions *ex post*. We use an orthogonality condition for this purpose, namely an assumption that ABI's marginal costs do not change differently from those of Modelo and Heineken, on average, with the Miller-Coors merger.

We then summarize the empirical results and analyze equilibrium in Section IV. Using the timing parameter that best satisfies the identifying assumption, we recover average supermarkups of \$1.20 in fiscal year 2007, just before Miller-Coors, and \$1.80 in fiscal year 2010, just after. The change in supermarkups between 2007 and 2010 reflects that the merger created slack in the binding IC constraint and also greater symmetry among coalition firms, though new cost and demand conditions also contribute. The difference between industry profits under price leadership and profits under static Bertrand competition is 17 percent and 22 percent of Bertrand profits in 2007 and 2010, respectively. The reduction in consumer surplus is 154 percent and 170 percent of the producer surplus gain in those two years.

Supermarkups tend to be higher in regions where ABI has large market shares, and lower in regions where Coors (in 2007) and MillerCoors (in 2010) have large market shares. This reflects the Kuhn-Tucker conditions that characterize the solution to ABI's constrained maximization problem: ABI benefits more from a higher supermarkup if it has a large market share, and the effect of a higher supermarkup on the binding IC constraint is greater if Coors and MillerCoors have large market shares. Relatedly, we use counterfactual simulation to explore the role of multi-market (here, multiregion) contact. The results indicate that multi-market contact affects the spatial dispersion of supermarkups, but less so after the Miller-Coors merger due to the enhanced symmetry among coalition firms.

In Section V, we use the model to examine the coordinated effects of the Miller-Coors merger and ABI's proposed acquisition of Modelo, which was approved in 2013 by the DOJ only after the Modelo brands were sold to a third party. We model the latter merger as it would have occurred without the divestiture. In both instances we find that the merger loosens the IC constraint of the binding firm, resulting in an increase for the supermarkup on domestic beers of \$0.50 due to Miller-Coors and an increase of \$0.40 due to ABI-Modelo. Miller-Coors combines the two smaller firms from the pricing coalition into one firm that then

faces demand and cost conditions that are more similar to the leader, ABI. The ABI-Modelo merger, as originally proposed, brings the largest outside firm into the coalition, which lessens competition from the pricing fringe. In neither case are the price effects mitigated by marginal cost efficiencies.

We conclude in Section VI with a short summary and a discussion of some of the more important modeling assumptions and limitations, with an eye toward informing future research efforts. As one example, the framework we introduce can be applied to model how mergers affect pricing in a coordinated equilibrium, but it is less well suited to evaluate whether equilibrium play would shift from a static Bertrand equilibrium to a coordinated equilibrium. The online Appendix includes additional details on the data, a set of theoretical results and proofs, a description of computational methods, and assorted additional analyses.

Literature Review

Our research is methodologically most similar to Igami and Sugaya (forthcoming), which studies the vitamin C cartel of the 1990s. Among the main findings of that paper 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 notable differences in the models. For example, Igami and Sugaya assume all firms engage in maximal collusion or revert to Cournot equilibrium forevermore. Some interesting aspects of our model, such as partial coalitions, multi-market contact, and the leader's ability to adjust the prices to satisfy IC constraints, are not present.

Eizenberg, Shilian, and Blanga (2020) and Eizenberg and Shilian (2019) also estimate IC constraints in empirical settings. The first of these estimates demand for hummus salad and instant coffee in the Israeli grocery sector, and recovers marginal costs with an assumption of Bertrand competition. It then evaluates hypothetical coordination and determines that multi-market contact would not substantially relax IC constraints, a result that is attributed to symmetry across the two product categories. In our setting, the degree of symmetry also affects the impact of multi-market contact. The second paper estimates conduct in 40 food categories and imputes the minimum discount factor necessary to support the conduct in SPE.

More broadly, our research relates to articles that seek to understand the equilibrium concept that governs competition in specific markets. Two of the more prominent articles focus on Bertrand equilibrium and joint profit maximization (Bresnahan 1987, Nevo 2001), though Stackelberg leadership and various other possibilities also have been examined (e.g., Gasmi, Laffont, and Vuong 1992; Slade 2004; Rojas 2008). Other studies use a conduct parameter approach to identify changes in the intensity of competition, without taking a stance on the precise equilibrium concepts (e.g., Porter 1983, Igami 2015, Miller and Weinberg 2017, Michel and Weiergraeber 2018). Finally, there is an empirical literature that tests whether multi-market contact leads to higher prices, without modeling the underlying game (e.g., Evans and Kessides 1994, Ciliberto and Williams 2014, Khwaja and Shim 2017).

The price leadership model itself draws on a number of theoretical contributions. An important precursor is the canonical Rotemberg and Saloner (1986) model of collusion, in which there is perfect information and collusive prices adjust so that deviation does not occur along the equilibrium path.⁴ Our treatment of multi-market contact extends a model of differentiated-products price competition developed in Bernheim and Winston (1990). We develop the connections to the theoretical literature more extensively in Section II, after presenting the model of price leadership.

I. The US Beer Market

A. Background

Beer is differentiated along multiple dimensions, including taste, calories, brand image, and package size. Most beer sold in the United States is lager style, and a handful of brewers dominate this market. This is despite recent growth of craft brewers, which tend to specialize in ales. Retailers and distributors also play a role in the supply chain, but in a different way than most industries because of regulation on the sales and distribution of alcohol dating back to prohibition. 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.⁵ 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 percent of revenue while Grupo Modelo and Heineken (importers) account for another 12–16 percent of revenue.⁶ Midway through the sample, in June 2008, SABMiller and Molson Coors consolidated their US operations into the MillerCoors joint venture. The DOJ elected not to challenge consolidation on the basis that cost savings in distribution likely would offset any loss of competition.⁷ In the same year, InBev purchased Anheuser-Busch (forming ABI). As InBev previously sold brands with limited sales in the United States, such as Stella Artois, the transaction had only minor consequences for market structure.

There have been other notable transactions after the sample period. In 2013, ABI acquired Grupo Modelo. The DOJ obtained a settlement under which the rights to the Grupo Modelo brands in the United States transferred to Constellation, at that

⁴A repeated game in which oligopolistic price leadership emerges is provided in Rotemberg and Saloner (1990). As their model incorporates asymmetric information, price announcements have informational and strategic content. Our model does not include asymmetric information.

⁵The 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).

⁶We refer to the first three firms as “domestic” because their beer is brewed in the United States.

⁷Subsequent academic research suggests that sizable cost savings were realized but were dominated by adverse competitive effects (Ashenfelter, Hosken, and Weinberg 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 and InBev to form A-B InBev (ABI), which closed in April 2009, and the acquisition by Heineken of the Fomento Económico Mexicano (FEMSA) brands in April 2010. All statistics are based on supermarket sales recorded in Information Resources, Inc. (IRI) scanner data. Reproduced from Miller and Weinberg (2017, p. 1767), printed with permission of the Econometric Society.

time a major distributor of wine and liquor.⁸ In 2016, ABI acquired SABMiller. In order to gain DOJ approval, SABMiller sold its stake in the MillerCoors joint venture to Molson Coors. Finally, ABI has purchased a number of craft breweries over the last decade, and now owns Goose Island, Devil's Backbone, and Kona, among a number of other craft brands.

B. Price Leadership in the Beer Industry

There is extensive qualitative evidence that competition among brewers involves price leadership behavior. Legal documents filed in 2013 by the DOJ to enjoin the ABI-Modelo acquisition allege that

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.⁹

The legal documents do not specify whether these pricing practices were used prior to the Miller-Coors merger in 2008. However, two prominent industry studies describe price leadership as occurring throughout the latter half of the twentieth century (Greer 1998, Tremblay and Tremblay 2005).¹⁰ Further, a recent enforcement action of the DOJ, related to ABI's acquisition of the Craft Brewers Alliance (CBA), suggests that price coordination is ongoing.¹¹ We interpret these descriptions as

⁸The 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.

⁹Paragraph 44 of the Complaint in *US v. Anheuser-Busch InBev SA/NV and Grupo Modelo S.A.B. de C.V.*

¹⁰For example, Tremblay and Tremblay (2005, p. 49) states that “Anheuser-Busch serves as the price leader for the industry” and that “[m]ost other brewers … key to Budweiser, usually matching Bud's price for premiums and going somewhat above it or below it for superpremiums and populars.” Tremblay and Tremblay (2005, p. 168) states that the purpose of leadership is to “maintain high prices.”

¹¹The DOJ press release describes the acquisition and divestiture, and states that “[b]y eliminating CBA's Kona brand as a competitive restraint, ABI would also likely have greater ability to facilitate price coordination, resulting in higher prices” See <https://www.justice.gov/opa/pr/justice-department-requires-divestiture-order-anheuser-busch-acquire-craft-brew-alliance>, last accessed September 19, 2020.

suggesting that price leadership occurs throughout the sample period, and maintain that assumption in our empirical analysis.¹²

The qualitative evidence guides other modeling assumptions that we make as well. First, price leadership behavior in the industry does not appear to 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.”¹³ Drennan, Magura, and Nevo (2013, p. 295), 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.”¹⁴

Second, trigger strategies may be important in sustaining supracompetitive pricing. Tremblay and Tremblay (2005, p. 168) state “Anheuser-Busch has used and threatened to use substantial price reductions to punish rivals....”¹⁵ Price wars appear to be relatively infrequent, however, as only one example is provided (occurring over 1953–1955).

Third, the legal documents filed by the DOJ in 2013 provide some support for our assumption that the leader’s price announcement serves as an equilibrium selection device. The following passage quotes from the business documents of ABI:

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.”¹⁶

We view this passage as suggesting 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.

C. Data and Prices

We use retail scanner data from the IRI Academic Database (Bronnenberg, Kruger, and Mela 2008), which contains weekly revenue and unit sales by Universal Product Code (UPC) for a sample of stores over 2001–2011. We restrict attention to supermarkets, which accounted for 26 percent of off-premise beer sales in 2011 (Beer Institute 2012).¹⁷ We aggregate the UP Cs to the brand \times size level. For convenience, we often refer to brand \times size combinations as “products.” We focus on

¹² Miller and Weinberg (2017) cite evidence culled from the annual reports of ABI and SABMiller that suggest competition was relatively tough over 2005–2008. This is consistent with our results, which indicate that the MillerCoors merger resulted in higher supermarkups.

¹³ Paragraph 49 of the Complaint in *US v. Anheuser-Busch InBev SA/NV and Grupo Modelo S.A.B. de C.V.*

¹⁴ The legal filings also speak to this. For example, the Competitive Impact Statement (p. 8) states that “[by] compressing the price gap between high-end and premium brands, Modelo’s actions have increasingly limited ABI’s ability to lead beer prices higher, resulting in higher prices....” The legal filings do not address Heineken specifically, though their prices are similar to Modelo’s in the data.

¹⁵ See also Greer (1998, p.50): “Anheuser-Busch’s strategy includes cutting price to discipline rivals....”

¹⁶ Para 46 of the Complaint in *US v. Anheuser-Busch InBev SA/NV and Grupo Modelo S.A.B. de C.V.*

¹⁷ The other major sources of off-premise beer sales are liquor stores (38 percent), convenience stores (26 percent), mass retailers (6 percent), and drugstores (3 percent). The price and quantity patterns that we observe for supermarkets also exist for drug stores, which are in the IRI Academic Database. See Miller and Weinberg (2017).

13 flagship brands sold as 6 packs, 12 packs, and 24 packs.¹⁸ We measure quantities based on 144-ounce equivalent units, the size of a 12 pack, and measure price as the ratio of revenue to equivalent unit sales. These choices comport with Miller and Weinberg (2017).

In our main supply-side analysis, we aggregate the weeks to quarters and use data from 37 distinct geographic regions. Thus, the primary unit of observation is a product-region-quarter. Because list price adjustments take effect in the fall (see the previous section), we further group quarters into “fiscal years” that begin in October and end in the following September. Thus, for example, fiscal year 2007 comprises the fourth quarter of 2006 and the first three quarters of 2007. We restrict attention to the fiscal years 2006, 2007, 2010, and 2011, each of which is fully contained in the sample period. We exclude earlier years because our demand model relies on household demographics from the Public Use Microdata Sample (PUMS) of the American Community Survey, which is available starting in 2005. We omit fiscal years 2008 and 2009 due to their proximity to the Miller-Coors merger, and fiscal year 2005 because it is only partially covered in the data.¹⁹

We rely on a number of other sources to complete the dataset. We use Google Maps to obtain the driving miles between each IRI region and the nearest brewery for each of the domestic products. For the imported brands, we obtain the driving miles between the regions and the nearest port into which the beer is shipped.²⁰ Our measure of “distance” is the multiplicative product of driving miles and diesel fuel prices, which we obtain from the Energy Information Agency of the Department of Energy. This allows us to capture variation in transportation costs that arises both cross-sectionally, based on the location of regions and breweries, and intertemporally, based on fluctuations in fuel costs. All prices and incomes are deflated using the consumer price index and are reported in 2010 dollars. See online Appendix A for additional details about the data.

Table 2 shows the average price of each product (brand \times size) in fiscal year 2011, along with the share of volume among the products in the sample. Domestic brands and larger package sizes tend to have lower prices. Volume shares increase in package sizes for the domestic brands, whereas 12 packs have the greatest volume shares for the imported brands. The most popular domestic brands are Bud Lite, Coors Light, Miller Lite, and Budweiser. The most popular imported brands are Corona Extra and Heineken.

Figure 1 shows the time path of average retail prices over 2001–2011 for each firm’s most popular 12 pack. The red vertical line at June 2008 marks the Miller-Coors merger. As shown, the prices of domestic beers increase abruptly after the merger,

¹⁸ We combine 24 packs and 30 packs in the construction of these products because whether 24 or 30 packs are sold tends to depend on region-specific historical considerations. We exclude 18 packs and promotional package sizes, which generate fewer sales.

¹⁹ We focus on quarterly aggregation because it is computationally less burdensome than monthly aggregation. However, the empirical variation we exploit is similar to what is used in the simpler model of Miller and Weinberg (2017), and those results are robust to monthly and quarterly aggregation.

²⁰ We obtain the location of Heineken’s primary ports from the website of BDP, a logistics firm hired by Heineken to improve its operational efficiency. See <http://www.bdpinternational.com/clients/heineken/>, last accessed on February 26, 2015. The ports include Baltimore, Charleston, Houston, Port of Long Beach, Miami, Seattle, Oakland, Boston, and New York. We measure the shipping distance for Grupo Modelo brands as the driving distance from each retail location to Ciudad Obregon, Mexico.

TABLE 2—PRICES AND VOLUME SHARES IN FISCAL YEAR 2011

Brand	Brewer	6 packs		12 packs		24 packs		All
		Share	Price	Share	Price	Share	Price	
Bud Light	ABI	0.020	11.61	0.074	10.02	0.193	8.16	0.288
Budweiser	ABI	0.011	11.60	0.031	10.01	0.074	8.15	0.116
Coors	MillerCoors	0.001	11.75	0.004	10.10	0.011	8.07	0.017
Coors Light	MillerCoors	0.010	11.63	0.042	10.08	0.107	8.12	0.159
Corona Extra	Modelo	0.010	15.85	0.040	13.02	0.018	12.56	0.067
Corona Light	Modelo	0.005	15.89	0.019	13.07	0.002	12.62	0.027
Heineken	Heineken	0.007	16.22	0.030	13.35	0.007	12.80	0.044
Heineken Light	Heineken	0.002	16.33	0.007	13.42	0.001	11.79	0.010
Michelob	ABI	0.002	12.43	0.005	10.81	0.004	8.03	0.011
Michelob Light	ABI	0.008	12.56	0.025	10.88	0.017	8.63	0.049
Miller Genuine Draft	MillerCoors	0.003	11.65	0.007	10.05	0.012	8.14	0.022
Miller High Life	MillerCoors	0.004	9.15	0.022	7.92	0.032	6.69	0.058
Miller Lite	MillerCoors	0.009	11.60	0.046	10.09	0.110	8.13	0.164

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

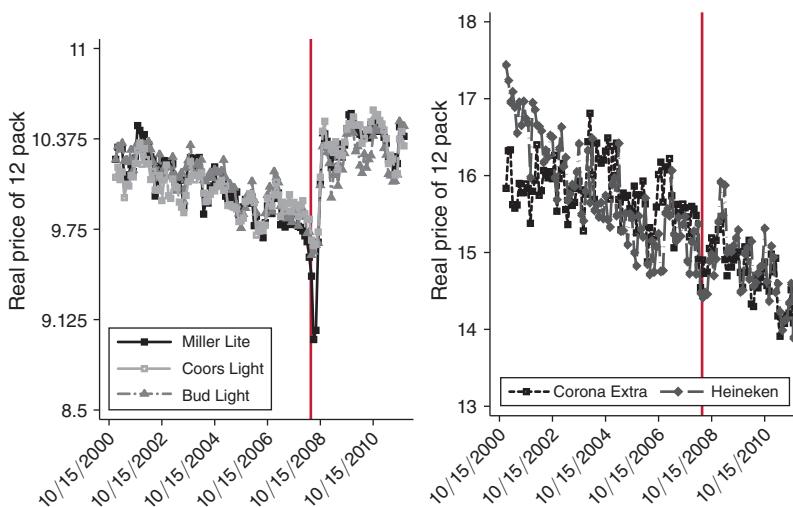


FIGURE 1. AVERAGE RETAIL PRICES OF FLAGSHIP BRAND 12 PACKS

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, p. 1769), printed with permission of the Econometric Society.

while import prices continue on trend. Notably, the price increases of ABI are commensurate with those of MillerCoors. Miller and Weinberg (2017) determines that the data are difficult to explain as a shift among Bertrand equilibria, absent a sizable ABI marginal cost increase. These price data also are a focus of the present study, and we show that they can be rationalized with a price leadership framework featuring binding IC constraints.

D. Demand

We rely on the random coefficient nested logit (RCNL) model of Miller and Weinberg (2017) to characterize consumer demand. Our preferred specifications are RCNL-1 and RCNL-2, which allow income to affect the price parameter. This incorporates consumer heterogeneity in the willingness-to-pay for more expensive imported beers and smaller package sizes. Among these two specifications, we focus on RCNL-2, so as to comport with our quarterly observations. The median product-level demand elasticity is -4.74 whereas the median market demand elasticity is -0.60 , indicating that consumers tend to substitute among beer products, rather than between beer and the outside good.²¹ Miller and Weinberg (2017) find these demand elasticities to be reasonably robust across a range of specification choices. We describe the RCNL demand model in greater detail in online Appendix E.

II. Model

A. Overview

We now develop the model of oligopoly price leadership. Let there be $i = 1, \dots, F$ firms and $j = 1, \dots, J$ differentiated, substitute products. Each firm i produces a subset \mathcal{J}_i of all products. We take as given the existence of a pricing *coalition*, a set \mathcal{C} which comprises a leader and at least one follower. Firms not in the coalition are considered part of the *fringe*. The allocation of firms to the coalition and fringe is fixed and predetermined. There are $r = 1, \dots, R$ distinct geographic regions. In each period ($t = 1, 2, \dots$) an economic state of demand and cost conditions is realized, and competition then plays out in two stages:

- (i) The leader announces nonbinding supermarkups, $m_{rt} \geq 0$ for each region.
- (ii) All firms price simultaneously, given $\mathbf{m}_t = (m_{1t}, m_{2t}, \dots, m_{Rt})$ and the history.

The game ends with probability $(1 - \phi)$ after each period. The nonbinding announcements are cheap talk that shape beliefs in the pricing stage. Thus, they are not a theoretical necessity and could be replaced with an assumption on equilibrium selection.²²

Firms maximize the present value of profit. Let $\mathbf{p}_{rt} = (p_{1rt}, p_{2rt}, \dots, p_{Jrt})$ be the prices in region r and period t , and let Ψ_t denote the economic state. The profit function of firm i is

$$(1) \quad \pi_{irt}(\mathbf{p}_{ri}; \Psi_t) = \sum_{j \in \mathcal{J}_i} (p_{jrt} - mc_{jrt}(\Psi_t)) q_{jrt}(\mathbf{p}_{ri}; \Psi_t),$$

²¹ Median firm-level elasticities are around three for the domestic brewers and five for the import brewers.

²² We assume that the supermarkup applies equally to all coalition products in its region for simplicity and computational tractability. In the empirical application, we consider a robustness exercise in which different supermarkups apply to products of different package sizes (online Appendix D).

where $mc_{jrt}(\cdot)$ is a constant marginal cost function and $q_{jrt}(\cdot)$ is a differentiable demand function. Firms apply a discount factor, $\delta \in (0, 1)$, to calculate present values. We assume that the economic state is common knowledge and that firms observe all previous states, prices, quantities, and supermarkup announcements. These elements constitute the history that is considered in stage (ii) of each period. Firm actions do not affect the economic state. The equilibrium concept is subgame perfection.

We turn next to the building blocks of the price leadership model: static first order conditions, slack functions, and the leader's maximization problem. We then provide the strategies which constitute the *price leadership equilibrium* (PLE). Finally, we relate our model to the theoretical literature.

B. Static First Order Conditions

Differentiating equation (1) with respect to price obtains static first order conditions for (stage-game) profit maximization. Let $(\mathbf{p}_{irt}, \mathbf{q}_{irt}, \mathbf{mc}_{irt})$ be vectors of firm i 's prices, quantities, and costs in region r and period t , and let \mathbf{p}_{-irt} contain the prices of competitors. Then the static first order conditions of firm i take the form

$$(2) \quad f(\mathbf{p}_{irt}, \mathbf{p}_{-irt}; \Psi_t) \equiv \mathbf{p}_{irt} + \left[\frac{\partial \mathbf{q}_{irt}(\mathbf{p}_{irt}, \mathbf{p}_{-irt}; \Psi_t)}{\partial \mathbf{p}_{irt}} \right]^{-1} \mathbf{q}_{irt}(\mathbf{p}_{irt}, \mathbf{p}_{-irt}; \Psi_t) - \mathbf{mc}_{irt}(\Psi_t) = 0.$$

We assume there exists a unique $\mathbf{p}_{irt}^*(\mathbf{p}_{-irt}; \Psi_t)$ that solves this system of equations, for each firm i and any competitor prices. This assumption can be verified for the special case of logit demand by adapting an argument of Nocke and Schutz (2018). A number of coding checks suggest existence and uniqueness in our empirical application, but with the RCNL demand system this is not guaranteed. If *all* firms solve their static first order conditions then Bertrand equilibrium obtains, with $\mathbf{p}_{irt}^*(\mathbf{p}_{-irt}^*; \Psi_t)$ for all i . We collect the Bertrand prices in the vector \mathbf{p}_{rt}^B for notational convenience.

Differentiated-products Bertrand pricing often is assumed in the empirical literature. If demand is known and the data contain Bertrand prices and quantities, then equation (2) identifies marginal costs (Rosse 1970). In our price leadership model, the static first order conditions also have empirical content. Take as given that, along the equilibrium path, coalition firms set prices according to $\mathbf{p}_{irt}^{PL}(m_{rt}) = \mathbf{p}_{irt}^B + m_{rt}$, and fringe firms solve equation (2) holding fixed the coalition prices. Then marginal costs can be recovered for any given m_{rt} . We formalize the result with the following proposition.

PROPOSITION 1 (Identification of Marginal Costs): *If the econometrician has knowledge of price leadership prices, the demand system, the identities of the coalition firms, and the supermarkup, 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 empirical implementation:

- (i) Infer mc_{jrt} for the products of fringe firms from equation (2). This can be done because fringe firms maximize stage-game profit.
- (ii) Obtain $p_{jrt}^B = p_{jrt}^{PL}(m_{rt}) - m_{rt}$ for the products of coalition firms.
- (iii) Compute p_{jrt}^B for the products of fringe firms by simultaneously solving the best response function, given the marginal costs inferred from step (i) and holding the prices of coalition firms fixed at the Bertrand levels obtained from step (ii).
- (iv) Infer mc_{jrt} for the products of coalition firms from equation (2), evaluated at the Bertrand prices \mathbf{p}_{rt}^B obtained in steps (ii) and (iii). ■

C. Slack Functions

We now define the *slack function* of each firm, which provides the net present value of price leadership less that of deviation, given a set of supermarkups. We use the slack function to characterize the IC constraints in the leader's maximization problem. For present purposes, we assume that deviation profit is earned for $\tau_1 \geq 1$ periods, and that punishment takes the form of Bertrand pricing for $\tau_2 \geq 1$ periods. We verify later that these timing assumptions are consistent with the strategies that characterize the PLE.²³

Some additional notation is necessary. Let the profit that firm i receives with price leadership be $\pi_{irt}^{PL}(m_{rt}; \Psi_t) \equiv \pi_{irt}(\mathbf{p}_{rt}^{PL}(m_{rt}); \Psi_t)$, where the prices in $\mathbf{p}_{rt}^{PL}(m_{rt})$ are as defined in the previous subsection. Let the deviation prices of firm i be those that arise if firm i solves its static first order conditions (equation (2)) and other firms set prices according to $\mathbf{p}_{rt}^{PL}(m_{rt})$. Denoting these prices as $\mathbf{p}_{rt}^{D,i}(m_{rt})$, we have deviation profit of $\pi_{irt}^{D,i}(m_{rt}; \Psi_t) \equiv \pi_{irt}(\mathbf{p}_{rt}^{D,i}(m_{rt}); \Psi_t)$. Finally, let profit in Bertrand equilibrium be $\pi_{irt}^B(\Psi_t) \equiv \pi_{irt}(\mathbf{p}_{rt}^B; \Psi_t)$.

The contribution of region r to the slack function of firm i is given by

$$(3) \quad \tilde{g}_{irt}(m_{rt}; \delta, \phi, \tau_1, \tau_2, \tilde{m}_{rt}, \Psi_t) \equiv \pi_{irt}^{PL}(m_{rt}; \Psi_t) + \frac{\phi\delta}{1 - \phi\delta} \pi_{irt}^{PL}(\tilde{m}_{rt}; \Psi_t) - \left(\sum_{s=0}^{\tau_1-1} (\phi\delta)^s \pi_{irt}^{D,i}(m_{rt}; \Psi_t) + \sum_{s=\tau_1}^{\tau_1+\tau_2-1} (\phi\delta)^s \pi_{irt}^B(\Psi_t) + \sum_{s=\tau_1+\tau_2}^{\infty} (\phi\delta)^s \pi_{irt}^{PL}(\tilde{m}_{rt}; \Psi_t) \right),$$

²³This rules out optimal punishments (e.g., Abreu 1986). Even with explicit collusion, often cartels do not employ complex punishments, other than making transfers (Harrington and Skrzypacz 2011).

where \tilde{m} refers to the supermarkup expected in future periods involving price leadership. The first line is the present value of price leadership in the region, and the next lines subtract the present value of deviation followed by punishment and an (eventual) return to price leadership. Summing across regions, the slack function is given by

$$(4) \quad \tilde{g}_{it}(\mathbf{m}_t; \delta, \phi, \tau_1, \tau_2, \tilde{\mathbf{m}}_t, \Psi_t) \equiv \sum_{r=1}^R \tilde{g}_{irt}(m_{rt}; \delta, \phi, \tau_1, \tau_2, \tilde{m}_{rt}, \Psi_t)$$

for the vectors of supermarkups $\mathbf{m}_t = (m_{1t}, m_{2t}, \dots, m_{Rt})$ and expected supermarkups $\tilde{\mathbf{m}}_t = (\tilde{m}_{1t}, \tilde{m}_{2t}, \dots, \tilde{m}_{Rt})$. The slack function is positive if the present value of price leadership exceeds the present value of deviation, and negative otherwise.

Using the slack function directly in empirical work presents identification challenges. Suppose that the profit terms are known and that $\tilde{g}_{it}(\mathbf{m}_t; \cdot) = 0$ for some coalition firm i . As the discount factor δ and the continuation probability ϕ enter as multiplicative factors, knowledge that $\tilde{g}_{it}(\mathbf{m}_t; \cdot) = 0$ is insufficient to disentangle the two parameters. Further, it can be verified that for any (τ_1, τ_2) there exists some $\phi\delta$ that satisfies the equality, meaning that $\phi\delta$ cannot be separately identified from (τ_1, τ_2) on the basis of $\tilde{g}_{it}(\mathbf{m}_t; \cdot) = 0$ alone. Complicating matters is that deviation and punishment do not occur on the equilibrium path, a result that we develop shortly, so τ_1 and τ_2 cannot be discerned from data on equilibrium outcomes. Absent other evidence about the strategies played by firms off the equilibrium path, $\tilde{g}_{it}(\mathbf{m}_t; \cdot) = 0$ provides only joint identification of $(\delta, \phi, \tau_1, \tau_2)$.

To help facilitate empirical progress, we construct an *equivalent slack function* in which the intertemporal trade off is governed by a single reduced-form timing parameter. The equivalent slack function is

$$(5) \quad g_{it}(\mathbf{m}_t; \eta, \tilde{\mathbf{m}}_t, \Psi_t) \equiv \sum_{r=1}^R g_{irt}(m_{rt}; \eta, \tilde{m}_{rt}, \Psi_t),$$

where

$$(6) \quad g_{irt}(m_{rt}; \eta, \tilde{m}_{rt}, \Psi_t) \equiv \pi_{irt}^{PL}(m_{rt}; \Psi_t) + \frac{\eta}{1-\eta} \pi_{irt}^{PL}(\tilde{m}_{rt}; \Psi_t) - \left(\pi_{irt}^{D,i}(m_{rt}; \Psi_t) + \frac{\eta}{1-\eta} \pi_{irt}^B(\Psi_t) \right),$$

and

$$(7) \quad \eta \equiv \eta(\delta, \phi, \tau_1, \tau_2) = \frac{(\phi\delta)^{\tau_1} - (\phi\delta)^{\tau_1+\tau_2}}{1 - (\phi\delta)^{\tau_1+\tau_2}}.$$

The equivalent slack function provides the net present value of price leadership less that of deviation in an infinitely repeated version of the game featuring a single period of deviation profit and grim trigger punishment strategies.

We now clarify our notion of equivalence formally.

PROPOSITION 2 (Equivalent Slack Function): *The slack function and the equivalent slack function are related according to*

$$\tilde{g}_{it}(\mathbf{m}_t; \delta, \phi, \tau_1, \tau_2, \tilde{\mathbf{m}}_t, \Psi_t) = \psi g_{it}(m_{rt}; \eta, \tilde{\mathbf{m}}_t, \Psi_t)$$

for some $\psi \in (0, 1]$. Further, if $\tau_1 = 1$ then $\psi = 1$.

PROOF:

See online Appendix B.

The two versions of the slack function both provide valid characterizations of the IC constraints because they share the same sign for any vector of supermarkups. Further, the timing parameter that solves the equation $g_{it}(\mathbf{m}_t; \eta, \tilde{\mathbf{m}}_t, \Psi_t) = 0$ summarizes the patience of firms, the continuation probability, and the durations of deviation and punishment that would solve the equation $\tilde{g}_{it}(\mathbf{m}_t; \delta, \phi, \tau_1, \tau_2, \tilde{\mathbf{m}}_t, \Psi_t) = 0$.

Two observations are appropriate at this point. First, our construction of the slack functions assumes that firms do not anticipate changes in the economic state. The assumption appears necessary for empirical work because the main alternatives are untenable—it is not clear how to model the processes that govern costs and demand, and a purely empirical approach is impossible because an infinite series of data would be required. A consequence is that the model does not generate countercyclical pricing à la Rotemberg and Saloner (1986), which arises because the present value of deviation is smaller if current demand is weak relative to future demand. However, we can show that if the profit function grows at a constant rate then the equivalent slack function obtains nonetheless.

The second observation is that the slack functions allow the leader's announced supermarkups (\mathbf{m}_t) to differ from expectations ($\tilde{\mathbf{m}}_t$) about what the leader will announce in the future, with the same economic state. Firms' expectations play an important role in the slack functions and, ultimately, the IC constraints. For example, if firms expect supermarkups of zero in the future, then they also expect punishment for deviation to be inconsequential, and it follows that positive supermarkups cannot be sustained.²⁴ To make progress, we assume that expectations are rational in a sense that we formalize in the next subsection.

D. The Leader's Problem

We assume the leader announces supermarkups that maximize its profit subject to IC constraints. Without loss of generality, let firm 1 be the leader. Each period the leader chooses markups for each region, $\mathbf{m}_t^* = (m_{1t}^*, m_{2t}^*, \dots, m_{Rt}^*)$, given that each firms' beliefs about future markups are fixed. The solution is then given by

$$(8) \quad \mathbf{m}_t^*(\tilde{\mathbf{m}}_t, \eta, \Psi_t) = \underset{\mathbf{m}_t \geq 0}{\operatorname{argmax}} \sum_{r=1}^R \pi_{1rt}^{PL}(m_{rt}; \Psi_t),$$

²⁴By inspection, we have $g_{irt}(\mathbf{m}_t) \leq 0$ for any m_{rt} in that case.

subject to

$$g_{it}(\mathbf{m}_t; \eta, \tilde{\mathbf{m}}_t, \Psi_t) \geq 0 \quad \forall i.$$

We require rational expectations, so that firms' expectations for future supermarkups align with the solution to the leader's constrained maximization problem given those expectations (i.e., $\mathbf{m}_t^*(\tilde{\mathbf{m}}_t, \eta, \Psi_t) = \tilde{\mathbf{m}}_t$). We assume the existence of a unique nondegenerate rational expectations solution.²⁵ We drop $\tilde{\mathbf{m}}_t$ as a function argument henceforth, as we have $\mathbf{m}_t^* = \tilde{\mathbf{m}}_t$. For empirical purposes, it is helpful to derive the first order conditions that characterize the rational expectations solution. This can be done under two mutually exclusive assumptions.

ASSUMPTION 1: *IC does not constrain the leader, i.e., $g_{it}(\mathbf{m}_t^*; \eta, \Psi_t) > 0, \forall i$.*

ASSUMPTION 2: *IC constrains the leader, i.e., $g_{kt}(\mathbf{m}_t^*; \eta, \Psi_t) = 0$ for some $k \in \mathcal{C}$.*

Under Assumption 1, the leader sets the supermarkups to maximize its profit, and there are region-specific first order conditions:

$$(9) \quad \tilde{h}_{rt}(m_{rt}; \Psi_t) \equiv \frac{\partial \pi_{1rt}^{PL}(m_{rt}; \Psi_t)}{\partial m_{rt}} = 0 \quad \text{for } r = 1, 2, \dots, R.$$

By inspection, the supermarkups that solve the leader's problem depend on the economic state, Ψ_t , but not on the timing parameter, η . The timing parameter enters the maximization problem only through the constraints, which do not bind in this case.

Under Assumption 2 there is a binding constraint, and the solution is characterized by the constraint itself and a series of $R - 1$ Kuhn-Tucker balancing equations. Without loss of generality, let the firm with the binding IC constraint be firm $k \in \mathcal{C}$ (this could be any coalition firm, including the leader). Then we obtain

$$(10) \quad g_{kt}(\mathbf{m}_t; \eta, \Psi_t) = 0,$$

and for each $r = 1, 2, \dots, R - 1$,

$$(11) \quad h_{rt}(\mathbf{m}_t; \Psi_t) \equiv \frac{\frac{\partial \pi_{1rt}^{PL}(m_{rt}; \Psi_t)}{\partial m_{rt}}}{\frac{\partial \pi_{krt}^{D,k}(m_{rt}; \Psi_t)}{\partial m_{rt}} - \frac{\partial \pi_{krt}^{PL}(m_{rt}; \Psi_t)}{\partial m_{rt}}} - \frac{\frac{\partial \pi_{1Rt}^{PL}(m_{Rt}; \Psi_t)}{\partial m_{Rt}}}{\frac{\partial \pi_{kRt}^{D,k}(m_{Rt}; \Psi_t)}{\partial m_{Rt}} - \frac{\partial \pi_{kRt}^{PL}(m_{Rt}; \Psi_t)}{\partial m_{Rt}}} = 0.$$

Equations (10) and (11) can be obtained with some algebra, after differentiating the Lagrangian of the leader's problem (online Appendix B). The numerators in the balancing equations are the benefit to the leader of a higher supermarkup; the denominators capture the shadow cost of a higher supermarkup, which arises from the binding firm's IC constraint. The solution depends on both the economic state and the timing parameter.

²⁵The degenerate solution involves $\mathbf{m}_t^* = \tilde{\mathbf{m}}_t = 0$.

E. Price Leadership Equilibrium

We now formally define the *price leadership equilibrium* (PLE), which is a subgame perfect equilibrium (SPE). The leader's strategy is $\sigma_1: \mathcal{H} \rightarrow \mathcal{M} \times \mathcal{R}^{J_1}$, where \mathcal{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 $i = 2, \dots, F$ are $\sigma_i: \mathcal{M} \times \mathcal{H} \rightarrow \mathcal{R}^{J_i}$.

For simplicity, we first examine grim trigger strategies in the infinitely repeated version of the game, and then map the result into the full model using Proposition 2. In the pricing subgame, let firms price according to \mathbf{p}_t^B if $g_{it}(\mathbf{m}_t; \eta, \Psi_t) < 0$ for any i or if, in any previous period s , prices differ from $p_s^{PL}(\mathbf{m}_s)$. Otherwise let firms price according to $p_t^{PL}(\mathbf{m}_t)$. It is easily verified that no firm can unilaterally improve its payoff in the pricing stage. For example, no firm would set deviation prices, $p_t^{D,i}(\mathbf{m}_t)$, because if any firm prefers deviation then this is known by all firms and play shifts to Bertrand. In the announcement stage, let the leader set supermarkups, $\mathbf{m}_t^*(\eta, \Psi_t)$, that solve its constrained maximization problem, taking as given rational expectations and the strategies proposed above for the pricing stage. This maximizes the leader's payoff, by construction, so the leader has no incentive to do otherwise. It also ensures that deviation does not occur in the pricing stage. Thus, the stated strategies constitute an SPE, and we label it the PLE.

An implication of Proposition 2 is an alternative set of strategies featuring τ_1 periods of deviation and τ_2 periods of punishment supports identical play along the equilibrium path. A full characterization of this equilibrium is unnecessary for our purposes, though we note that some pricing friction must exist in order to explain why punishment would not ensue immediately after deviation. Finally, it is worth highlighting that the PLE might not be Pareto optimal for the coalition firms because the leader acts in its own interest and side payments—which violate antitrust statutes—are not incorporated.²⁶

F. Relationship to the Theoretical Literature

The price leadership model resembles the canonical Rotemberg and Saloner (1986) model of collusion because information is perfect and deviation does not occur along the equilibrium path. The main distinction is that we incorporate the idea that price signaling can help support supracompetitive prices. The conditions under which it is reasonable to assume that beliefs respond to cheap talk, such as the leader's price announcement, are discussed in the literature (e.g., Aumann 1990, Farrell and Rabin 1996).²⁷ Harrington, Gonzalez, and Kujal (2016) develops experimental evidence that price announcements can help facilitate coordination in

²⁶ See Asker (2010) and Asker, Collard-Wexler, and De Loecker (2019) for two empirical examples of inefficient coordination.

²⁷ 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.

repeated oligopoly games, and qualitative evidence previously described provides support for our approach in the context of the beer industry.

To incorporate multiple regions and the possibility that multi-market contact affects equilibrium outcomes, we extend the two-firm, two-region model of Bernheim and Whinston (1990).²⁸ We maintain the assumption that IC constraints are pooled across regions because it generates leader profit at least weakly larger than the alternative of independent regions. That said, the feasibility of the pooled solution depends on the sophistication of firms, into which we have little visibility (insofar as it relates to multi-market contact). It is also possible to make empirical progress with independent regions, as the supermarkups that solve the constrained maximization problem would be determined by

$$(12) \quad g_{krt}(m_{rt}^*, \eta, \Psi_t) = 0,$$

where firm $k \in \mathcal{C}$ has the binding constraint. Later in this article, we explore the implications of multi-market contact in a counterfactual simulation. We also provide imputation results for the case of region-specific IC constraints (online Appendix C).

The price leadership model incorporates firm heterogeneity, which is important for applied work in most industries. This raises questions about which firms participate in the coalition and, among these, which is the leader. We impose *ex ante* that the coalition includes ABI, Miller, and Coors (or MillerCoors) and that ABI is the leader, an approach that is supported by the available qualitative and empirical evidence. In principle, a more theoretical approach to the coalition could be taken, under the assumption that each firm faces a decision whether to join the coalition (e.g., as in d'Aspremont et al. 1983; Donsimoni, Economides, and Polemarchakis 1986; and Bos and Harrington 2010). Similarly, in price leadership models slightly different than ours, Pastine and Pastine (2004) allows a war of attrition to determine the leader, and Ishibashi (2008) and Mouraviev and Rey (2011) show that coalition profits are higher if the leader is the firm with the greatest incentive to deviate.²⁹ As multiple theoretical assumptions appear to be available, we prefer the empirical approach.

Finally, our model abstracts away the retail and distribution sectors. However, it is isomorphic to a model that incorporates constant markups, or “cost-plus” pricing, downstream. The reason is that downstream markups and brewer marginal costs enter the profit functions in the same way, so that downstream markups are equivalent to a tax on production (online Appendix F). Recent research provides some support for cost-plus pricing among retailers in scanner data similar to ours.³⁰ The model would be misspecified for settings in which retailers exercise buyer power to obtain lower prices (e.g., as in Loertscher and Marx 2019). In our setting, buyer

²⁸ See Section VII on differentiated products. The necessary conditions are analogous to the Kuhn-Tucker balancing equations that characterize the solution to the leader's constrained maximization problem.

²⁹ Our model differs in that each period features an announcement followed by simultaneous pricing, rather than sequential pricing.

³⁰ DellaVigna and Gentzkow (2019) shows that retail prices often do not respond to local demand shocks; and Butters, Sacks, and Seo (2020) documents that retail prices change one for one with local cost shocks (generated by excise taxes). This combination would arise from cost-plus pricing.

power may be limited by the lack of private label store brands and the regulatory prohibition on slotting allowances, which makes it harder for retailers to discipline coordination by auctioning shelf space. Nonetheless, downstream markups appear to be sizable. Tremblay and Tremblay (2005) places them at about 35 percent of total revenue, based on industry studies conducted in the 1970s and 1990s, such that brewer revenues would be about 65 percent of total revenue.

III. Empirical Implementation

A. Assumptions and Imputation Algorithms

We make a number of assumptions to map the data into the model of price leadership. First, we assume that firms set prices according to the PLE, so that the price vector $\mathbf{p}_t^{PL}(\Psi_t)$ is observed in the data. Second, we assume that the coalition comprises ABI (the leader) and either Miller and Coors (in 2006 and 2007) or MillerCoors (in 2010 and 2011). These first two assumptions are supported by the qualitative evidence in Section I. Third, we assume that demand is given by the RCNL-2 model estimated in Miller and Weinberg (2017). Fourth, we assume that each period ($t = 1, \dots, \infty$) in the theoretical model corresponds to a fiscal year in the data. Our treatment incorporates that demand and cost conditions change within the fiscal year, while the supermarkups are fixed. As the empirical implementation uses the equivalent slack function, this also implies that deviation lasts for one year before punishment ensues. If deviation is longer or shorter than one year in reality, this would be subsumed into the timing parameter.

Marginal costs, Bertrand prices, and the supermarkups can be recovered from the data by applying the structure of the model. We have already described how marginal costs and Bertrand prices can be obtained given the supermarkup (Proposition 1). Thus, we focus on inferring the supermarkups from the leader's constrained maximization problem. Recall that the leader's first order conditions depend on (i) whether an IC constraint binds and (ii) in the case of a binding constraint, the value of the timing parameter. Our general approach is to make *ex ante* assumptions on (i) and (ii) and then evaluate these assumptions *ex post*. The algorithms that recover the supermarkup depend on the assumptions.

The constrained case is the most demanding from a computational standpoint because it requires solving a system of R nonlinear equations with R unknowns (with $R = 37$ in our application). We initially consider $\eta = (0.20, 0.25, 0.30, 0.35, 0.40)$, based on our experience with the data and the model. For each of these timing parameters in turn, we apply the following algorithm to each period:

- (i) Consider a candidate supermarkup vector, $\hat{\mathbf{m}}_t$.
- (ii) Obtain implied marginal costs and Bertrand prices (Proposition 1).
- (iii) Obtain $g_{it}(\hat{\mathbf{m}}_t; \eta, \Psi_t)$ for all $i \in \mathcal{C}$ by computing deviation prices using the static first order conditions (equation (2)), calculating profit under price leadership, deviation, and Bertrand, and then applying the proposed timing parameter.

- (iv) Identify the binding firm, k , such that $g_{kt}(\hat{\mathbf{m}}_t; \eta, \Psi_t) = \min_{i \in \mathcal{C}} g_{it}(\hat{\mathbf{m}}_t; \eta, \Psi_t)$.
- (v) Obtain $h(\hat{m}_{rt}; \Psi_t)$ for $r = 1, \dots, R - 1$, by numerically differentiating the price leadership profit of the leader and the binding firm, as well as the deviation profit of the binding firm, with respect to the supermarkups.
- (vi) Assess the loss function, which we construct based on the observation that if $\hat{\mathbf{m}}_t = \mathbf{m}_t$ then $g_{kt}(\hat{\mathbf{m}}_t; \eta, \Psi_t) = 0$ and $h_{rt}(\hat{\mathbf{m}}_t; \Psi_t) = 0$ for $r = 0, 1, \dots, R - 1$.
- (vii) Update $\hat{\mathbf{m}}_t$ if needed, and repeat to convergence.

The computational burden of such an algorithm typically increases nonlinearly in the number of unknowns (here, supermarkups). Thus, we develop an approach that exploits the structure of the model, and which makes the computational burden roughly linear in the number of supermarkups. The details are provided in online Appendix C.

The unconstrained case is simpler computationally because each region can be considered in isolation, yielding R distinct problems, each with one equation and one unknown. We apply a standard equation solver to identify the supermarkup, yielding $\tilde{h}_{rt}(m_{rt}; \Psi_t) = 0$. Finally, we also consider the case of Bertrand competition, for which marginal costs can be recovered from the static first order conditions of equation (2).

B. Identification

Each of the ex ante modeling assumptions considered above implies corresponding supermarkups and marginal costs, and this provides a path to ex post model selection. We observe that timing parameters closer to one imply higher supermarkups because they create more slack in the IC constraints. In turn, higher supermarkups imply lower marginal costs for coalition firms, holding fixed the observed prices. We illustrate this latter point in Figure 2, which plots the Bertrand prices and marginal costs one would infer from different supermarkups. Putting these observations together, in our application there is a one-to-one mapping between the timing parameter and both (i) the costs of coalition firms and (ii) the relative costs of coalition and fringe firms. This can allow the relative quality of the ex ante modeling assumptions to be evaluated.³¹

In particular, if one has prior knowledge of marginal costs, then the timing parameter and the associated supermarkups are identified. We see at least two workable approaches. The first exploits engineering or business data on the magnitude of marginal costs (e.g., as in Nevo 2001), and simply selects the ex ante modeling assumptions that imply the correct marginal costs. This does not require multiple periods or regions, and may be particularly useful in merger review because anti-trust authorities can use subpoena power to gain access to confidential business

³¹ We suspect this identification argument would extend in similar applications. Fringe costs are invariant to the candidate supermarkup, which can be seen in Figure 2 and verified with the proof to Proposition 1.

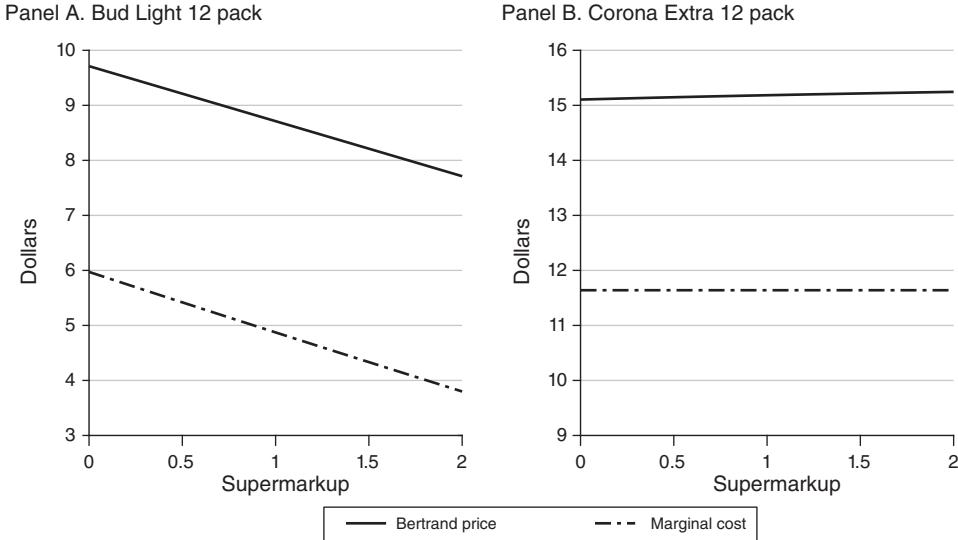


FIGURE 2. BERTRAND PRICES AND MARGINAL COSTS UNDER DIFFERENT SUPERMARKUPS

Notes: The figures plot the average Bertrand prices and supermarkups that one would infer given candidate supermarkups in the range of $m \in [0, 2]$. The statistics are obtained for Bud Light 12 packs (panel A) and Corona Extra 12 packs (panel B) in the fiscal year 2007.

documents. Unfortunately, we do not have access to publicly available information on the marginal costs of the beers in our sample.

We take a second approach, which involves applying prior knowledge about how marginal costs change with the economic state. To implement this, we specify the following marginal cost function for product j in region r , fiscal year t , and quarter s :

$$\begin{aligned}
 (13) \quad mc_{jrt} &= \beta_1 \mathbf{1}\{j \in ABI\} \times \mathbf{1}\{t \in (2010, 2011)\} \\
 &\quad + \beta_2 \mathbf{1}\{j \in Miller\} \times \mathbf{1}\{t \in (2010, 2011)\} \\
 &\quad + \beta_3 \mathbf{1}\{j \in Coors\} \times \mathbf{1}\{t \in (2010, 2011)\} \\
 &\quad + \beta_4 Distance_{jrt} + \mu_j + \mu_r + \mu_{ts} + \epsilon_{jrt}.
 \end{aligned}$$

The first three terms allow for the marginal costs of ABI, Miller, and Coors products to shift after the Miller-Coors merger. Distance is between the brewery and the region in question, and accounts for transportation costs. The terms μ_j , μ_r , and μ_{ts} are product, region, and time (year \times quarter) fixed effects, respectively. Of particular relevance for our identification strategy, the time effects account for any changes in the prices of labor and ingredients that affect all firms equally.³²

³²There are four fiscal years, each with four quarters, so there are 16 time fixed effects.

Our identifying assumption is $\beta_1 = 0$, which yields the marginal cost function of Miller and Weinberg (2017). Given the presence of product and time fixed effects, the identifying assumption implies that ABI's marginal costs do not change differently from those of Modelo and Heineken, on average, with the 2008 Miller-Coors merger. Equivalently, recalling Figure 1, the assumption is that the true model generates the ABI price increases that are observed in the data after the Miller-Coors merger without ABI-specific cost increases.

We estimate the marginal cost function with ordinary least squares (OLS) after stacking the observations from each fiscal year. This allows us to evaluate the *ex ante* modeling assumptions on the IC constraints, and reject those that generate a $\hat{\beta}_1$ that is statistically different than zero. Further, we select as the preferred assumptions those that generate the $\hat{\beta}_1$ that is closest to zero. In principle, one could estimate the timing parameter using a nested method-of-moments procedure, but that is computationally infeasible in our context.³³

We view the identifying assumption as a reasonable approximation given the institutional details of the market. There are a number of reasons the assumption might not hold exactly. Changes in the prices of ingredients (e.g., water, hops, barley, and yeast) would affect ABI costs differently than import costs if ingredients are used in different proportions. Similarly, it is feasible that the wages paid by ABI could trend differently from those paid by Modelo and Heineken, especially as their brewing occurs in different countries. However, we are skeptical that changes in input prices would substantially impact the validity of the identifying assumption. For example, with respect to ingredients, the brands in our sample are all simple lagers, so proportions are likely to be roughly similar and, furthermore, some information suggests that ingredient and labor costs are small relative to the retail price.³⁴

Also worth discussing at this point is the InBev acquisition of Anheuser-Busch, which closed in 2009. The merger was cleared by the DOJ after minor divestitures because, in most local markets, there was little competition between Anheuser-Busch and InBev.³⁵ As best we can discern, the postacquisition efforts to reduce costs would not have affected marginal costs. InBev revised the pay system, ended pension contributions and life insurance for retirees, and transferred all foreign beer operations to InBev (Ascher 2012). The domestic distribution of the brands in our sample was unaffected, so transportation cost would not have changed, and in any event the cost specification controls for distance. An interesting possibility is that the acquisition could have affected buyer power in the hops market, reducing ABI's input costs relative those of its competitors.³⁶ However, corroborating our understanding of

³³ If we run the imputation algorithm for constrained price leadership on a single processor, it takes three to four weeks to converge, depending on the timing parameter. A number of factors contribute to the time costs, including the number of regions and the 500 demographic draws used in the RCNL demand system. We reduce the computational burden somewhat by using different processors for different fiscal years.

³⁴ We draw on Tremblay and Tremblay (2005) which places ingredients costs at 4–6 percent of retail revenues and labor costs as 5–12 percent of retail revenues, based on industry studies from the 1970s and 1990s. According to Tremblay and Tremblay, distribution and packaging accounts for the bulk of brewer costs.

³⁵ The DOJ press release can be obtained at https://www.justice.gov/archive/atr/public/press_releases/2008/239430.htm, last accessed March 16, 2021.

³⁶ Our marginal cost specification accounts for the analogous possibility that the Miller-Coors merger amplified buyer power in the hops market, through the β_2 and β_3 parameters.

TABLE 3—MODEL SELECTION AND IMPLIED SUPERMARKUPS

	Bertrand	$\eta = 0.20$	$\eta = 0.25$	$\eta = 0.26$	$\eta = 0.30$	$\eta = 0.35$	$\eta = 0.40$	Unconstrained
<i>Panel A. OLS regression results</i>								
β_1	0.657 (0.094)	0.204 (0.117)	0.024 (0.122)	-0.017 (0.123)	-0.203 (0.125)	-0.492 (0.127)	-0.858 (0.126)	0.607 (0.147)
β_2	0.189 (0.074)	-0.286 (0.096)	-0.476 (0.102)	-0.519 (0.103)	-0.716 (0.107)	-1.023 (0.113)	-1.411 (0.137)	-0.017 (0.137)
β_3	-0.067 (0.095)	-0.557 (0.107)	-0.748 (0.113)	-0.791 (0.114)	-0.987 (0.120)	-1.288 (0.129)	-1.663 (0.173)	-0.242 (0.173)
β_4	0.220 (0.066)	0.229 (0.065)	0.234 (0.066)	0.235 (0.067)	0.240 (0.069)	0.251 (0.073)	0.269 (0.082)	0.340 (0.120)
<i>Panel B. Other statistics</i>								
\bar{m}_{2006}	0	0.84	1.10	1.15	1.39	1.71	2.07	4.51
\bar{m}_{2007}	0	0.87	1.14	1.20	1.44	1.77	2.15	4.65
\bar{m}_{2010}	0	1.27	1.70	1.80	2.22	2.83	3.57	4.64
\bar{m}_{2011}	0	1.31	1.75	1.85	2.28	2.90	3.64	4.79
$mc < 0$	<0.001	0.002	0.005	0.006	0.011	0.023	0.050	0.170

Notes: Panel A summarizes the results from OLS estimation of the marginal cost function, based on the marginal costs obtained under different modeling assumptions. There are 20,162 observations at the product-region-fiscal year-quarter level, combining data from fiscal years 2006, 2007, 2010, and 2011. The intermediate timing parameters are generated under the assumption that the IC constraint binds. Regressors include indicators for ABI brands, Miller brands, and Coors brands, in the fiscal years 2010 and 2011 (corresponding to β_1 , β_2 , and β_3), distance from the brewery to the region (β_4), as well as product, region, and time fixed effects. Standard errors are clustered at the region level and shown in parentheses. Panel B provides the average supermarkups by fiscal year, and the proportion of marginal cost values which are negative.

the acquisition, we have been unable to find references to substantial variable costs changes in the ABI annual reports or the popular press.

IV. Results

A. Model Selection

Table 3 summarizes the results of the model selection exercise. Each column corresponds to one set of *ex ante* modeling assumptions. Results for Bertrand and unconstrained price leadership are shown in the leftmost and rightmost columns, respectively, and results for constrained price leadership under various timing parameters are shown in the middle columns. Panel A provides the main regression coefficients that we obtain from OLS estimation of the marginal cost function. Standard errors are clustered at the region level and shown in parentheses. Panel B provides the average supermarkups, $\bar{m}_t = (1/R)\sum_r m_{rt}$, and the proportion of marginal costs that are negative.

The timing parameter of $\eta = 0.26$ generates the $\hat{\beta}_1$ closest to zero, and so it is our preferred model.³⁷ The Bertrand and unconstrained price leadership models are easily rejected. There are timing parameters near $\eta = 0.26$ for which the null of $\beta_1 = 0$ cannot be rejected, with $\eta = 0.20$ and $\eta = 0.30$ being near the boundaries of this range.³⁸ Among the constrained models, the average supermarkups increase

³⁷ After our examination of the initial timing parameters $\eta = (0.20, 0.25, 0.30, 0.35, 0.40)$, we determined that $\eta = 0.26$ would probably generate the OLS estimate of β_1 closest to zero, and indeed that is the case.

³⁸ The relevant *t*-statistics are -1.74 for the $\eta = 0.20$ model and 1.62 for the $\eta = 0.30$ model.

TABLE 4—WELFARE EFFECTS OF PRICE LEADERSHIP

	2006	2007	2010	2011
Producer surplus				
Δ Total profit (%)	17.28	17.43	22.18	22.10
Consumer surplus (CS)				
Δ CS/ Δ Total profit	−1.55	−1.62	−1.84	−1.84

Notes: The table summarizes the effects of price leadership, based on a comparison of the observed equilibrium to a counterfactual that features Bertrand prices. The profit statistics are calculated by aggregating product-region-quarter observations, subtracting Bertrand profit from price leadership profit and then dividing by Bertrand profit. The consumer surplus statistics are calculated by aggregating over region-quarter combinations, and then normalizing by the change in total profit.

with the timing parameter, as does the average change before versus after the Miller-Coors merger. This pattern arises because more slack is inferred in IC constraints (for any given price) if the timing parameter is greater. At the preferred model, we infer average supermarkups of \$1.15, \$1.20, \$1.80, and \$1.85 in the four fiscal years, respectively.³⁹

Turning to the remaining parameters, we estimate that the marginal cost intercepts of Miller and Coors decrease with the joint venture by \$0.52 and \$0.79 (β_2 and β_3) with the preferred model. As the distance estimate is positive (β_4), 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) estimates similar marginal cost parameters and analyzes the efficiencies in greater detail.

Table 4 analyzes the welfare consequences of price leadership. We benchmark against Bertrand equilibrium, which we compute holding fixed the imputed marginal costs from the preferred model. Across all firms, we find that the difference in profit under price leadership and Bertrand is about 17 percent of Bertrand profits in fiscal years 2006 and 2007, and 22 percent in fiscal years 2010 and 2011. The adverse impact of price leadership on consumer surplus is about 1.6 times larger than this profit gain in 2006 and 2007, and about 1.8 times larger in 2010 and 2011.⁴⁰ Thus, the loss of total welfare is just greater than half of the profit gain.

Table 5 provides the average markup for each product in the data in fiscal years 2007 and 2010, again based on the preferred model. Across all 20,162 brand-size-region-quarter-year observations, the median markup is \$4.53 on an equivalent-unit basis, and accounts for 44 percent of the retail price.⁴¹ The average

³⁹ Unconstrained price leadership generates the largest implied supermarkups. These supermarkups do not change much with the Miller-Coors merger because there is only a small effect on ABI's preferred supermarkup. This also explains why $\beta_1 > 0$ for that model—without an increase in supermarkups, the model requires ABI-specific marginal cost increases to rationalize the price patterns of Figure 1.

⁴⁰ We report effects in this manner because consumer surplus is identified only up to an additive constant, so we cannot recover percentage changes.

⁴¹ Tremblay and Tremblay (2005) reports that downstream markups account for about 35 percent of total revenue. The brewer markups we recover would account for 63 percent of the remainder, on average, which would correspond to the margins or Lerner index. In our experience, this is somewhat high but not exceptional for consumer products, and is certainly plausible with suprareactive pricing. As a point of comparison, Nevo (2001) reports

TABLE 5—BREWER MARKUPS

Brand	6 packs		12 packs		24 packs	
	2007	2010	2007	2010	2007	2010
Bud Light	5.19	5.73	5.04	5.61	4.92	5.48
Budweiser	5.35	5.87	5.19	5.74	5.04	5.65
Coors	4.15	5.72	3.97	5.61	3.82	5.45
Coors Light	3.90	5.55	3.77	5.44	3.66	5.33
Corona Extra	3.65	3.28	3.36	2.99	3.34	3.06
Corona Light	3.41	3.09	3.09	2.77	3.17	2.91
Heineken	3.53	3.28	3.28	3.00	3.34	3.26
Heineken Light	3.16	2.97	2.86	2.66	2.97	2.90
Michelob	5.27	5.94	5.19	5.84	4.39	5.70
Michelob Light	5.19	5.75	5.06	5.62	5.00	5.24
Miller Genuine Draft	4.42	5.61	4.24	5.47	4.13	5.36
Miller High Life	4.43	5.65	4.24	5.46	4.15	5.38
Miller Lite	4.39	5.58	4.25	5.48	4.13	5.35

Note: This table provides the average markups for each product (brand-size combination) in fiscal years 2007 and 2010.

markups on ABI 12 packs tend to be about \$0.55 higher in 2010 as compared to 2007, which reflects higher Bertrand prices and supermarkups. The markups on Miller 12 packs tend to be about \$1.20 higher and the markups on Coors products tend to be about \$1.67 higher, reflecting the combined impact of higher Bertrand prices, higher supermarkups, and lower marginal costs.

Before proceeding, it is worth discussing whether the timing parameter that emerges from our analysis, $\eta = 0.26$, comports with reasonable priors about oligopoly supergames. As a purely mathematical observation, it is not difficult to generate such a timing parameter if the duration of deviation exceeds that of punishment. For example, the combination $(\delta = 0.9, \phi = 0.9, \tau_1 = 2, \tau_2 = 1)$ yields $\eta = 0.27$, through an application of equation (7), and is consistent with a simple form of imperfect monitoring in which punishment ensues after a one-period lag. We are agnostic about whether deviation might indeed last longer than punishment, especially as firms seemingly should not punish longer than is necessary, and so view our result as plausible. Higher timing parameters obtain if punishment is relatively longer than in the above example. The combination $(\delta = 0.9, \phi = 0.9, \tau_1 = 1, \tau_2 = 1)$ yields $\eta = 0.45$ and the combination $(\delta = 0.9, \phi = 0.9, \tau_1 = 1, \tau_2 = 5)$ yields $\eta = 0.74$.

An alternative interpretation of our timing parameter is that it embeds the influence of some pricing friction that is not modeled explicitly. Different microfoundations are available—the friction could arise if the leader is not certain about the IC constraints of other firms and wishes to avoid an accidental punishment phase, or if price leadership creates some risk of antitrust penalties, for example. Let the augmented slack function be

$$(14) \quad g_i \equiv \frac{1}{1 - \eta^*} \pi_i^{PL} - \pi_i^{D,i} - \frac{\eta^*}{1 - \eta^*} \pi_i^B - \zeta,$$

margin estimates of 64.4 percent and 57.4 percent for manufacturers in the ready-to-eat cereals industry, based on the Annual Survey of Manufacturers and an industry study, respectively.

where ζ is the friction. There is a continuum of (η, ζ) combinations that would generate a binding IC constraint of $g_i = 0$. To illustrate, we express the timing parameter in terms of the friction:

$$(15) \quad \eta^* = \frac{\pi_i^{D,i} - \pi_i^{PL} + \zeta}{\pi_i^{D,i} - \pi_i^B + \zeta}.$$

Our baseline analysis implicitly imposes $\zeta = 0$ and obtains $\eta = 0.26$. However, these two objects are not separately identifiable, and the presence of a friction would lead us to underestimate the timing parameter. This would not affect our imputation results or the analysis of equilibrium because the binding IC constraint crosses zero at the same supermarkup—indeed that is what defines the continuum of jointly identified (η, ζ) combinations.⁴² It could generate different counterfactual inferences if the friction depends on the supermarkup.

B. Analysis of Equilibrium

We now examine the preferred model in greater detail. To start, we provide some intuition for how the supermarkups are selected and why they change with the Miller-Coors merger. We exploit that it is possible to compute profit under price leadership, deviation, and Bertrand (π^{PL}, π^D, π^B) for any counterfactual supermarkups, holding fixed marginal costs that are recovered with the preferred model. We consider counterfactual supermarkups $\tilde{\mathbf{m}}_t = \iota \mathbf{m}_t$, for $\iota = (0, 0.01, 0.02, \dots, 1.50)$. This scales all of the region-specific supermarkups by the same multiplicative factor; with $\iota = 0$ the outcomes are identical to Bertrand, and with $\iota = 1$ the supermarkups are those that arise in the PLE.

Figure 3 plots price leadership and deviation profit as indices relative to Bertrand profit. Results are provided for fiscal years 2007 and 2010, which immediately predate and postdate the Miller-Coors merger. The vertical blue line marks the PLE (equivalently, $\iota = 1.00$). The profit functions take a value of one at $\tilde{\mathbf{m}}_t = 0$ because outcomes are equivalent to Bertrand. From there, they increase in the supermarkup, with deviation profit increasing at a faster rate than price leadership profit. If we were to extend these graphs to large enough supermarkups, profit under price leadership would flatten and (eventually) start to fall, whereas the slope of deviation profit would converge to zero.

At the PLE in fiscal year 2007, price leadership increases profit above Bertrand by 15 percent for ABI, 20 percent for Miller, and 25 percent for Coors. Deviation increases static profit even more, and this is especially true for Coors, for which deviation profit exceeds Bertrand profit by 35 percent at the PLE. Thus, among the coalition firms, it appears that Coors has the greatest incentive to deviate, and we verify this momentarily. There is more symmetry in fiscal year 2010. Price leadership increases the profit of ABI and MillerCoors by 21 percent, relative to Bertrand, evaluated at the PLE, and deviation increases profit by 29 percent for both firms.

⁴²In online Appendix Figure G.4, we plot the (η, ζ) combinations under which the MillerCoors IC constraint binds in fiscal year 2010. With $\eta = 0.45$, the magnitude of the friction is equivalent to 8.25 percent of MillerCoors' profit in the PLE.

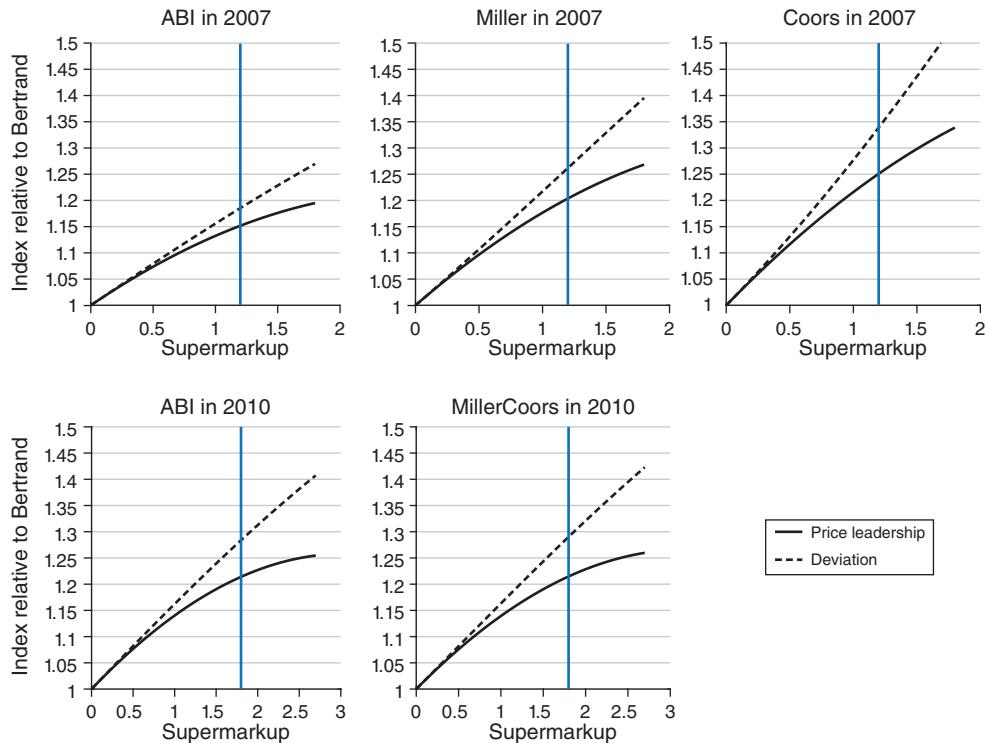


FIGURE 3. PROFIT FUNCTIONS OF COALITION FIRMS

Notes: The figure provides the profit under price leadership and deviation for coalition firms in fiscal years 2007 and 2010, reported as an index relative to Bertrand profit. Statistics are computed for a range of supermarkups. At the vertical blue line, the region-specific supermarkups are those recovered from the imputation procedure with a timing parameter of 0.26 ($\bar{m}_{2007} = 1.20$ and $\bar{m}_{2010} = 1.80$). These supermarkups are scaled by the multiplicative factors $\iota = (0.00, 0.01, \dots, 1.50)$ in the evaluation of the profit functions.

Figure 4 plots the equivalent slack functions for the coalition firms in fiscal years 2007 and 2010, which we obtain from equations (5) and (6) using the timing parameter of 0.26 and the profit functions just analyzed. The vertical blue line marks the PLE. The IC constraints in the leader's maximization problem are satisfied if all of the slack functions are weakly positive (i.e., at or above the horizontal blue line). The slack functions are positive for small supermarkups and negative for large supermarkups. In 2007, the slack function of Coors crosses zero at the PLE, even as the slack functions of Miller and ABI are positive. Thus, Coors provides the binding constraint. The MillerCoors slack function (in 2010) crosses zero with larger supermarkups than either the Miller or Coors slack functions (in 2007), which explains why average supermarkups are higher in 2010 than in 2007. This suggests that the Miller-Coors merger is the main driver of the higher supermarkups in the latter half of our sample, though the illustration is not definitive because demand and cost conditions also change. We use a counterfactual later in the article to isolate the effect of the merger.

We turn now to the dispersion of supermarkups across regions. Figure 5 shows that supermarkups tend to be higher in regions where ABI has a large market share and Coors (in 2007) or MillerCoors (in 2010) have a smaller market share. This

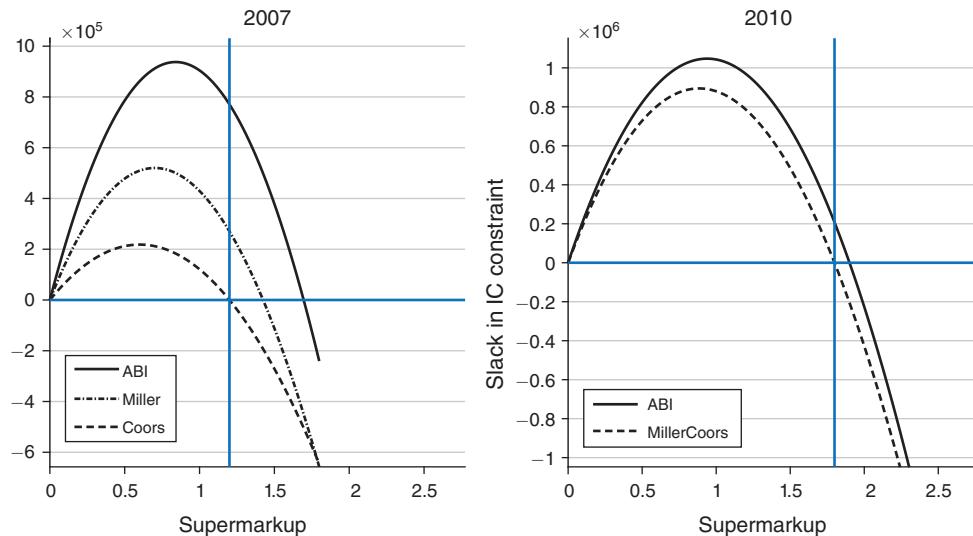


FIGURE 4. SLACK FUNCTIONS OF THE COALITION FIRMS

Notes: The figure provides the equivalent slack functions of the coalition firms in each fiscal year. At the vertical blue lines, the region-specific supermarkups are those recovered from the imputation procedure with a timing parameter of 0.26 ($\bar{m}_{2007} = 1.20$ and $\bar{m}_{2010} = 1.80$). These supermarkups are scaled by the multiplicative factors $\iota = (0.00, 0.01, \dots, 1.50)$ in the evaluation of the slack functions.

reflects the Kuhn-Tucker conditions that characterize the solution to ABI's constrained maximization problem: ABI benefits more from a higher supermarkup if it has a large market share, and the effect of a higher supermarkup on the binding IC constraint is greater if Coors and MillerCoors have large market shares. As the Herfindahl-Hirschman Index (HHI) tends to be high if ABI is large and other firms are small, the region-specific supermarkups are also correlated with the HHI. These are not structural relationships, but are informative nonetheless because market shares provide some composite measure of consumer willingness-to-pay and marginal cost.

To explore the role of multi-market contact in generating this variation across regions, we recompute equilibrium under the alternative assumption that the leader faces multiple, distinct, region-specific constrained maximization problems. This amounts to finding the supermarkup, m_r , for each region r , that satisfies $g_{ir}(m_r; \eta, \Psi_t) = 0$. We hold fixed the marginal costs and timing parameter that we recover from the baseline model. Figure 6 provides two scatter plots of the counterfactual supermarkups (vertical axis) and the baseline supermarkups (horizontal axis). There are noticeable differences in fiscal year 2007, but these mostly disappear in fiscal year 2010, reflecting the greater symmetry among the coalition firms after the Miller-Coors merger. Interestingly, multi-market contact does not affect the average supermarkup or the profitability of price leadership much.⁴³

⁴³ In fiscal year 2007, the profits of ABI, Miller, and Coors, are about 0.5 percent, 0.5 percent, and 1 percent higher, respectively, in the baseline model than in the counterfactual with independent regions. In fiscal year 2010, the profit of both ABI and MillerCoors is about 0.5 percent higher in the baseline model. In online Appendix D, we explore imputation under the alternative assumption that IC constraints are not pooled across regions.

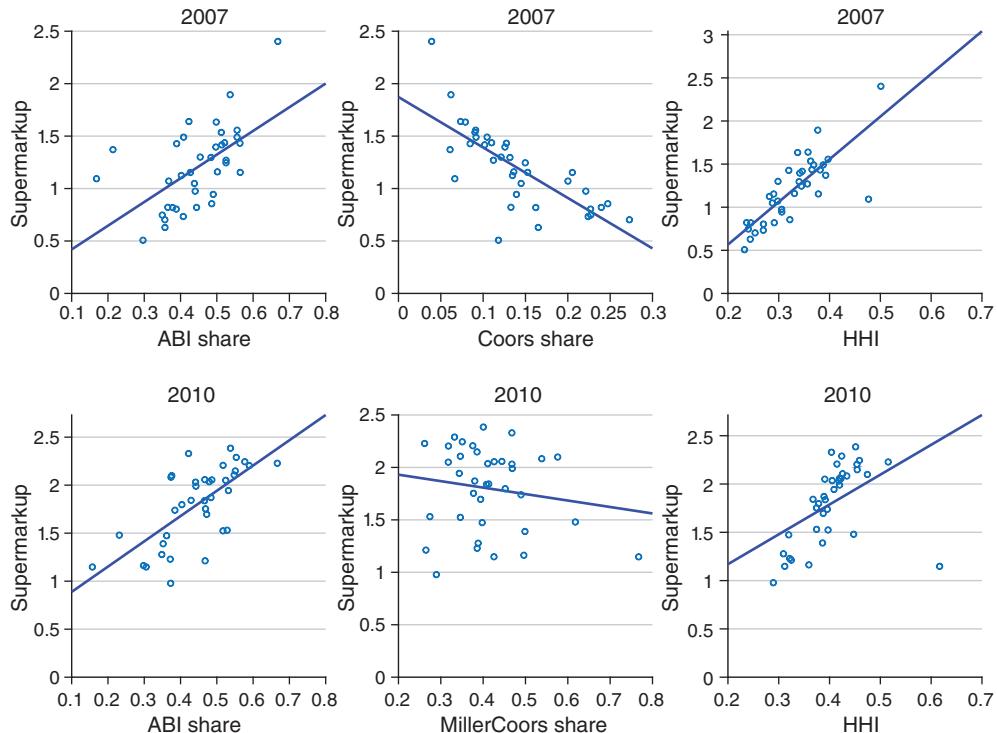


FIGURE 5. REGION-SPECIFIC SUPERMARKUPS IN 2007 AND 2010

Notes: The figure provides scatter plots of the region-specific supermarkups in fiscal years 2007 and 2010 against the ABI market share (in 2007 and 2010), the Coors market share (in 2007), the MillerCoors market share (in 2010), and the HHI (in 2007 and 2010). Lines of best fit are provided to assist with interpretation.

V. Application to Merger Review

A. Background

In this section, we analyze the Miller-Coors and ABI-Modelo mergers using the price leadership model. This covers two interesting merger scenarios. First, with Miller-Coors, the merger involves two coalition firms, including the binding firm. This combination creates slack in the IC constraints at premerger prices, and allows the leader to support higher supermarkups in equilibrium. Second, with ABI-Modelo, the merger involves the acquisition of an important fringe firm by the leader. By raising the prices of the acquired fringe firm, the leader can create slack in the binding IC constraint, and thereby support higher supermarkups in equilibrium.⁴⁴ We adopt standard antitrust parlance and refer to the change in the supermarkup as

⁴⁴This mechanism is consistent with the allegations of the DOJ regarding ABI-Modelo:

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 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.

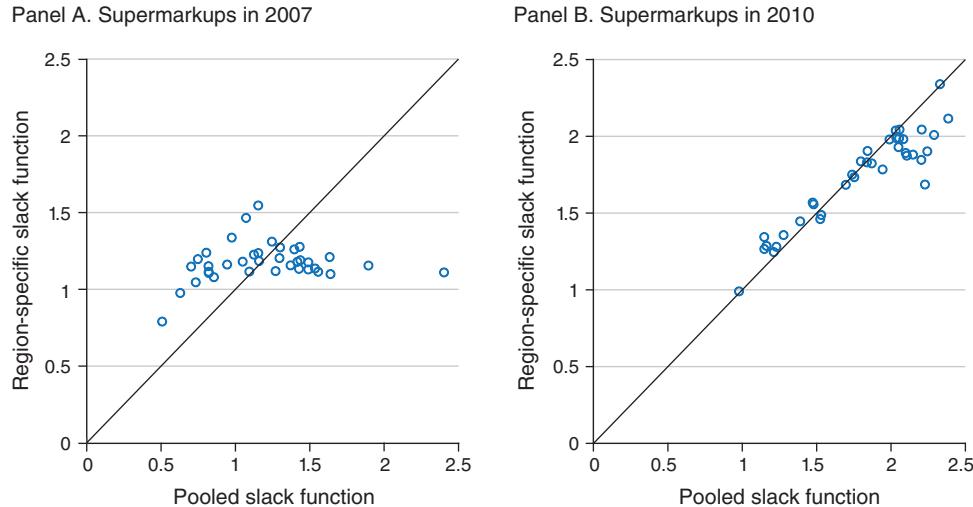


FIGURE 6. REGION-SPECIFIC SUPERMARKUPS AND MULTI-MARKET CONTACT

Notes. The figure provides scatter plots of the region-specific supermarkets in fiscal years 2007 and 2010 for the baseline model (horizontal axis) and a counterfactual simulation in which the leader solves region-specific constrained maximization problems. Each panel has a 45° line to assist with interpretation.

the coordinated effect of the merger. The overall price changes also reflect shifts in underlying Bertrand prices, which we refer to as the unilateral effects of the merger.

B. Implementation

We focus on fiscal year 2010 for Miller-Coors, which is the first complete fiscal year after the merger. The data contain PLE outcomes with merger efficiencies. We obtain the “no merger” scenario by simulating the PLE under the assumptions that (i) the Miller and Coors brands are owned by separate firms, and (ii) the estimated changes to the Miller and Coors marginal cost functions do not occur. The latter of these involves adding $\hat{\beta}_2$ and $\hat{\beta}_3$ to the marginal costs of Miller and Coors products (see Table 3), and accounting for the greater shipping distances that arise with separate ownership. We also obtain a “merger without efficiencies” scenario by simulating the PLE with joint pricing but no merger efficiencies. A comparison of these scenarios isolates the impact of the merger.

For ABI-Modelo, we focus on fiscal year 2011, which is the closest year in our sample to the acquisition date. The data contain the “no merger” scenario. We simulate PLE outcomes in which ABI and Modelo brands are owned by the same firm. We consider three specific cases: (i) no merger efficiencies, (ii) a “minor” efficiency comprising a \$0.50 reduction in Modelo costs, and (iii) a “major” efficiency comprising product-specific cost reductions which leave Bertrand prices exactly unchanged due to the merger.⁴⁵ A comparison of these scenarios against the baseline modeling results isolates the impact of the merger.

⁴⁵The major efficiency is a multiproduct version of the compensating marginal cost reductions derived in Werden (1996). On average, we reduce ABI marginal costs by \$0.29 and Modelo marginal costs by \$1.80.

There is an additional question about how to model the postmerger prices of Modelo. Because Modelo sells imported beers that are relatively more expensive than those of ABI and MillerCoors, successful postmerger coordination probably would not require that the same supermarkup apply to Modelo brands. Instead, we think it more likely that ABI would set Modelo prices to maximize the present value of its profit, accounting for impacts on IC constraints. As computing such an equilibrium would be computationally prohibitive, we work with an approximation where postmerger Modelo prices are equal to Bertrand prices plus some region-specific amount that applies to all Modelo products. This essentially adds a second set of supermarkups to the model.

Thus, in the postmerger equilibrium, there are $2R$ supermarkups that must be computed. For notational purposes, we denote the supermarkups that apply to ABI and MillerCoors brands as $\mathbf{m}^1 = (m_1^1, m_2^1, \dots, m_R^1)$ and the supermarkups that apply to Modelo products as $\mathbf{m}^2 = (m_1^2, m_2^2, \dots, m_R^2)$. Letting ABI be firm 1 and MillerCoors be firm k , postmerger equilibrium satisfies the following first order conditions:

$$(16) \quad g_k(\mathbf{m}^1, \mathbf{m}^2; \eta) \equiv \sum_{r=1}^R \left(\frac{1}{1-\eta} \pi_{kr}^{PL}(m_r^1, m_r^2) - \pi_{kr}^{D,2}(m_r^1, m_r^2) - \frac{\eta}{1-\eta} \pi_{kr}^B \right) = 0,$$

$$(17) \quad h_r^1(m_r^1, m_R^1, m_r^2, m_R^2) \equiv \frac{\frac{\partial \pi_{1r}^{PL}(m_r^1, m_r^2)}{\partial m_r^1}}{\frac{\partial \pi_{kr}^{D,k}(m_r^1, m_r^2)}{\partial m_r^1}} - \frac{\frac{\partial \pi_{1R}^{PL}(m_R^1, m_R^2)}{\partial m_R^1}}{\frac{\partial \pi_{kr}^{D,k}(m_R^1, m_R^2)}{\partial m_R^1}} - \frac{\frac{\partial \pi_{1R}^{PL}(m_R^1, m_R^2)}{\partial m_R^1}}{\frac{\partial \pi_{kr}^{PL}(m_R^1, m_R^2)}{\partial m_R^1}} = 0,$$

$$(18) \quad h_r^2(m_r^1, m_R^1, m_r^2, m_R^2) \equiv \frac{\frac{\partial \pi_{1r}^{PL}(m_r^1, m_r^2)}{\partial m_r^2}}{\frac{\partial \pi_{kr}^{D,k}(m_r^1, m_r^2)}{\partial m_r^2}} - \frac{\frac{\partial \pi_{1R}^{PL}(m_R^1, m_R^2)}{\partial m_R^2}}{\frac{\partial \pi_{kr}^{D,k}(m_R^1, m_R^2)}{\partial m_R^2}} - \frac{\frac{\partial \pi_{1R}^{PL}(m_R^1, m_R^2)}{\partial m_R^2}}{\frac{\partial \pi_{kr}^{PL}(m_R^1, m_R^2)}{\partial m_R^2}} = 0.$$

An important detail is that the second ratios in the expressions for h^1 and h^2 are identical, and involve differentiating with respect to m_R^1 . There are $R - 1$ identifying equations for h^1 as, by inspection, the equation is always satisfied for $r = R$. However, there are R equations for h^2 , so combining with the binding slack function (i.e., $g_k(\mathbf{m}^1, \mathbf{m}^2; \eta) = 0$) there are $2R$ equations in total, and the postmerger supermarkups are exactly identified.

C. Simulation Results

Table 6 summarizes the effects of the mergers on prices. The first two columns consider Miller-Coors with and without efficiencies. The final three columns consider ABI-Modelo under the three different efficiency scenarios. We report the average supermarkup change, the average Bertrand price change, and the total price change.⁴⁶

⁴⁶We calculate the average supermarkups as means across the 37 regions, for comparability with Table 3. However, the average price changes are means across the product-region-quarter observations, so the average total price change is not always equal to the sum of the average supermarkup and Bertrand price changes.

TABLE 6—MERGER PRICE EFFECTS

Merger:	Miller-Coors		ABI-Modelo		
	Yes (1)	No (2)	None (3)	Minor (4)	Major (5)
Δ Supermarkup					
Domestic brewers	0.51	0.39	0.40	0.37	0.32
Modelo	—	—	1.46	1.45	1.44
Δ Bertrand price					
ABI	-0.01	0.16	0.20	0.21	0.00
MillerCoors	—	—	0.08	0.07	0.00
Miller	0.16	0.50	—	—	—
Coors	0.01	0.89	—	—	—
Modelo	-0.02	-0.04	1.84	1.31	0.00
Δ Total price					
ABI	0.50	0.55	0.59	0.58	0.32
MillerCoors	—	—	0.48	0.44	0.32
Miller	0.67	0.90	—	—	—
Coors	0.52	1.28	—	—	—
Modelo	-0.03	-0.04	3.39	2.85	1.53

Notes: The table summarizes the price effects of mergers. Column 1 is a comparison of the observed equilibrium to a counterfactual without the Miller-Coors merger. Column 2 is a comparison of a counterfactual in which the Miller-Coors merger occurs without efficiencies to a counterfactual in which the Miller-Coors merger does not occur. Columns 3–5 are comparisons of counterfactuals in which an ABI-Modelo merger occurs, with varying levels of efficiencies, to the observed equilibrium. The Miller-Coors merger is evaluated in fiscal year 2010, whereas the ABI-Modelo merger is evaluated in fiscal year 2011. Statistics for the supermarket change are averages across the 37 regions. Statistics for the changes in Bertrand prices, total prices, and market shares are averages over observations at the product-region-quarter level.

Starting with Miller-Coors, we find that the merger increases the supermarket by 0.51 in fiscal year 2010 (column 1). Our imputation results indicate that the average supermarket increases from 1.20 in fiscal year 2007 to 1.80 in fiscal year 2010 (Table 3). Thus, the Miller-Coors merger appears to account for 85 percent ($0.51/0.60 = 0.85$) of the change. We attribute the remainder to changing cost and demand conditions. Due to the merger efficiencies, Bertrand prices are largely unaffected, so the changes in total price primarily reflect coordinated effects. Analyzing the merger without efficiencies, we find that the increase in the supermarket decreases slightly, but the Bertrand price increases are higher, especially for Miller and Coors brands (column 2). Total prices are higher on average.

Turning to the ABI-Modelo merger, column 3 of Table 6 shows that Bertrand prices of ABI and Modelo products increase by \$0.20 and \$1.84, respectively, with the magnitude of the latter reflecting an incentive to steer customers toward higher-markup ABI brands. Prices also increase due to a higher supermarket, which rises by \$0.40 for ABI and MillerCoors, and \$1.46 for Modelo. Efficiencies offset the unilateral incentive to raise prices, as in standard merger analysis, but in this case do little to reduce the supermarket. This occurs because the impact of the marginal costs of ABI and Modelo on the supermarket is indirect, coming through the (binding) MillerCoors slack function. The changes in total price reflect both the change in Bertrand price and the change in the supermarket.

Table 7 decomposes the binding slack function across various scenarios to explore in greater detail how the mergers impact coordination incentives. Columns 1 and 5

TABLE 7—DECOMPOSITION OF BINDING SLACK FUNCTIONS

Ownership: Efficiencies: Supermarkup:	Miller-Coors merger				ABI-Modelo merger			
	Pre	Post	Post	Post	Pre	Post	Post	Post
	No	No	Yes	Yes	No	No	Yes	Yes
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	
(1) π^{PL}	11.85	35.08	41.12	42.78	39.70	42.85	42.67	43.73
(2) π^D	12.77	36.45	42.53	45.44	42.15	45.06	44.88	47.06
(3) π^B	9.22	29.84	35.22	35.22	32.72	34.41	34.24	34.24
(4) $\frac{1}{1-\eta} \pi^{PL}$	16.01	47.41	55.57	57.82	53.65	57.90	57.66	59.09
(5) $\pi^D + \frac{\eta}{1-\eta} \pi^B$	16.01	46.93	54.90	57.82	53.65	57.15	56.91	59.09
(6) $g(m)$	0	0.474	0.668	0	0	0.749	0.747	0

Notes: The table shows the components of the binding slack function for Miller-Coors and ABI-Modelo. Units are millions of dollars. Slack functions are calculated using the timing parameter $\eta = 0.26$. The Miller-Coors merger is evaluated in fiscal year 2010, and the ABI-Modelo merger is evaluated in fiscal year 2011. Columns 3 and 4 apply the estimated efficiencies of the Miller-Coors merger while columns 7 and 8 apply a reduction in Modelo costs of \$0.50 (“minor efficiencies”).

contain premerger values for the Miller-Coors and ABI-Modelo mergers, respectively. Coors is the binding firm prior to the Miller-Coors merger, while MillerCoors is the binding firm prior to the ABI-Modelo merger.

Columns 2 and 6 present the slack function and its components after each merger but at premerger supermarkups and costs. With Miller-Coors, the slack function changes due to the unilateral effect on Bertrand prices and because the binding postmerger slack function incorporates Miller products. With ABI-Modelo, the slack function changes only due to the effect on Bertrand prices. In both mergers, row 5 shows that the present value of deviating increases as both punishment phase competition is softened (row 3) and profits from deviating increase (row 2). This mechanism has been highlighted in several empirical and theoretical articles that show mergers can make tacit collusion more difficult when collusion is at joint monopoly prices (e.g., Davidson and Deneckere 1984, Werden and Baumann 1986, Davis and Huse 2010). However, in our model collusion is not at the joint monopoly level, so even a merger between tacitly colluding firms can make collusion easier. Indeed, each merger increases price leadership profits by more than deviation profits, allowing for higher supermarkups.

Columns 3 and 7 illustrate how efficiencies can change incentives to collude. In the case of Miller-Coors, efficiencies increase supermarkups. Comparing column 3 to column 2 shows that while efficiencies increase each component of the slack function, the value of price leadership increases by more than the total deviation value. This allows for greater supermarkups. In contrast, comparing columns 6 and 7 demonstrates that efficiencies very slightly reduce the slack function and as a result supermarkups fall in the case of ABI-Modelo. (Here, we focus on the minor efficiencies scenario.) Together, these simulation results indicate that efficiencies have ambiguous implications for the magnitude of coordinated effects.

Table 8 summarizes the welfare effects of the two mergers. We consider a range of assumptions about the presence and magnitude of efficiencies, and whether supermarkups are allowed to adjust, as they would in equilibrium of the price leadership

TABLE 8—WELFARE EFFECTS ACROSS MERGER SCENARIOS

	Miller-Coors merger				ABI-Modelo merger			
	No Fix	Yes Fix	No Adjust	Yes Adjust	None Fix	None Adjust	Minor Adjust	Major Adjust
Efficiencies:								
Supermarkup:								
$\Delta\bar{m}$	0.00 (1)	0.00 (2)	0.39 (3)	0.51 (4)	0.00 (5)	0.40 (6)	0.37 (7)	0.32 (8)
% Δ Profit								
All firms	4.86	7.36	8.38	12.21	5.34	9.12	9.12	10.88
ABI	7.06	0.22	9.72	3.95	—	—	—	—
ABI + Modelo	—	—	—	—	2.65	6.37	7.11	14.59
Miller + Coors	1.08	18.48	4.48	23.27	7.93	10.89	10.15	5.73
Modelo	8.55	0.01	16.71	10.58	—	—	—	—
Δ CS/ Δ Profit	-1.74	-0.16	-1.95	-0.99	-3.03	-2.53	-2.46	-1.73

Notes: The table shows the percentage change in profit and the ratio of consumer surplus (CS) change to profit change due to the Miller-Coors and the ABI-Modelo mergers, under different assumptions about efficiencies and whether the supermarkup adjusts. The comparisons for Miller-Coors are for fiscal year 2010. Column 4 features prices observed in the data, whereas columns 1–3 and the baseline “no merger” scenario feature simulated prices. The comparisons for ABI-Modelo are for fiscal year 2011. The “no merger” scenario features prices observed in the data, whereas columns 5–8 feature simulated prices. The premerger supermarkups used in columns 1, 2, and 5 solve the leader’s constrained maximization problem without the merger, and are applied for illustrative purposes. The postmerger supermarkups used in columns 3, 4, 6, 7, and 8 solve the leader’s constrained maximization problem, taking into account efficiencies, and so differ across columns even for the same merger.

model. Column 4 shows that the Miller-Coors merger with the estimated efficiencies was roughly welfare neutral, as industry profit increased by 12.21 percent and consumer surplus fell by 99 percent of profit increase. Comparisons to columns 1–3 establish that consumer surplus loss increases relative to profit gains in the absence of the efficiencies, and that allowing the supermarkup to adjust after the merger contributes to profit gains and consumer surplus losses. Columns 6–8 examine the ABI-Modelo merger under no efficiencies, minor efficiencies, and major efficiencies. The merged ABI-Modelo firm’s profit gains increase in the size of the efficiency, but MillerCoors’ profit gains decrease. Total industry profit increases by 9.12 percent, 9.12 percent, and 10.88 percent, respectively, and consumer surplus loss far exceeds the profit gain in each scenario. A comparison of columns 5 and 6 establishes that allowing the supermarkup to adjust after the merger contributes to profit gains and consumer surplus losses.

VI. Conclusion

This study is an attempt to apply methodologies that have become standard in industrial organization over the previous two decades to a repeated pricing game of perfect information. The particular setting—price leadership in the beer industry—is advantageous in part because documentary evidence in the public record helps inform the timing of actions and the strategies that firms play along the equilibrium path. Additional assumptions, more difficult to verify, are required in fully specifying strategies. In our setting they include that (i) ABI, the leader, picks supermarkups subject to fringe firm prices and the constraint that the followers, Miller and Coors, would rather adopt the supermarkups than deviate; (ii) fringe firms maximize their static profit functions; and (iii) Miller and Coors incentives to match are based

on trigger strategies with reversion to static Bertrand prices. Different assumptions could be adapted to other repeated pricing games of perfect information, including but not limited to other markets which exhibit price leadership.

A practical benefit of our modeling approach is that it allows for the prospective evaluation of mergers. This is especially important in US brewing, where there have been many recent mergers and pricing behavior at odds with the standard framework of static price competition (Miller and Weinberg 2017). However, a few limitations of our counterfactuals should be noted. First, they require coordination to be present premerger so IC constraints can be evaluated, and then adjusted to postmerger conditions. The model does not predict when a coordinated equilibrium would emerge due to a merger. Second, our simulations hold the timing parameter—and thus the duration of deviation and punishment—fixed. If timing considerations are endogenously determined by market structure, a version of the Lucas (1976) critique applies. Additionally, our counterfactuals hold fixed the identities of the coalition members. A way to relax this might be to simulate counterfactual profits for each potential coalition. Then a rule for cartel formation, for example that in Bos and Harrington (2010), could determine which are stable. Finally, the set of fringe firms is held fixed, even though a number of articles address the possibility that mergers could induce entry both theoretically (e.g., Spector 2003; Caradonna, Miller, and Sheu 2021) and empirically (e.g., Collard-Wexler 2014, Fan and Yang 2020). The connection between coordination and postmerger entry is an interesting avenue for future research.

REFERENCES

Abreu, Dilip, David Pearce, and Ennio Stacchetti. 1986. "Optimal Cartel Equilibrium with Imperfect Monitoring." *Journal of Economic Theory* 39 (1): 251–69.

Ascher, Bernard. 2012. *Global Beer: The Road to Monopoly*. Washington, DC: American Antitrust Institute.

Ashenfelter, Orley C., Daniel S. Hosken, and Matthew C. Weinberg. 2015. "Efficiencies Brewed: Pricing and Consolidation in US Brewing." *RAND Journal of Economics* 46 (2): 328–61.

Asker, John. 2010. "A Study of the Internal Organization of a Bidding Cartel." *American Economic Review* 100 (3): 724–62.

Asker, John. 2016. "Diagnosing Foreclosure Due to Exclusive Dealing." *Journal of Industrial Economics* 64 (3): 375–410.

Asker, John, Allan Collard-Wexler, and Jan De Loecker. 2019. "(Mis)Allocation, Market Power, and Global Oil Extraction." *American Economic Review* 109 (4): 1568–1615.

Aumann, Robert J. 1990. "Nash Equilibrium Are Not Self-Enforcing." In *Economic Decision-Making: Games, Econometrics and Optimisation*, edited by Jean Jaskold Gabszewicz, Jean François Richard, and Laurence A. Wolsey, 201–06. Amsterdam: North Holland.

Bain, Joe S. 1960. "Price Leaders, Barometers, and Kinks." *Journal of Business* 33 (3): 193–203.

Baker, Jonathan B. 2002. "Mavericks, Mergers, and Exclusion: Proving Coordinated Competitive Effects under the Antitrust Laws." *New York University Law Review* 77 (1): 135–203.

Baker, Jonathan B. 2010. "Market Concentration in the Antitrust Analysis of Horizontal Mergers." *Beer Institute. 2012. 2011–2012 Beer Institute Annual Report*. Washington, DC: Beer Institute.

Bernheim, B. Douglas, and Michael D. Whinston. 1990. "Multimarket Contact and Collusive Behavior." *RAND Journal of Economics* 21 (1): 1–26.

Bos, Iwan, and Joseph E. Harrington Jr. 2010. "Endogenous Cartel Formation with Heterogeneous Firms." *RAND Journal of Economics* 41 (1): 92–117.

Bresnahan, Timothy F. 1987. "Competition and Collusion in the American Automobile Oligopoly: The 1955 Price War." *Journal of Industrial Economics* 35 (4): 457–82.

Bronnenberg, Bart J., Michael W. Kruger, and Carl F. Mela. 2008. "Database Paper: The IRI Marketing Data Set." *Marketing Science* 27 (4): 745–48.

Butters, R. Andrew, Daniel W. Sacks, and Boyoung Seo. 2020. "How Do National Firms Respond to Local Shocks? Evidence from Excise Taxes." Unpublished.

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

Caradonna, Peter, Nathan H. Miller, and Gloria Sheu. 2021. "Mergers, Entry, and Consumer Welfare." Available at SSRN: <https://ssrn.com/abstract=3537135>.

Chilet, Jorge Alé. 2018. "Gradually Rebuilding a Relationship: The Emergence of Collusion in Retail Pharmacies in Chile." Unpublished.

Ciliberto, Federico, and Jonathan 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, Robert, and Jean-François Houde. 2013. "Collusion with Asymmetric Retailers: Evidence from a Gasoline Price-Fixing Case." *American Economic Journal: Microeconomics* 5 (3): 97–123.

Collard-Wexler, Allan. 2014. "Mergers and Sunk Costs: An Application to the Ready-Mix Concrete Industry." *American Economic Journal: Microeconomics* 6 (4): 407–47.

Compte, Olivier, Frédéric Jenny, and Patrick Rey. 2002. "Capacity Constraints, Mergers and Collusion." *European Economic Review* 46 (1): 1–29.

d'Aspremont, Claude, Alexis Jacquemin, Jean Jaskold Babszewicz, and John A. Weymark. 1983. "On the Stability of Collusive Price Leadership." *Canadian Journal of Economics* 16 (1): 17–25.

Davidson, Carl, and Raymond Deneckere. 1984. "Horizontal Mergers and Collusive Behavior." *International Journal of Industrial Organization* 2 (2): 117–32.

Davis, Peter J., and Cristian Huse. 2010. "Estimating the 'Coordinated Effects' of Mergers." Unpublished.

DellaVigna, Stefano, and Matthew Gentzkow. 2019. "Uniform Pricing in U.S. Retail Chains?" *Quarterly Journal of Economics* 134 (4): 2011–84.

Donsimoni, Marie-Paule, Nicholas S. Economides, and Herakles M. Polemarchakis. 1986. "Stable Cartels." *International Economic Review* 27 (2): 317–27.

Drennan, Ronald, Matthew Magura, and Aviv Nevo. 2013. "The Year in Review: Economics at the Antitrust Division 2012–2013." *Review of Industrial Organization* 43 (4): 291–302.

Eizenberg, Alon, and Dalia Shilian. 2019. "Structure, Conduct, and Contact: Competition in Closely-Related Markets." Unpublished.

Eizenberg, Alon, Dalia Shilian, and Daniel D. Blanga. 2020. "Estimating the Potential Impact of Multi-market Contact on the Intensity of Competition." Unpublished.

Evans, William N., and Ioannis N. Kessides. 1994. "Living by the 'Golden Rule': Multimarket Contact in the U.S. Airline Industry." *Journal of Quarterly Economics* 109 (2): 341–66.

Fan, Ying, and Chenyu Yang. 2020. "Merger, Product Variety and Firm Entry: The Retail Craft Beer Market in California." Unpublished.

Farrell, Joseph, and Matthew Rabin. 1996. "Cheap Talk." *Journal of Economic Perspectives* 10 (3): 103–18.

Gasmi, Farid, Jean-Jacques Laffont, and Quang Vuong. 1992. "Econometric Analysis of Collusive Behavior in a Soft-Drink Market." *Journal of Economics & Management Strategy* 1 (2): 277–311.

Greer, Douglas F. 1998. "Beer: Causes of Structural Change." In *Industry Studies*. 2nd ed., edited by Larry L. Duitsch, 28–64. Armonk, NY and London: M. E. Sharpe.

Harrington, Joseph E., Jr. 2013. "Evaluating Mergers for Coordinated Effects and the Role of 'Parallel Accommodating Conduct'." *Antitrust Law Journal* 78 (3): 651–68.

Harrington, Joseph E. 2018. "Developing Competition Law for Collusion by Autonomous Artificial Agents." *Journal of Competition Law and Economics* 14 (3): 331–63.

Harrington, Joseph E., and Andrzej Skrzypacz. 2011. "Private Monitoring and Communication in Cartels: Explaining Recent Collusive Practices." *American Economic Review* 101 (6): 2425–49.

Harrington, Joseph E., Jr., Roberto Hernan Gonzalez, and Praveen Kujal. 2016. "The Relative Efficiency of Price Announcements and Express Communication for Collusion: Experimental Findings." *Journal of Economic Behavior & Organization* 128: 251–64.

Igami, Mitsuru. 2015. "Market Power in International Commodity Trade: The Case of Coffee." *Journal of Industrial Economics* 63 (2): 225–48.

Igami, Mitsuru, and Takuo Sugaya. Forthcoming. "Measuring the Incentive to Collude: The Vitamin Cartels, 1990–1999." *Review of Economic Studies*.

Ishibashi, Ikuo. 2008. "Collusive Price Leadership with Capacity Constraints." *International Journal of Industrial Organization* 26 (3): 704–15.

Ivaldi, Marc, Bruno Jullien, Patrick Rey, Paul Seabright, and Jean Tirole. 2007. "The Economics of Tacit Collusion: Implications for Merger Control." In *The Political Economy of Antitrust*, Vol. 282

of *Contributions to Economic Analysis*, edited by Vivek Ghosal, and Johan Stennek, 217–40. Bingley, UK: Emerald Group Publishing.

Khwaja, Ahmed, and Beomjoon Shim. 2017. “The Collusive Effect of Multimarket Contact on Prices: Evidence from Retail Lumber Markets.” Paper presented at the Society for Economic Dynamics Annual Meeting, Edinburgh, Scotland, June 24.

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

Lemus, Jorge, and Fernando Luco. 2021. “Price Leadership and Uncertainty about Future Costs.” *Journal of Industrial Economics* 69 (2): 305–37.

Lucas, Robert E., Jr. 1976. “Econometric Policy Evaluation: A Critique.” *Carnegie-Rochester Conference Series on Public Policy* 1: 19–46.

Loertscher, Simon, and Leslie M. Marx. 2019. “Merger Review for Markets with Buyer Power.” *Journal of Political Economy* 127 (6): 2967–3017.

Markham, Jesse W. 1951. “The Nature and Significance of Price Leadership.” *American Economic Review* 41 (5): 891–905.

Michel, Christian, and Stefan Weiergraeber. 2018. “Estimating Industry Conduct in Differentiated Products Markets: The Evolution of Pricing Behavior in the RTE Cereal Industry.” Unpublished.

Miller, Nathan H., Gloria Sheu, and Matthew C. Weinberg. 2021. “Replication Data for: Oligopolistic Price Leadership and Mergers: The United States Beer Industry.” American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. <https://doi.org/10.3886/E140341V1>.

Miller, Nathan H., and Matthew C. Weinberg. 2017. “Understanding the Price Effects of the MillerCoors Joint Venture.” *Econometrica* 85 (6): 1763–91.

Miller, Nathan H., and Matthew C. Weinberg 2021. “Data on Driving Distances Between Breweries and IRI Media Markets.” Unpublished Data, (accessed August, 2021).

Mouraviev, Igor, and Patrick Rey. 2011. “Collusion and Leadership.” *International Journal of Industrial Organization* 29 (6): 705–17.

Nevo, Aviv. 2001. “Measuring Market Power in the Ready-to-Eat Cereal Industry.” *Econometrica* 69 (2): 307–42.

Nocke, Volker, and Nicolas Schutz. 2018. “Multiproduct-Firm Oligopoly: An Aggregative Games Approach.” *Econometrica* 86 (2): 523–57.

Oxenfeldt, Alfred R. 1952. “Professor Markham on Price Leadership: Some Unanswered Questions.” *American Economic Review* 42 (3): 380–84.

Pastine, Ivan, and Tuvana Pastine. 2004. “Cost of Delay and Endogenous Price Leadership.” *International Journal of Industrial Organization* 22 (1): 135–45.

Porter, Robert H. 1983. “A Study of Cartel Stability: The Joint Executive Committee, 1880–1886.” *Bell Journal of Economics* 14 (2): 301–14.

Rojas, Christian. 2008. “Price Competition in U.S. Brewing.” *Journal of Industrial Economics* 56 (1): 1–31.

Rosse, James N. 1970. “Estimating Cost Function Parameters without Using Cost Data: Illustrated Methodology.” *Econometrica* 38 (2): 256–75.

Rotemberg, Julio J., and Garth Saloner. 1986. “A Supergame-Theoretic Model of Price Wars during Booms.” *American Economic Review* 76 (3): 390–407.

Rotemberg, Julio J., and Garth Saloner. 1990. “Collusive Price Leadership.” *Journal of Industrial Economics* 39 (1): 93–111.

Scherer, Frederic. 1980. *Industrial Market Structure and Economic Performance*. Boston: Houghton Mifflin.

Seaton, Jonathan S., and Michael Waterson. 2013. “Identifying and Characterising Price Leadership in British Supermarkets.” *International Journal of Industrial Organization* 31 (5): 392–403.

Slade, Margaret E. 2004. “Market Power and Joint Dominance in U.K. Brewing.” *Journal of Industrial Economics* 52 (1): 133–63.

Spector, David. 2003. “Horizontal Mergers, Entry, and Efficiency Defences.” *International Journal of Industrial Organization* 21 (10): 1591–1600.

Stigler, George J. 1947. “The Kinky Demand Curve and Rigid Prices.” *Journal of Political Economy* 55 (5): 432–49.

Tremblay, Victor J., and Carol Horton Tremblay. 2005. *The U.S. Brewing Industry*. Cambridge, MA: MIT Press.

US Bureau of Labor Statistics. “Consumer Price Index for All Urban Consumers: All Items in U.S. City Average [CPIAUCNS].” FRED, Federal Reserve Bank of St. Louis. <https://fred.stlouisfed.org/series/CPIAUCNS>, (accessed August 9, 2021).

US Census Bureau. 2010. "American Community Survey (ACS): Public Use Microdata Sample (PUMS), 2005-2010." United States Department of Commerce. Inter-university Consortium for Political and Social Research.

US Energy Information Administration. "U.S. No 2 Diesel Retail Prices." retrieved from U.S. EIA; https://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=EMD_EP2D_PTE_NUS_DPG&f=M, (accessed August 9, 2021).

Vasconcelso, Helder. 2005. "Tacit Collusion, Cost Asymmetries, and Mergers." *RAND Journal of Economics* 36 (1): 39–62.

Werden, Gregory J. 1996. "A Robust Test for Consumer Welfare Enhancing Mergers among Sellers of Differentiated Products." *Journal of Industrial Economics* 44 (4): 409–13.

Werden, Gregory J., and Michael G. Baumann. 1986. "A Simple Model of Imperfect Competition in which Four Are Few but Three Are Not." *Journal of Industrial Economics* 34 (3): 331–35.