Oligopolistic Price Leadership and Mergers: An Empirical Model of the Beer Industry*

Nathan H. Miller†
Georgetown University

Gloria Sheu‡
U.S. Department of Justice

Matthew C. Weinberg§
The Ohio State University

March 25, 2019

Abstract

We study an infinitely-repeated game of oligopolistic price leadership in which one firm, the leader, proposes a supermarkup over Nash-Bertrand prices to a coalition of rivals. The supermarkup is chosen to maximize the leader's profit subject to incentive compatibility (IC) constraints and in anticipation of fringe firms' responses. We provide conditions under which the supermarkup can be recovered from aggregate scanner data. We apply the model to the U.S. beer industry over 2005-2011 and estimate that ABI and MillerCoors implemented supermarkups of \$0.60 in the wake of the Miller/Coors merger. Counterfactuals demonstrate that IC binds, as profit is greater with even higher supermarkups. The implied equality constraint jointly identifies a discount factor and an antitrust risk coefficient, the remaining structural parameters. We explore the coordinated effects of ABI/Modelo merger. Without divestitures, the merger would have relaxed IC and resulted in substantially higher prices. Finally, we return to the Miller/Coors merger. For many parameterizations, no supermarkup satisfies IC without the merger. Thus, it may be pivotal in generating price leadership.

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

*We thank Karl Schurter for helpful comments, as well as seminar participants at Federal Trade Commission, Harvard Business School, New York University, Pennsylvania State University, Princeton University, Texas A&M. All estimates and analyses in this paper based on IRI data are by the authors and not by IRI. The views expressed herein are entirely those of the authors and should not be purported to reflect those of the U.S. Department of Justice.

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

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

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

1 Introduction

Firms in concentrated industries sometimes repeatedly change their prices together in quick succession and by similar magnitudes, with these changes initiated by a single firm. In some cases, these changes are incommensurate with plausible changes in costs or demands, suggesting they are driven by changes in competition. Recent studies utilizing extremely detailed data document this phenomenon in retail industries ranging from supermarkets, pharmacies, and gasoline (Seaton and Waterson (2013), Chilet (2018), Lemus and Luco (2018), Byrne and de Roos (2019)). Anecdotal examples are discussed in Scherer (1980) and an older economics literature (e.g., Stigler (1947), Markham (1951), Oxenfeldt (1952)).[1] We follow Bain (1960) in referring to the pricing practice as *oligopolistic price leadership*.

One way to understand price leadership is in the context of a repeated pricing game. Economic theory indicates that repeated interaction within oligopolies can support collusive equilibria if firms are sufficiently patient (e.g., Friedman (1971), Abreu (1988)). Often in these models, however, many equilibria are available, raising the question of how firms go about selecting one. We argue that price leadership provides a natural resolution to the coordination problem, as the leader provides a focal point for other firms. Further, information can be disseminated through normal market interactions, allowing firms to avoid the explicit agreements frequently targeted by antitrust authorities.[2]

Proceeding along these lines, this paper presents an empirical model of oligopolistic price leadership. We take the model to data on U.S. brewing markets, for which there is documentary evidence of price leadership behavior (Miller and Weinberg (2017)). Once the model is estimated, we use simulations to explore the economic consequences of mergers. We focus especially on *coordinated effects*, which we conceptualize as a shift from a one-shot Nash equilibrium to a supracompetitive equilibrium, or from one supracompetitive equilibrium to another. We analyze both the ABI acquisition of Grupo Modelo in 2013 and the merger of Miller and Coors in 2008 (forming the MillerCoors joint venture). Our research is among the first to formally model coordinated effects in real-world markets.[3]

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

[2]Modern interpretations of antitrust law typically classify price leadership as tacit collusion, though the distinction between (legal) tacit collusion and (illegal) express collusion may be unclear (Kaplow (2013)).

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

We first provide a description of U.S. brewing markets (Section 2). In scanner data spanning 2001-2011, a handful of brewers account for about 80% of retail revenue. We cite to legal documents filed by the Department of Justice (DOJ) alleging that Anheuser Busch (ABI) pre-announces its annual list price changes as a signal to competitors, and that MillerCoors tends to follow. We show that the data support an abrupt increase in the prices of ABI and MillerCoors shortly after the consummation of the Miller/Coors merger, both in absolute terms and relative to the prices of Grupo Modelo and Heineken, the other large brewers. Finally, we describe the differentiated-products model of consumer demand estimated in Miller and Weinberg (2017), which we take as given in this paper.

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

We discuss identification and estimation in Section 4. Our main identification result is that the marginal costs that rationalize prices can be recovered for any candidate supermarkup. The connection flows through the Nash-Bertrand first order conditions (e.g., Rosse (1970)), though multiple numerical steps are required in implementation as Nash-Bertrand prices are unobserved. With this result in hand, a structural error term in the marginal cost function can be isolated, allowing for estimation with GMM. Valid instruments shift marginal revenue in the Nash-Bertrand equilibrium and are orthogonal to residual costs. A final complication is that the objects of interest in estimation (the supermarkups) are a choice variable rather than a structural parameter. Thus, a fully unrestricted model is underidentified as theory indicates supermarkups adjust with variation in valid instruments. In our application, we assume Nash-Bertrand competition prior to the Miller/Coors merger, which is sufficient for exact identification of the post-merger supermarkup.

We estimate post-merger supermarkups ranging from \$0.60 to \$0.74 (Section 5), depending on the specific demand specification employed. For context, \$0.60 is about six percent of the average price of a 12 pack. In counterfactual simulations, we find that higher supermarkups would increase ABI's per-period profit. Thus, to rationalize post-merger pricing within the context of the model, an IC constraint must bind, such that the present values of coordination and deviation are equal at the estimated supermarkup. To make progress, we obtain the profit under deviation and punishment using counterfactual simulations. We then recover combinations of the discount factor and an antitrust risk coefficient (which captures some disutility of coordination) such that one IC constraint binds. The results indicate that the IC of MillerCoors is the constraint on post-merger prices.

In Section 6, we use the model to examine the coordinated effects of mergers. We start with ABI's acquisition of Grupo Modelo, approved in 2013 by the DOJ only after the Modelo brands were divested to a third party. An initial DOJ Complaint characterizes Modelo as a maverick, defined in the 2010 Horizontal Merger Guidelines as "a firm that has often resisted otherwise prevailing industry norms to cooperate on price setting or other terms of competition." Mavericks are naturally modeled as fringe firms in our framework. Our simulation results indicate that incorporating Modelo into the coalition (as part of ABI) loosens IC constraints and results in substantially higher supermarkups.

We then turn to the 2008 Miller/Coors merger, and address the outstanding question of whether it was pivotal in generating a price leadership equilibrium. To implement, we unwind the merger in the post-merger periods and evaluate incentive compatibility across a wide range of supermarkups. Our results depend on the particular combinations of (the jointly identified) discount factor and antitrust risk used in the analysis. For most parameterizations, we find that Coors prefers deviation to coordination for any supermarkup. Further, Miller prefers deviation with an ABI-Miller coalition. Thus price leadership unravels without the MillerCoors joint venture. By contrast, the coalitions are stable if antitrust risk is negligible; unwinding the joint venture merely tightens IC constraints and reduces the supermarkup. The total price changes are not commensurate with what is observed in the raw data, however, which supports parameterizations with greater antitrust risk.

Finally, in Section 7, we conclude with a short summary and a discussion of the maintained assumptions about strategies off the equilibrium path, with an eye toward informing future research efforts.

1.1 Literature Review

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

A number of theoretical articles develop results on the organization of coalitions. Ishibashi (2008) and Mouraviev and Rey (2011) analyze repeated games in which (each period) the leader sets price in an initial stage and other firms set price in a subsequent stage; cartel profits are maximized by having the firm with the greatest incentive to deviate serve as the leader. Pastine and Pastine (2004) analyze a similar game in which a war of attrition determines the leader. Our model differs in that each period features an announcement followed by simultaneous pricing, rather than sequential pricing.[6] Under the timing and informational assumptions we maintain, any coalition firm could serve as the leader, and thus we assume the leader is exogenously determined. In allowing for partial coalitions, we build on a literature that considers homogeneous-product quantity games (e.g., d'Aspremont et al. (1983), Donsimoni et al. (1986), and Bos and Harrington (2010)).

With respect to the empirical literature, our research is methodologically most similar to Igami and Sugaya (2018) on the vitamin C cartel of the 1990s. The main result is that unexpected shocks to demand and fringe supply undermined incentive compatibility and led to the collapse of the cartel. As in our research, Igami and Sugaya estimate the structural parameters of a supergame in which trigger strategies sustain supracompetitive prices, and rely on counterfactual simulations to recover the profit terms that enter the IC constraints. There are also important differences. Igami and Sugaya assume all firms either engage in maximal collusion or revert to Nash-Cournot equilibrium. Thus, some interesting aspects of our model, such as partial coalitions and the leader's ability to adjust the supermarkup to satisfy incentive compatibility, are not present in their setup.

Also similar is contemporaneous research of Eizenberg and Shilian (2019). The paper

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

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

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

tests for Nash-Bertrand pricing in each of 40 different Israeli food sectors. Marginal costs are obtained from supply-side first order conditions, which allows for the recovery of the profit terms that enter IC constraints, via counterfactual simulations. The critical discount factors necessary to sustain maximal collusion under grim trigger strategies are then calculated and compared across sectors. Mergers are not considered. Furthermore, it is not clear whether mergers would generate coordinated effects in their framework, as they increase punishment payoffs (assuming Nash reversion) yet do not affect the benefit of maximal collusion. On this point, see also the paper of Davis and Huse (2010) on the network server industry.

Finally, our research relates to a number of articles that use a conduct parameter approach to test for changes in the equilibrium concept (e.g., Porter (1983), Ciliberto and Williams (2014), Igami (2015), Miller and Weinberg (2017)). The closest of these is Miller and Weinberg, as it examines the MillerCoors joint venture. Our methodology also tests for a shift in the equilibrium concept, but does so using an explicit model of the underlying supergame. As the objects governing incentive compatibility can be recovered, our methodology allows for the consideration of a much broader set of counterfactual scenarios, and may contribute to an improved understanding of post-merger coordinated effects.

2 The U.S Beer Market

2.1 Background

Most beer sold in the Unites States is produced by a handful of large brewers that compete across the country. These brewers compete in prices, product introduction, advertising, and periodic sales. The product offerings typically are characterized as differentiated along multiple dimensions, including taste, calories, brand image, and package size. The beer industry differs from typical retail consumer product industries in its vertical structure because of state laws regulating the sales and distribution of alcohol. Large brewers are prohibited from selling beer directly to retail outlets. Instead, they typically sell to state-licensed distributors, who, in turn, sell to retailers. Payments along the supply chain cannot include slotting fees, slotting allowances, or other fixed payments between firms.[7] 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)).

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

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

Table 1 summarizes the revenue shares of the major brewers over 2001-2011. In the early years of the sample, Anheuser Busch, SABMiller, and Molson Coors (domestic brewers) account for 61%-69% of revenue while Grupo Modelo and Heineken (importers) account for another 12%-16% of revenue.[8] Midway through the sample, in June 2008, SABMiller and Molson Coors consolidated their U.S. operations into the MillerCoors joint venture. The DOJ reviewed the transaction as a merger and elected not to challenge on the basis that cost savings in distribution likely would offset any loss of competition. Subsequent academic research suggests that sizable costs savings were realized but were dominated by adverse competitive effects (Ashenfelter et al. (2015), Miller and Weinberg (2017)).

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

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

> ABI and MillerCoors typically announce annual price increases in late summer

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

> 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]

In the model, price leadership acts as an equilibrium selection device, essentially resolving the coordination problem that firms may face due to the folk theorem (Whinston (2006)). The legal documents are helpful in ascertaining whether such a mechanism is consistent with empirical setting. 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."[10]

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

2.2 Data

We use the retail scanner data of the IRI Academic Database (Bronnenberg et al. (2008)), which provides weekly revenue and unit sales by UPC code for a sample of stores over 2001-2011. We restrict attention to supermarkets, which account for 20% of off-premise beer sales (McClain (2012)).[11] We aggregate the data to the product-region-period-year level. Products are brand×size combinations. We consider alternative period definitions—months and quarters—to provide some robustness to any sales and consumer stockpiling behavior. We restrict attention to 13 flagship brands sold as six packs, 12 packs, 24 packs, and 30 packs. We measure quantities based on 144-ounce equivalent units, the size of a 12-pack, and measure price as the ratio of revenue to equivalent unit sales. Table 2 provides summary statistics on prices and volume shares.[12]

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

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

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

[12] The flagship brands include Bud Light, Budweiser, Michelob, Michelob Light, Miller Lite, Miller Genuine Draft, Miller High Life, Coors Light, Coors, Corona Extra, Corona Extra Light, Heineken, and Heineken Light. The most popular omitted brands tend to be regional brands or subpremium brands with lower

Table 2: Prices and Conditional Volume Shares in 2011

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

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

We combine the scanner data with demographic information obtained from the Public Use Microdata Sample (PUMS) of the American Community Survey to help incorporate consumer heterogeneity into the demand model. The PUMS data are available over 2005-2011. Finally, we rely on the driving miles between each IRI region and the nearest brewery for each product and the price of diesel fuel. The former is obtained using Google Maps and the latter is obtained from the U.S. Energy Information Administration. We model transportation costs as varying with the interaction of driving miles and the fuel price. All prices and incomes are deflated using the Consumer Price Index and reported in 2010 dollars. The final sample comports with that of Miller and Weinberg (2017), and we refer readers to that article for more extensive details on the data.

2.3 Prices

Figure 1 shows the time path of average prices over 2001-2011 for each firm's most popular 12 pack: Bud Light, Miller Lite, Coors Light, Corona Extra, and Heineken. The red vertical line at June 2008 marks the closing of the Miller/Coors merger. As shown, the prices of the domestic beers increase starkly after the merger, breaking a downward pre-merger trend. The prices of the more expensive import brands continue on trend before and after the merger, suggesting that the change in domestic prices may not be due to common

prices. We combine 24 packs and 30 packs in the construction of our products because whether 24 packs or 30 packs are sold tends to depend on region-specific historical considerations.

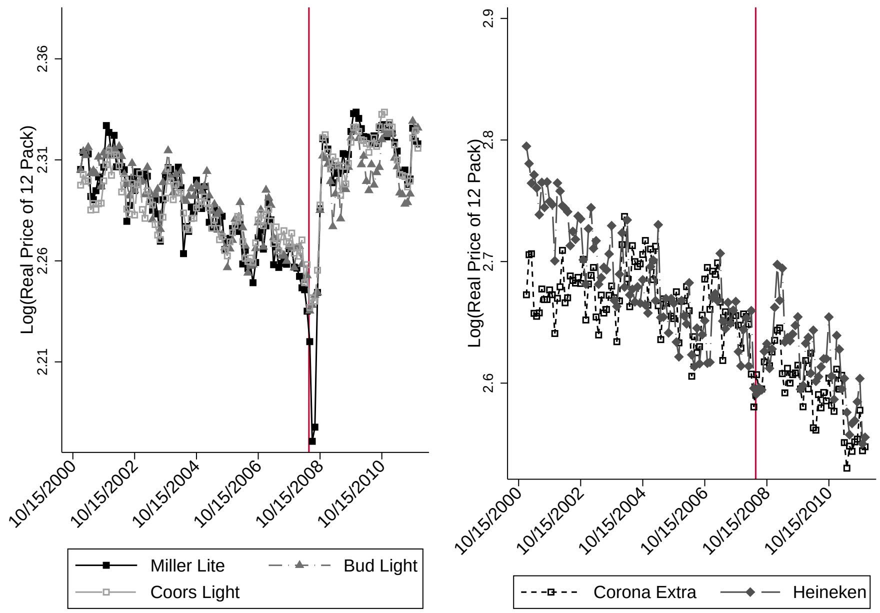

Figure 1: Average Retail Prices of Flagship Brand 12-Packs

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).

cost or demand factors. Miller and Weinberg (2017) provide econometric evidence that the visual break apparent in the figure is not due artificial sample selection, trends, or omitted variables. Further, qualitative evidence culled from regulatory filings is consistent with softened competition among the domestic brewers in the post-merger period.[13]

Of particular interest is that the post-merger price increases of ABI are commensurate with those of MillerCoors. It can be difficult to explain such a pattern with standard static models of differentiated products competition. Indeed, the main result of Miller and Weinberg (2017) is that coordinated pricing remains a plausible explanation after one accounts for the unilateral effects of the Miller/Coors merger and market conditions. In the present

[13]The 2005 SABMiller annual report describes "intensified competition" and an "extremely competitive environment." The 2005 Anheuser-Busch report states that the company was "collapsing the price umbrella by reducing our price premium relative to major domestic competitors." SABMiller characterizes price competition as "intense" in its 2006 and 2007 reports. Language changes markedly after the merger. In its 2009 report, SABMiller attributes increasing earnings before interest, taxes, and amortization expenses to "robust pricing" and "reduced promotions and discounts." In its 2010 and 2011 reports, it references "sustained price increases" and "disciplined revenue management with selected price increases." See SABMiller's Annual Report of 2005 (p. 13), 2006 (p. 5), 2007 (pp. 4 and 8), 2009 (pp. 9 and 24), 2010 (p. 29), and 2011 (p. 28) and Anheuser-Busch's Annual Report in 2005 (p. 5).

paper, we push further, and explore the conditions under which the higher post-merger prices shown in the figure can be attributed to price leadership behavior on the part of the domestic brewers, as well as why importers may not have participated in coordinated pricing.

2.4 Demand

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

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

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

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

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

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

The quantity sold of good j in region r and period t is given by

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

where I_{igrt} and I_{irt} are the McFadden (1978) inclusive values, M_r is the market size of the region, $\delta_{jrt} = x_j\beta + \alpha p_{jrt} + \sigma_j^D + \tau_t^D + \xi_{jrt}$, and $\mu_{ijrt} = [p_{jrt}, x_j]' * \Pi D_i$.[14] We assume market sizes 50% greater than the maximum observed unit sales within each region. Expressions for the price derivatives of demand are supplied in Grigolon and Verboven (2014).

Table 3 presents the results from four specifications examined in Miller and Weinberg (2017). The first two (RCNL-1 and RCNL-2) allow income to affect the price parameter, thereby relaxing cross-price elasticities between more affordable domestic beers and the more expensive imported beers. The latter two (RCNL-3 and RCNL-4) allow income to affect tastes for imported beers directly. Most coefficients are precisely estimated and take the expected signs. The median own price elasticities range from -4.45 to -6.10. The price elasticities of market demand are much smaller, ranging from -0.60 to -0.72, due to the magnitude of the nesting parameter. Most substitution occurs among the inside goods, rather than between the inside goods and the outside good. We provide additional summary statistics on product-level and firm-level elasticities in Appendix Tables B.1 and B.2.[15]

3 Empirical Model of Price Leadership

3.1 Primitives

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

The game features $t = 0, \ldots, \infty$ periods. At the beginning of the game, $t = 0$, the leader designates a set of firms, $\mathbb{C}$, as the coalition. The leader is always in the coalition. Other firms in the coalition are "followers," and firms outside the coalition are "fringe firms." In each subsequent period, $t = 1, \ldots, \infty$, an economic state Ψ_t is realized and observed by all firms. Competition then plays out in two stages:

[14] The normalization on the mean indirect utility of the outside good yields $I_{i0rt} = 0$, while the inclusive value of the inside goods is $I_{i1rt} = (1-\rho)\log\sum_{j=1}^{J_{rt}}\exp((\delta_{jrt} + \mu_{ijrt})/(1-\rho))$ and the inclusive value of all goods is $I_{irt} = \log(1 + \exp I_{i1rt})$.

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

Table 3: Demand Estimates

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

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

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

(ii) All firms set prices simultaneously, given the announced supermarkup m_t and history h_t, and receive payoffs according to a profit function we introduce below.

We have chosen the timing of the game to mimic a common practice in which one firm announces a price change before the new prices become available to consumers.[16] However, given common knowledge of the economic state, the first stage is not a theoretical necessity. The price leadership equilibrium (defined later) also can be obtained in a standard repeated pricing game with a particular assumption on equilibrium selection.

[16]Not all leadership/follower behavior has this feature (e.g., Byrne and de Roos (2019)).

Payoffs are determined by a standard profit function and a reduced-form expression for antitrust risk that coalition firms incur by adopting the supermarkup. The profit function of firm f in period $t = 1 \ldots, \infty$ is given by

$$\sum_{j \in \mathbb{J}_f} \pi_j(p_t, \Psi_t) = \sum_{j \in \mathbb{J}_f} (p_{jt} - mc_j(W_t)) q_j(p_t, X_t) \tag{4}$$

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

$$p_{ft} + \left(\frac{\partial q_f(p_t, X_t)}{\partial p_f}^T \right)^{-1} q_f(p_t, X_t) = mc_f(W_t) \tag{5}$$

where we apply the f subscript to refer to vectors of firm f's prices, quantities, and marginal costs. If all firms maximize per-period profit then the Nash-Bertrand prices, $p_t^{NB}(\Psi_t)$, obtain. We assume a unique solution exists, which can be verified with nested logit demand (Mizuno (2003)) but need not hold with the RCNL demand model. Coalition firms that set $p_{ft} = p_{ft}^{NB} + m_t$ also incur a fixed cost, $R(m_t)$, with $R(0) = 0$ and $R'(m) \geq 0$, which we motivate as arising from antitrust risk. We discuss micro-foundations in Section 5.3.

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

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

This treatment of the economic states is theoretically appealing because it avoids certain scenarios in which price leadership unravels due to an adverse realization of Ψ_t.[17] As will be developed, deviation from the leader's proposed supermarkup does not occur on the equilibrium path because the leader adjusts the supermarkup to satisfy incentive compatibility constraints. Finally, we assume that firm actions do not affect the economic states.

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

3.2 Equilibrium

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

Consider the pricing stage in some arbitrary period t. Each coalition firm $f \in \mathbb{C}$ "accepts" the leader's proposed supermarkup m_t if it prices according to $p^{PL}_{ft}(m_t; \Psi_t) = p^{NB}_{ft}(\Psi_t) + m_t$. Fringe firms accept simply by pricing on their best reaction functions. Thus, let $p^{PL}_{ft}(m_t; \Psi_t)$ for $f \notin \mathbb{C}$ solve the first order conditions of equation (5), taking as given the coalition prices and the prices of other fringe firms. Firms "reject" m_t if they select some other price. Given the beliefs to be enumerated below, two particular forms of rejection are relevant. First, let the vector $p^{D,f}_t(m_t; \Psi_t)$ collect the prices that arise if firm f solves equation (5) with the (correct) anticipation that other firms accept. Second, let the vector $p^{NB}_t(\Psi_t)$ collect the Nash-Bertrand prices that solve equation (5) for all firms. We refer to $p^{D,f}_t(\cdot)$ and $p^{NB}_t(\cdot)$ as deviation and Nash-Bertrand prices, respectively.

Let the *slack function* capture the present value of price leadership less the present value of deviation, under the assumption that deviation is punished in all future periods. For a coalition firm, this difference can be expressed

$$\begin{aligned} s_{ft}(m_t; \Psi_t) &= \overbrace{\frac{\delta}{1-\delta} E_\Psi \left[\sum_{j \in \mathbb{J}_f} \pi^{PL}_j(\Psi) - R^*(\Psi) - \sum_{j \in \mathbb{J}_f} \pi^{NB}_j(\Psi) \right]}^{\text{Expected Future Net Benefit of Price Leadership}} \\ &- \underbrace{\left[\sum_{j \in \mathbb{J}_f} \pi_{jt}\left(p^{D,f}_t(m_t, \Psi_t); \Psi_t\right) - \sum_{j \in \mathbb{J}_f} \pi_{jt}\left(p^{PL}_t(m_t, \Psi_t); \Psi_t\right) + R(m_t) \right]}_{\text{Immediate Net Benefit of Deviation}} \end{aligned} \quad (6)$$

where $\delta \in (0, 1)$ denotes a common discount factor, $\pi^{NB}(\phi) \equiv \pi(p^{NB}(\Psi); \psi)$ is the profit from Nash-Bertrand, $\pi^{PL}(\Psi) \equiv \pi\left(p^{PL}(m^*(\Psi), \Psi); \Psi\right)$ is price leadership profit evaluated at $m^*(\Psi)$, defined below as the leader's optimal supermarkup, and $R^*(\Psi) \equiv R(m^*(\Psi))$.

The slack functions of fringe firms do not include the antitrust risk terms but otherwise are identical. The slack functions can take positive or negative values for coalition firms, depending on m_t and Ψ_t, but are weakly positive for fringe firms by construction.

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

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

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

$$\begin{aligned} m_t^*(\Psi) \quad &= \quad \underset{m \geq 0}{\arg\max} \sum_{j \in \mathbb{J}_1} \pi_{jt}\left(p_t^{PL}(m, \Psi_t); \Psi_t\right) - R(m) \\ &s.t. \quad s_{ft}(m; \Psi_t) \geq 0 \quad \forall f \in \mathbb{C} \end{aligned} \tag{7}$$

As the slack function equals zero at $m_t = 0$, a solution to the leader's constrained maximization problem always exists.[18] It follows that punishment never occurs on the equilibrium path because the leader can always find some supermarkup that satisfies IC of coalition firms, even if this implies Nash-Bertrand prices for some realizations of the economic state.

Finishing, in the coalition selection stage ($t = 0$), the leader selects the coalition that maximizes the present value of its payoffs, under the belief of equilibrium play in subsequent periods. In numerical experiments, we have confirmed that partial coalitions can be optimal for the leader. Typically this occurs if there is substantial heterogeneity in the slack func-

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

tions, which can allow for higher supermarkups with a partial coalition as IC constraints are relaxed. However, heterogeneity is not necessary for partial coalitions generally (e.g., as in d'Aspremont et al. (1983), Donsimoni et al. (1986), and Bos and Harrington (2010)).

The existence of non-degenerate price leadership outcomes, which we define as involving multiple coalition firms and positive supermarkups, is not guaranteed. To help frame the empirical analysis, we provide a pair of existence results in the following proposition:

Proposition 1 (Existence): *Given the set of coalition firms $\mathbb{C}$, suppose the solution to the leaders' constrained maximization problem generates*

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

for all $f \in \mathbb{C}$. Then, for any candidate supermarkup m, there exists some $\tilde{\delta}(m)$ such that if $\delta > \tilde{\delta}(m)$ then $s_f(m;\Psi) \geq 0$ for all $f \in \mathbb{C}$. Furthermore, for any $\delta \in (0,1)$, if $R(m) = 0$ then there exists some $m > 0$ such that $s_f(m;\Psi) \geq 0$ for all $f \in \mathbb{C}$.

The proof is provided in Appendix A. The first part of the proposition is standard: any supermarkup that raises profit above the Nash-Bertrand level satisfies IC in the pricing stage provided that firms are sufficiently patient. The second part of the proposition states that, in the absence of antitrust risk, there exists a strictly positive supermarkup that satisfies IC. Thus, antitrust risk is important insofar as it creates the theoretical possibility that some market structures cannot support positive supermarkups.

3.3 Discussion

Figure 2 provides a numerical illustration of a PLE, based on a market with two coalition firms and logit demand.[19] The Nash-Bertrand equilibrium is easily identifiable as the intersection of the coalition firms' reaction functions. In selecting the supermarkup, leader essentially considers symmetric price increases above Nash-Bertrand equilibrium, plotted in the figure as the 45-degree line connecting to the Nash-Bertrand equilibrium. Here the supermarkup that maximizes the leader's profit (the "Unconstrained PLE") violates IC, so the PLE features a smaller supermarkup. The PLE is not generally Pareto optimal for the

[19] Notes: The figure shows results obtained with two coalition firms and one fringe firm. Demand is given by $q_i = \frac{\exp(\beta_i - \alpha p_i)}{1+\sum_k \exp(\beta_k - \alpha p_k)}$, with the parameterizations $\beta_1 = \beta_2 = 3$, $\beta_3 = 1$, $\alpha = 1.5$. Marginal costs are $mc_1 = mc_2 = 0$ and $mc_3 = 2$, the discount factor is $\delta = 0.4$, and antitrust risk is $R(m) = 0$.

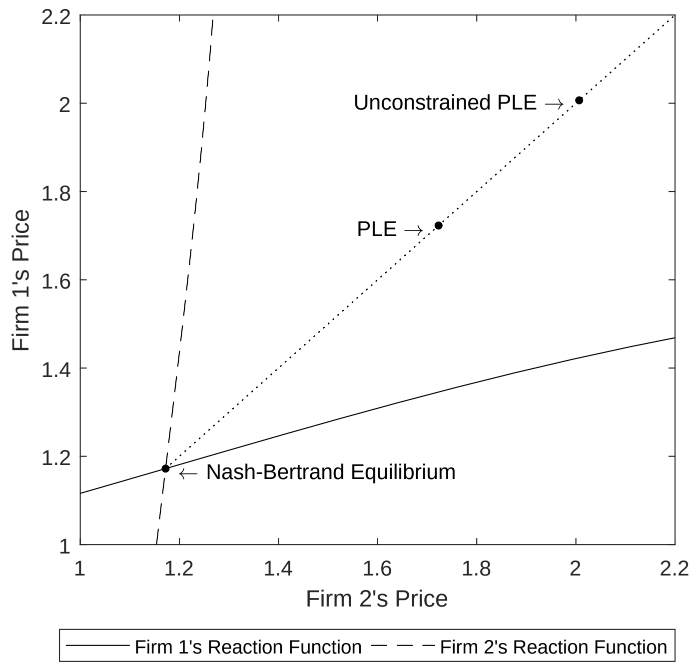

Figure 2: Illustration of the Price Leadership Equilibrium

coalition because the leader acts in its own interest.[20] Otherwise, the model closely resembles the canonical Rotemberg and Saloner (1986) model of collusion.

Importantly, the leader's price announcement selects an equilibrium in the model because (by assumption) firm beliefs are shaped by the announcement. To see this, consider that if firms believe their competitors will price ala Nash-Bertrand, then the announcement would have no bearing on equilibrium outcomes. The conditions under which it is reasonable to assume cheap talk—such as the price announcement—affects beliefs have been debated in the literature (e.g., Aumann (1990), Farrell and Rabin (1996)).[21] In support of our assumption, we cite to recent experimental evidence that price announcements can help facilitate coordination in repeated oligopoly games (Harrington et al. (2016)).

[20]See Asker (2010) and Asker et al. (2018) for two empirical examples of inefficient coordination.

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

4 Empirical Implementation

In this section, we discuss the conditions under which the supermarkups can be estimated with data on prices and quantities. The estimation procedure tracks standard industrial organization methodologies: for any candidate set of supermarkups, one can recover marginal costs, isolate a residual from the cost function, and evaluate a loss function by interacting the residual with instruments taken from the demand-side of the model. Estimation does not require an evaluation of IC. Nonetheless, with the supermarkups in hand, one can test whether IC binds. In the affirmative case, it also is possible to jointly identify the discount factor and the antitrust risk, a matter to which we return in Section 5.3.

4.1 Identification of Marginal Costs

The identification strategy is a variant on the standard methodology of inferring marginal costs from the Nash-Bertrand first order conditions, as introduced in Rosse (1970). To illustrate, we stack equation (5) for each firm and evaluate at Nash-Bertrand prices, which obtains the familiar solution that marginal revenue equals marginal cost:

$$mr_t(p_t^{NB}, X_t, \Omega_t) \equiv p_t^{NB} + \left[\Omega_t \circ \left(\frac{\partial q_t(p_t, X_t)}{\partial p_t} \bigg|_{p=p_t^{NB}} \right)^T \right]^{-1} q_t(p_t^{NB}, X_t) = mc_t(W_t) \tag{8}$$

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

In settings which feature Bertrand competition, equation (8) allows marginal costs to be recovered given knowledge of demand and data on prices. Our application is more complicated. As competition may not be Bertrand, observed prices (p_t) may not correspond with Bertrand prices (p_t^{NB}). It follows that equation (8) cannot be evaluated directly. Nonetheless, if the econometrician has knowledge of the supermarkup, then Bertrand prices and marginal costs can be recovered. We state this result as a proposition:

Proposition 2 (Identification). *Suppose the econometrician has knowledge of the demand system, the identities of the coalition firms (i.e., $\mathbb{C}$), and the supermarkup (m). Then Bertrand prices and marginal costs are identified.*

The proof is constructive and proceeds in four steps, each of which is easily verified given the maintained assumptions. We enumerate the steps here as they are central to the estimation procedure. Suppressing time subscripts, the steps are:

1. Infer mc_j for each fringe firm $j \notin \mathbb{C}$ from the first order conditions of equation (5). This can be done with observed prices because fringe firms maximize per-period profit.

2. Obtain $p_k^{NB} = p_k - m$ for each coalition firm $k \in \mathbb{C}$.

3. Compute p_j^{NB} for each fringe firm $j \notin \mathbb{C}$ by simultaneously solving the first order conditions of equation (5), given the inferred marginal costs mc_j and holding the prices of coalition firms fixed at the Bertrand level (i.e., $p_k = p_k^{NB}$ for each $k \in \mathbb{C}$).

4. Infer mc_k for each coalition firm $k \in \mathbb{C}$ from the first order conditions of equation (5), evaluated at the Bertrand prices p^{NB} obtained in steps 2 and 3.

4.2 Specification of Marginal Costs

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

$$mc_{jrt}(W_{rt}) = w_{jrt}\gamma + \sigma_j^S + \tau_t^S + \mu_r^S + \eta_{jrt} \tag{9}$$

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

4.3 Estimation

The objects of interest in estimation are $\theta_0 = (m_t, \gamma, \sigma_j^S, \tau_t^S, \mu_r^S)$. For each candidate $\widetilde{\theta}$, one can apply the four steps necessary to recover Bertrand prices and marginal costs (Proposition 2). The implied residuals then obtain:

$$\eta_{jrt}^*(\widetilde{\theta}; \Psi_t) = mr_{jrt}(p_{rt}^{NB}(\widetilde{m}_t; \Psi_t); X_t, \Omega_t) - w_{jrt}\widetilde{\gamma} - \widetilde{\sigma}_j^S - \widetilde{\tau}_t^S - \widetilde{\mu}_r^S \tag{10}$$

Our notation emphasizes that inferences on Bertrand prices depend on the candidate supermarkup and the economic state. Importantly, marginal revenue is endogenous because residual costs enter implicitly through Bertrand prices. Valid instruments can be constructed from aspects of the economic state that enter demand (X_t) or ownership (Ω_t) and that meet the population moment condition $E[Z' \cdot \eta^*(\theta_0)] = 0$, where $\eta^*(\theta_0)$ is a stacked vector of residuals and Z is the matrix of instruments.[22]

The corresponding generalized method-of-moments estimate is

$$\widehat{\theta} = \arg\min_{\theta} \eta^*(\theta; X, W, \Omega)' Z A Z' \eta^*(\theta; X, W, \Omega) \tag{11}$$

where A is some positive definite weighting matrix. We have exact identification in our application, given instruments that we define below, so A is an identity matrix. We concentrate the fixed effects and the marginal cost parameters out of the optimization problem using OLS to reduce the dimensionality of the nonlinear search. We cluster the standard errors at the region level and make an adjustment to account for the incorporation of demand-side estimates into the economic state, following Wooldridge (2010).[23]

4.4 Instruments

An important departure from the literature is that the objects of interest in estimation include the supermarkup, which is not a structural parameter but instead a strategic choice variable that solves a constrained maximization problem. A simple example illustrates the ramifications for identification: Suppose that the econometrician attempts to use a single binary variable, Z_1, taken from the economic state, as the excluded instrument. The model is under-identified because variation in Z_1 implies the existence of two supermarkups that must be estimated. Adding a second instrument, Z_2, does not solve the under-identification problem as any additional variation provided by the second instruments implies the existence of at least one additional supermarkup. Iterating, it follows that no set of instruments is sufficient to identify the supermarkups absent additional restrictions on the model.

We make progress by assuming Bertrand pricing ($m_t = 0$) in periods predating the

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

[23]See also Appendix E of Miller and Weinberg (2017) for details on how the adjustment is implemented in a similar estimation model.

Miller/Coors merger, which resolves the otherwise intractable under-identification problem.[24] The empirical and qualitative evidence presented in Section 2.3 provides support for the restriction. The instruments we employ include separate post-merger indicator variables for the ABI, Miller, and Coors brands, respectively. Of these, only the ABI post-merger instrument is excluded from the marginal cost function. Thus, identification exploits a change in ownership structure and the resulting shifts in ABI's marginal revenue schedules that occur as MillerCoors adjusts its prices (under the Bertrand assumptions). Given the product and period fixed effects in the marginal cost function, the excluded instrument is valid if the average residual costs of ABI do not change contemporaneously with the Miller/Coors merger, relative to the average residual costs of the fringe firms.

The ABI post-merger instrument is sufficient to identify a single supermarkup, and indeed our main results are developed under the assumption that the coalition sets the same supermarkup in every post-merger period and region. Alternatively, it is possible to estimate region-specific or period-specific supermarkups by interacting the ABI post-merger instrument with region or period fixed effects, respectively, so as to maintain exact identification.[25] Doing so does not materially affect our conclusions, however, so we provide results for the simpler model. To the extent there is heterogeneity in the true supermarkup, our estimates can be interpreted as providing something of an average.

5 Estimation Results

5.1 Baseline Specifications

Table 4 contains the baseline supply-side estimates. As described, each column corresponds to one of the baseline demand specifications. The marginal cost functions incorporate product (i.e., brand×size), period (month or quarter), and region fixed effects in all cases. The estimates of the supermarkup are range from 0.596 to 0.738 and are precisely estimated. To put these supermarkups in context, a supermarkup of \$0.60 above Nash-Bertrand prices is

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

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

Table 4: Baseline Supply Estimates

	Parameter	RCNL-1	RCNL-2	RCNL-3	RCNL-4
Estimation Results					
Supermarkup	m	0.643	0.596	0.738	0.709
		(0.025)	(0.027)	(0.034)	(0.033)
Miller×Post-Merger	γ_1	-0.540	-0.533	-0.583	-0.416
		(0.007)	(0.007)	(0.005)	(0.002)
Coors×Post-Merger	γ_2	-0.826	-0.831	-0.914	-0.666
		(0.009)	(0.009)	(0.006)	(0.004)
Distance	γ_3	0.168	0.164	0.172	0.153
		(0.001)	(0.001)	(0.001)	(0.001)
Supplementary Results					
Unconstrained supermarkup		2.69	2.57	3.25	2.56
		[2.64,2.77]	[2.49, 2.66]	[3.18, 3.31]	[2.48,2.63]
Negative marginal costs		0.12%	0.09%	0.26%	0.03%

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

equivalent to about six percent of the average price of a 12 pack. The marginal cost intercepts of Miller and Coors brands drop with the joint venture. As the distance estimate is positive, a second source of efficiencies from Miller/Coors arises as production of Coors brands and, to a lesser extent Miller brands, is moved to breweries closer to retail locations.[26] The unconstrained supermarkups, which we obtain with a post-estimation simulation, greatly exceeds the estimated supermarkups. Thus, at least one IC constraint must bind.

Table 5 provides the average markup for each product in the data both before and after the Miller/Coors merger, based on the RCNL-1 specification. Across all 89,619 brand–size–month–region observations, the average markup is \$3.37 on an equivalent-unit basis, which accounts for 32% of the retail price. The average markups on ABI 12 packs tend to be about \$0.70 higher in the post-merger periods, which reflects the combination of higher Nash-Bertrand prices and the supermarkup. The markups on Miller 12 packs increase by about \$1.35 and the markups on Coors products increase by about \$1.75. Those changes reflect the combined impact of higher Nash-Bertrand prices, the supermarkup, and lower marginal

[26]Miller and Weinberg (2017) estimate similar marginal cost parameters, and we refer reader to that article for a more in depth analysis of the merger efficiencies.

Table 5: Brewer Markups from RCNL-1

Brand	6 Packs		12 Packs		24 Packs	
	Pre	Post	Pre	Post	Pre	Post
Bud Light	3.62	4.36	3.52	4.27	3.45	4.16
Budweiser	3.76	4.51	3.65	4.41	3.55	4.28
Coors	2.64	4.37	2.53	4.32	2.43	4.20
Coors Light	2.44	4.22	2.35	4.14	2.28	4.05
Corona Extra	3.18	3.05	2.95	2.83	2.92	2.90
Corona Light	2.93	2.80	2.69	2.60	2.78	2.72
Heineken	3.08	3.03	2.89	2.83	3.03	3.15
Heineken Light	2.80	2.73	2.56	2.47	2.65	2.62
Michelob	3.68	4.54	3.62	4.45	3.36	4.36
Michelob Light	3.61	4.37	3.53	4.26	3.47	4.03
Miller Gen. Draft	2.88	4.27	2.77	4.17	2.70	4.10
Miller High Life	2.93	4.27	2.81	4.21	2.75	4.14
Miller Lite	2.88	4.26	2.78	4.18	2.70	4.08

Notes: This table provides the average markups for each brand–size combination separately for the pre-merger and post-merger periods, based on the RCNL-1 specification results shown in Tables 3 and 4.

costs. The markups on imported beers do not change much over the sample period.

To build intuition about the model, we conduct an exercise in which we use counterfactual simulations to examine a series of alternative supermarkups, $\tilde{m} = [0.00, 0.01, 0.02, \ldots, 3.00]$. For each $\tilde{m}$ we obtain economic outcomes under price leadership and deviation. The exercise requires knowledge of demand and marginal costs. While we recover profit functions that inform the IC constraints, an analysis of IC is not required for the exercise.

Figure 3 summarizes the simulation results. Panel A shows the profit functions of ABI. Price leadership profit is an inverted-U in the supermarkup, with a maximum just over \$2.50 (see also Table 4). Deviation profit increase monotonically in the supermarkup because higher supermarkups correspond to higher MillerCoors prices, and the static best-response profits of ABI increase in MillerCoors prices. The gap between the two profit function grows in the supermarkup. At our estimate, marked with the blue line, MillerCoors profits are about seven percent higher than Nash-Bertrand and deviation profits are about eight percent higher. That deviation is only slightly more profitable than price leadership is striking; we show in the next section that the finding seems to be driven by the demand elasticities rather that the logit demand assumption itself.

Panel B shows the prices and shares of ABI. Under leadership, the price and share functions have slopes of quite similar magnitudes (though opposite in sign). As the functions are indexed relative to Nash-Bertrand, this implies a coalition elasticity of demand around one. At our estimate, ABI prices are about eight percent higher than Nash-Bertrand, and

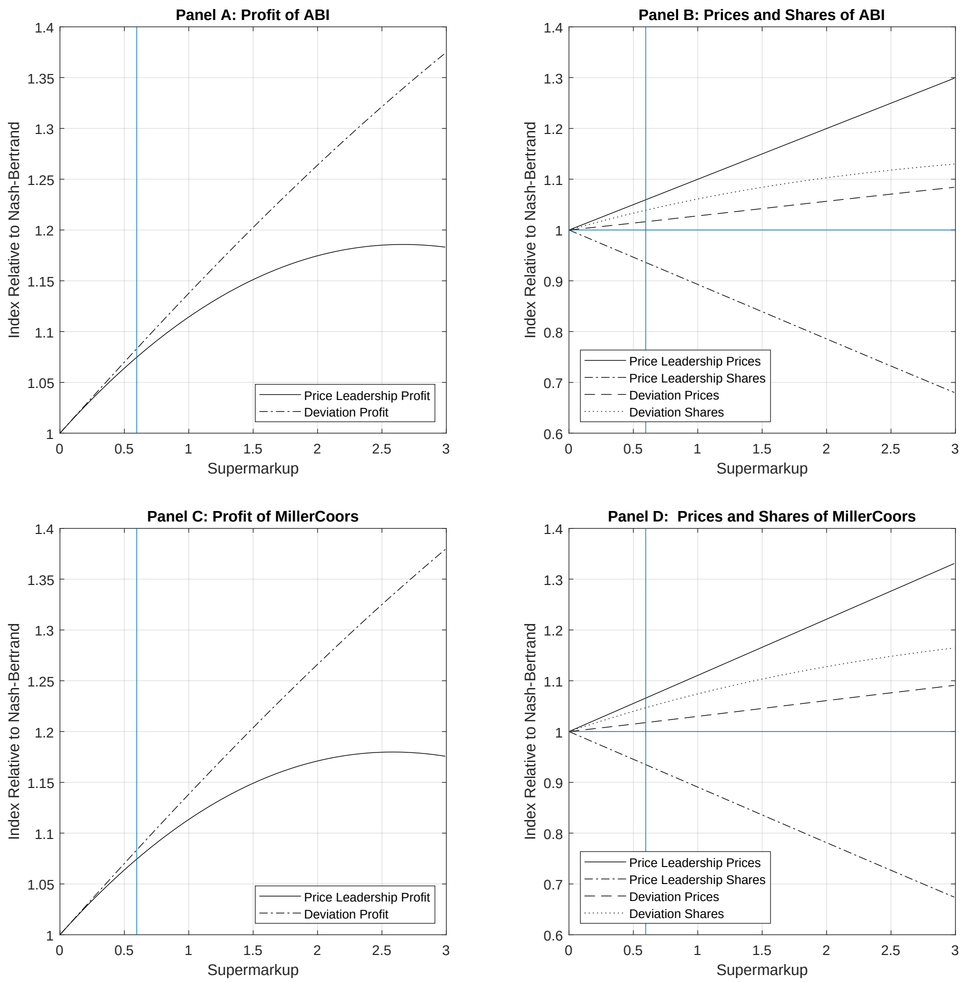

Figure 3: Profit, Prices and Shares with Price Leadership and Deviation

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

shares are about eight percent lower. The deviation price and share functions increase with the supermarkup. Thus, the prices of ABI and MillerCoors appear to be strategic complements across a wide support. Panels C and D show that the statistics for MillerCoors are essentially the same, which reflects that ABI and MillerCoors have similar markups and firm elasticities in the post-merger periods (e.g., Table 5 and Appendix Table B.2).

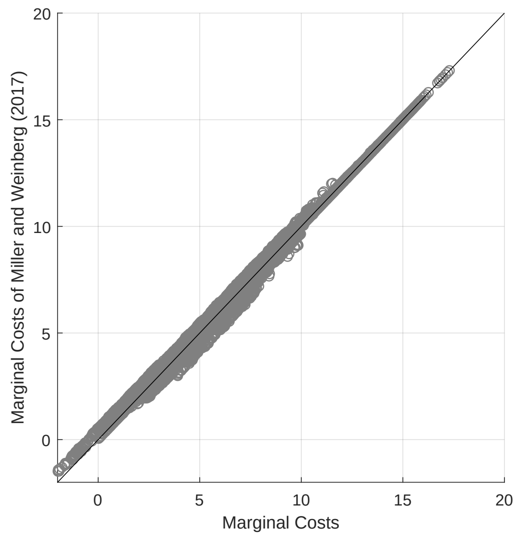

Figure 4: Empirical Distribution of Marginal Costs

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

5.2 Alternative Models

We perform two additional sets of analyses to probe the sensitivity of the results to alternative modeling choices. First, we compare the marginal costs we recover with the price leadership model to those obtained in the conduct parameter model of Miller and Weinberg (2017) using identical demand estimates. For reference, Miller and Weinberg assume the brewers set prices according to satisfy

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

where Ω_t is an ownership matrix, κ is a conduct parameter, and the operation $\circ$ is element-by-element matrix multiplication. The $(j,\, k)$ element of the ownership matrix equals one if products j and k are produced by the same firm, κ if they are sold by ABI and MillerCoors and the period postdates the merger, and zero otherwise. The model nests post-merger Nash-Bertrand ($\kappa = 0$) and joint profit maximization for ABI/MillerCoors ($\kappa = 1$). Figure 4 plots the marginal costs of the two models. The dots, each representing a product-region-year observation, fall along the 45-degree line. Thus, our structural framework generates similar results to models that rely on alternative behavioral assumptions.

The second exercise examines whether the logit assumption is important in generating the results that deviation does not increase profit much for supermarkups of the magnitude we estimate (e.g., Figure 3). To that end, we calibrate a linear demand system

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

such that the elasticities exactly match those of the RCNL when evaluated at the average prices and quantities in 2011.[27] We then repeat the counterfactual exercise of the preceding section, simulating price leadership and deviation for $m \in [0, 2]$. The results, provided in Appendix Figure B.2, closely resemble the those developed with the RCNL. Thus, we conclude that our findings are not overly dependent on the logit assumption.

5.3 Calibrating the Slack Functions

We make three modifications to the slack functions before bringing them to the data. First, we replace the assumption of a stochastic economic state with an assumption that the entire sequence $(\Psi_\tau)_{\tau=1}^{\infty}$ is common knowledge in every period. This raises the theoretical possibility that price leadership could unravel if positive supermarkups cannot be sustained beyond some future date, as in Igami and Sugaya (2018). However, unraveling does not occur in our application by construction, as we model the future using infinite repetitions of the year 2011. Second, we assume that deviation profit is earned for a full calendar year before punishment ensues, which we motivate based on the observed practice of annual list price adjustments. We discuss timing assumptions below. Finally, we sum the functions across regions, creating a single IC constraint for each coalition firm.[28]

Among the objects in the slack functions, the profit terms are easily recovered via counterfactual simulations given knowledge of $(\Psi_\tau)_{\tau=1}^{\infty}$, leaving the discount factor and the antitrust risk as the only unknowns (see equation (6)). Antitrust risk plays an important role in the model because it creates the theoretical possibility that some market structures cannot support positive supermarkups (Proposition 1). There are a variety of reasons that tacit coordination may impose explicit or implicit costs on firms. A common interpretation is legal risk. For instance, evidence of price leadership has been considered in a number of price-fixing lawsuits when courts have weighed whether discovery should be granted to

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

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

the plaintiffs.[29] Further, historical evidence of pricing coordination sometimes is cited by antitrust authorities as contributing to a decision to challenge a merger.[30]

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

Figure 5 plots the values of (δ, ϕ) that balance the MillerCoors IC constraints in 2011:Q4. With $\phi = 0$, an annualized discount factor of 0.11 balances IC, and greater values of ϕ require higher discount factors. We attempt to remain agnostic about what constitutes an economically reasonable discount factor. The main reason is that the IC constraints incorporate strong timing assumptions about deviation and punishment that are impossible to verify as they are off the equilibrium path (and therefore not observed in the data). In particular, there are a number of reasons to suspect firms cannot or would not implement infinite Nash reversion. Thus, we interpret the discount factor as a reduced-form parameter that summarizes both the patience of firms and the timing of the game.[31]

Figure 6 plots the slack in IC of ABI (Panel A) and MillerCoors (Panel B) over the range of supermarkups $m \in [0, 0.8]$. Four alternative assumptions are used to calibrate the IC constraints: $\delta = 0.7$, $\delta = 0.5$, $\delta = 0.3$, and $\phi = 0$. In each case, we select the free parameter such that IC of MillerCoors binds at the estimated supermarkup of 0.596 (Figure 5 provides the mapping). We consider a number of candidate supermarkups, $m = 0.00, 0.01, 0.02, \ldots$, and for each we use counterfactual simulations to obtain profit with price leadership, deviation, and punishment. Pairing this with the calibrated (δ, ϕ) parameters, we recover firm-specific slack functions. The figure shows that slack exists in the IC constraints

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

[30] Interestingly, a prime example is ABI's attempted acquisition of Modelo in 2012-2013, which the DOJ challenged in part due to a concern it would eliminate a constraint on coordinated price increases. We return to the economic effects of the proposed ABI/Modelo merger in Section 6.1. A second example is the Tronox/Cristal merger in the titanium dioxide industry (*Re: Fed. Trade Comm'n v. Tronox Ltd.*, Case No. 1:18-cv-01622 (TNM)(D.D.C. 2018)).

[31] Rotemberg and Saloner (1986) point out that an infinite punishment period with a low discount factor is equivalent to a finite punishment period with a high discount factor, an argument we formalize in Appendix A.2. In our application, with $\delta = 0.9$ and $\phi = 0$, about three months of punishment are sufficient to ensure incentive compatibility. That such a brief punishment period is required can be attributed to the results shown in Figure 3: the gap between price leadership and Nash-Bertrand per-period profit is much larger than the gap between deviation and price leadership per-period profit.

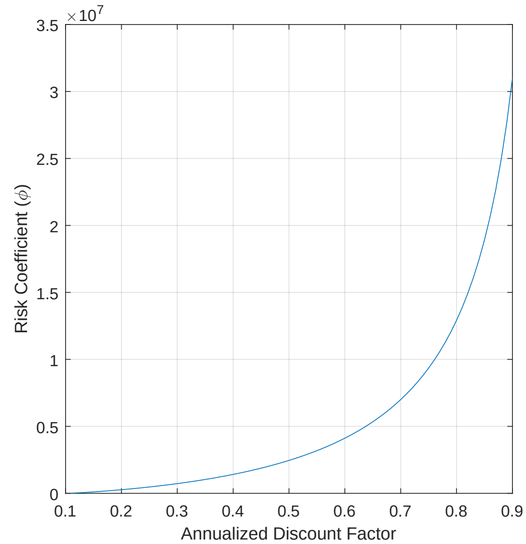

Figure 5: Joint Identification of Antitrust Risk and the Discount Factor

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

for any supermarkup less than 0.596. MillerCoors would prefer to deviate for any higher supermarkup. ABI, by contrast, still has slack in its IC constraint at $m = 0.596$. Thus we conclude that MillerCoors constrains coalition pricing in the observed equilibrium.[32]

6 Counterfactual Analysis of Mergers

6.1 The ABI and Modelo Merger

An June 28, 2012, ABI agreed to acquire Grupo Modelo for about $20 billion. The acquisition was reviewed by DOJ, which sued in January 2013 to enjoin the acquisition.[33] Prior to trial the merging firms and DOJ reached a settlement under which Modelo's entire U.S. business was divested to Constellation Brands, a major distributor of wine and liquor.[34] In

[32]Readers may wonder why a lower discount factor is associated with more slack for some supermarkups, as increasing the discount factor unambiguously loosens IC constraints in the model, *ceterus parabis*. Here not all else is held constant—a lower discount factor requires a greater risk coefficient to balance IC.

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

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

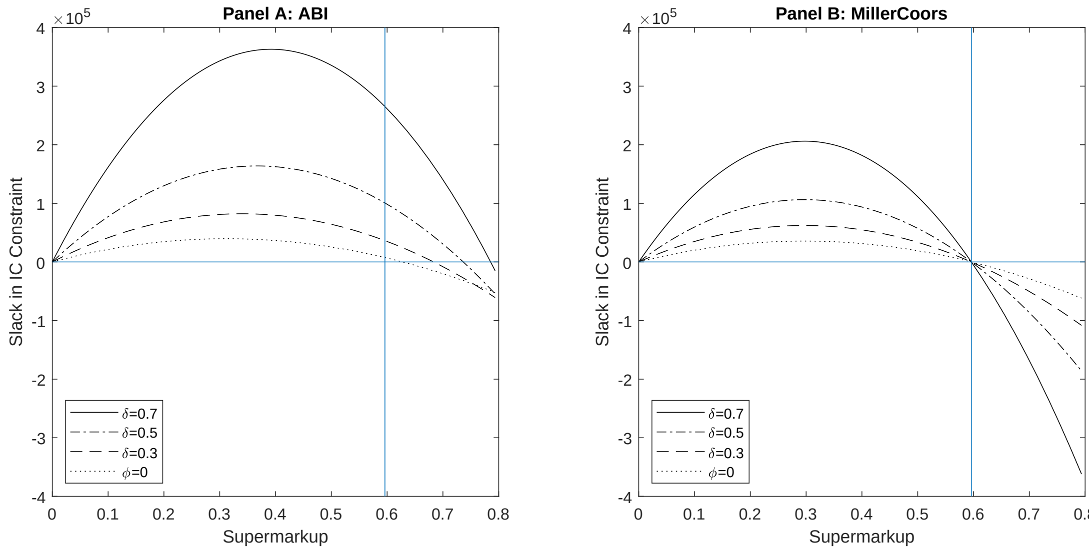

Figure 6: Slack Functions Given Observed Post-Merger Market Structure

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

its Complaint, the DOJ alleged that Modelo constrains the (coordinated) prices of ABI and MillerCoors:

> Defendant's combined national share actually understates the effect that eliminating Modelo would have on competition in the beer industry... because of the interdependent pricing dynamic that already exists between the largest brewers. As the two largest brewers, 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.[35]

In our first set of counterfactuals, we examine this theory of harm using the price leadership model. To implement, we assume that the Modelo products are priced by ABI, that is, we model the merger as it would have occurred without the divestiture. We do not incorporate any merger efficiencies in the result we present.[36] We consider a number of

last accessed February 13, 2019.

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

[36]Results are surprisingly similar if we assume that Modelo brands are produced in ABI breweries (reducing

candidate supermarkups, $m = 0.00, 0.01, 0.02, \ldots$, and for each we obtain the profit of each firm with price leadership, deviation, and punishment. We then calculate the slack in the IC constraints of each firm under the calibrated (δ, ϕ) combinations. We focus on the year 2011. As the final year of our data, it is the nearest to the acquisition date.

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

Table 6 provides greater detail on our estimates for the unilateral ("Δ Bertrand Price") and coordinated ("Δ Supermarkup") effects of the merger. As shown, the Nash-Bertrand prices of ABI and Modelo brands increase by \$0.29 and \$1.76 on average, with the magnitude of the latter change reflecting a strong incentive to steer customers toward higher-markup ABI brands. Prices also increase due to a higher supermarkup. For ABI and MillerCoors the magnitude of this change ranges from \$0.21 to \$1.01 across the calibrations selected for (δ, γ). For Modelo the change also reflects an adoption of the initial supermarkup of 0.596. The total changes in price ("Total Δ Price") equal the sum of these effects for the coalition firms. The changes should be interpreted cautiously as they move the equilibrium well beyond the range of the data. However, they do support that the merged firm would raise the prices of Modelo substantially and that coordinated effects account for much of the total change, consistent with the DOJ theory of harm. The average market share of Modelo brands decreases by more than 50% in all of the specifications we consider.[37]

shipping distances) and also benefit from a \$0.50 downward shift in marginal costs, in line what we estimate for MillerCoors. This robustness exists because it is the MillerCoors slack function that constrains coalition pricing. The marginal cost of ABI does not matter much, so long as IC does not bind.

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

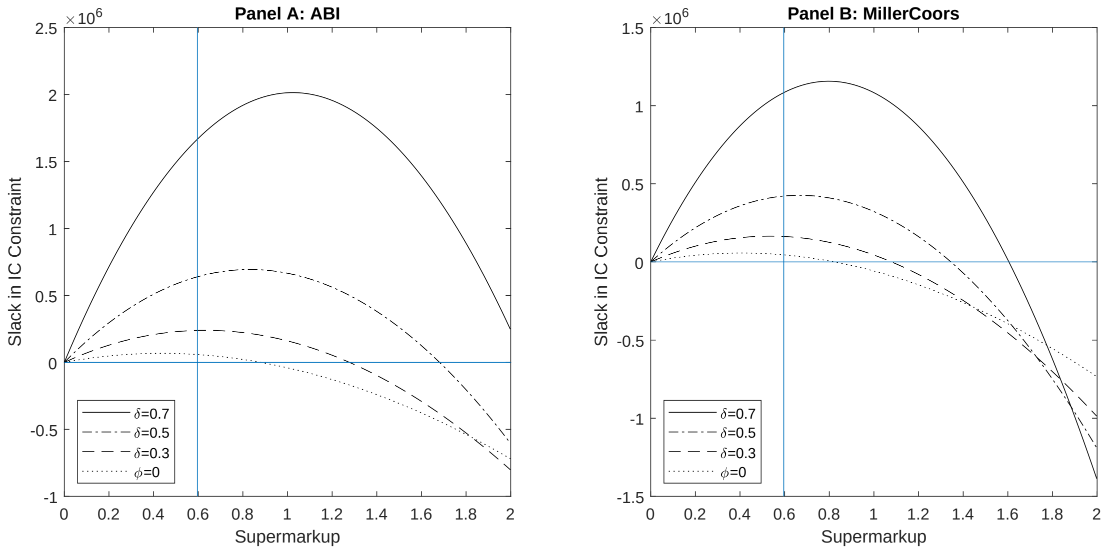

Figure 7: Slack Functions with an ABI/Modelo Merger

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

6.2 The MillerCoors Joint Venture

In the second set of counterfactual exercises, we explore the role of the MillerCoors joint venture in facilitating a price leadership equilibrium. To do so, we unwind the joint venture by assigning the Miller and Coors brands to separate firms and applying the pre-merger cost structure. We focus on the year 2011, which isolates the effects of the joint venture as other demand and cost factors are unchanged.[38] We consider a number of candidate supermarkups, $m = 0.00, 0.01, 0.02, \ldots$, and for each we obtain the profit of each firm with price leadership, deviation, and punishment. We then calculate the slack in the IC constraints of each firm under the calibrated (δ, ϕ) combinations.

Figure 8 plots the results for the calibrations that use $\delta = 0.7$ (Panel A), $\delta = 0.5$ (Panel B), $\delta = 0.3$ (Panel C), and $\gamma = 0$ (Panel D). Unwinding the joint venture has little effect on the slack functions of ABI, which remains positive for supermarkups less than 0.596. Effects are more pronounced for Miller and Coors. In Panels A and B, IC of both firms is violated for

[38] The marginal cost specification allows the merger to affect marginal costs by reducing shipping distances and via separate vertical shifts for Miller and Coors (e.g., see the discussion under equation (9)). To conduct the counterfactual, we recalculate distribution costs for the year 2011 using pre-merger brewery ownership and 2011 gasoline prices. We also eliminate the estimated vertical shifts in marginal cost.

Table 6: Effects of ABI/Modelo on Average Prices and Quantities

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

Notes: The table shows unweighted averages for the year 2011. Results are based on the RCNL-2 demand specification.

any positive supermarkup. In Panel C, the slack function of Miller is positive for $m \leq 0.05$ but Coors IC is violated for any positive supermarkup. Thus, in these first three panels, any ABI/Miller/Coors coalition is unsustainable. In Panel D, all IC constraints are satisfied for $m \leq 0.48$. The result is unsurprising as theory indicates some positive supermarkup can always be sustained in the absence of antitrust risk (Proposition 1).

While we cannot rule out the ($\gamma = 0, \phi = 0.11$) calibration on theoretical grounds, it appears to explain the prices changes in the raw data less well than the calibrations that employ higher discount factors. To demonstrate, we run difference-in-difference OLS regressions based on the specification

$$\begin{aligned} \log p_{jrt} &= \beta_1 \mathbb{1}\{\text{MillerCoors}\}_{jt} \times \mathbb{1}\{\text{Post-Merger}\}_t \\ &+ \beta_2 \mathbb{1}\{\text{ABI}\}_{jt} \times \mathbb{1}\{\text{Post-Merger}\}_t + \psi_{jr} + \tau_t + \epsilon_{jrt} \end{aligned}$$

where ψ_{jr} represents product $\times$ region fixed effects and τ_t represents period fixed effects.

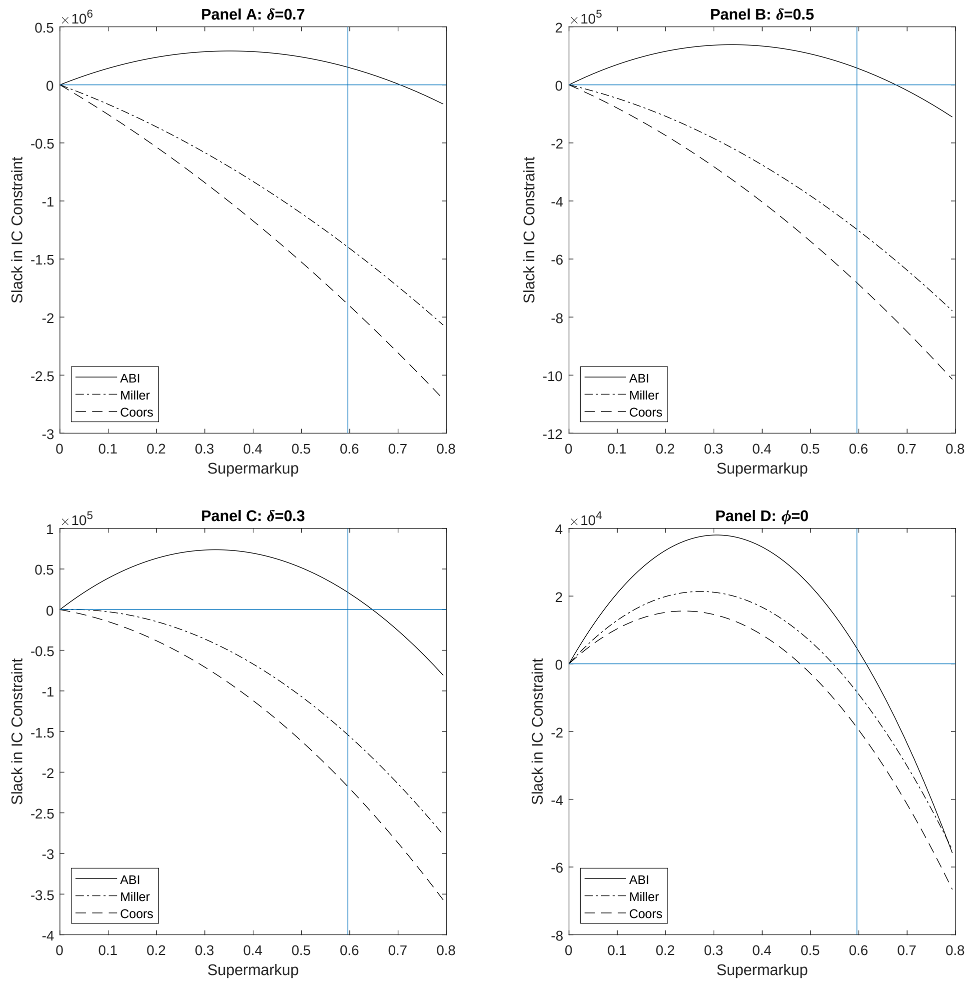

Figure 8: Slack Functions with an ABI/Miller/Coors Coalition

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

We construct the sample three ways to provide the comparison. Each sample uses the raw data for the post-merger prices. The first sample also uses raw data for the pre-merger prices. The second sample replaces these with simulated Nash-Bertrand prices computed for the "no merger" scenario. The third sample instead replaces with the "no merger" PLE obtained with $(\gamma = 0, \phi = 0.11)$ calibration.[39] As Modelo/Heineken provide the control

[39] In the second and third samples, we draw from the post-merger periods.

Table 7: Results from Differences-in-Differences Regression

	(i)	(ii)	(iii)
$\mathbb{1}\{\text{MillerCoors}\}_{jt} \times \mathbb{1}\{\text{Post-Merger}\}_t$	0.067	0.075	0.019
	(0.005)	(0.010)	(0.009)
$\mathbb{1}\{\text{ABI}\}_{jt} \times \mathbb{1}\{\text{Post-Merger}\}_t$	0.057	0.058	0.009
	(0.005)	(0.001)	(0.001)

Notes: The table provides the results of OLS regression. The dependent variable is log(price). The specification includes region×product and period fixed effects. The columns are produced from different samples. Column (i) uses the raw data from 2005-2011. Column (ii) uses data from 2009-2011 for post-merger prices and simulated Bertrand prices for pre-merger prices. Column (iii) uses data from 2009-2011 for post-merger prices and PLE prices from 2009-2011 for the pre-merger prices. The counterfactual used in the latter two columns are computed assuming that the Miller/Coors merger did not occur. All columns exclude months from June 2008 to May 2009. Standard errors are clustered at the region level and shown in parenthesis. Results are based on the RCNL-2 demand specification.

observations in the regression equation, the (β_1, β_2) coefficients provide the average change in ABI/Miller/Coors prices relative to those of Modelo/Heineken. We are interested in whether the change in the raw data (first sample) is closer to the change from a shift in equilibria from Nash-Bertrand to PLE (second sample) or by the change from shift from one PLE to another according to the $(\gamma = 0, \phi = 0.11)$ calibration (third sample).

Table 7 summarizes the regression results. As shown, the average price effect in the raw data is about 6.7%. A transition from Nash-Bertrand equilibrium to a PLE produces similar average price effects of about 7.5%. By contrast, the shift in PLE obtained from the $(\gamma = 0, \phi = 0.11)$ produces much smaller average price changes of about 1.9%. Thus, the regressions support a transition from Bertrand to PLE that is contemporaneous with the merger. Of course, we cannot rule out a transition from Bertrand to a PLE even with the $(\gamma = 0, \phi = 0.11)$ calibration, as Bertrand is an SPE of the repeated game. In that case, however, the model does not provide guidance on why equilibrium change occurs. The other calibrations do provide such an explanation.[40]

To complete the analysis, we also explore a coalition structure that leaves Coors in the fringe. The IC constraints of ABI and Miller shift down substantially (see Appendix Figure B.3). With a discount factor of $\delta = 0.3$, for example, both IC constraints are violated for

[40]The results shown in the table support an important assumption that we employ in estimation, namely that the pre-merger equilibrium is Bertrand. We do not think this consistency is pre-determined.

any positive supermarkup. Considered together, the results developed herein suggest the MillerCoors joint venture is pivotal for successful price leadership, under a reasonably broad set of discount factors. Thus, it may be reasonable to interpret the price changes shown in Section 2.3 as reflecting in part the coordinated effects of the Miller/Coors merger.

7 Conclusion

There is a longstanding concern that horizontal mergers can facilitate tacit collusion. Such "coordinated effects" are discussed extensively in the 2010 U.S. DOJ/Federal Trade Commission Merger Guidelines. Despite this concern, there is little empirical work that estimates models in which tacit collusion can arise. Addressing this gap is important to the advancement of antitrust policy and research on concentrated industries.

This paper makes progress by deriving and estimating a repeated oligopoly model of the U.S. beer industry. We motivate assumptions about firms' strategies along the equilibrium path with publicly-available court documents that character ABI as a price leader and MillerCoors as a follower. We estimate that these brewers charge about \$0.60, or 5% of the average price of a 12 pack of beer, in excess of the competitive price. Counterfactual simulations of ABIs attempted purchase of Grupo Modelo support the view that Grupo Modelo acted as a maverick, constraining interdependent pricing between the industry leaders ABI and MillerCoors. We also show that for many parameterizations the MillerCoors joint venture is pivotal in facilitating supracompetitive pricing.

As we are among the first to estimate a structural model of price coordination, we conclude with a brief discussion about certain limitations of our approach, with an eye toward assisting future research efforts. As is typical of games with perfect information, deviation and punishment do not occur along the equilibrium path in our model. Thus, so long as the model aligns with the data generating process, the researcher cannot make empirical inferences about the duration and severity of punishment. Yet often some inference is needed, as any analysis of incentive compatibility depends on punishment payoffs.

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

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

References

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

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

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

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

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

Asker, J., A. Collard-Wexler, and J. D. Loecker (2018). (mis)Allocation, market power, and global oil extraction. *American Economic Review f.*

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

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

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

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

Bernheim, B. D. and M. D. Whinston (1990). Multimarket contact and collusive beh. *RAN 21*(1), 1–26.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Igami, M. and T. Sugaya (2018). Measuring the Incentive to Collude: The Vitamin Cartels, 1990-1999. Available at SSRN: `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2889837`.

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

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

Kaplow, L. (2013). *Competition Policy and Price Fixing.* Princeton University Press.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Whinston, M. D. (2006). *Lectures in Antitrust Economics*. MIT Press.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data* (2nd ed.). MIT Press.

Appendix

A Theoretical Details

A.1 Proof of Proposition 1

The proof of the first claim is standard. It is immediately apparent from equation (6) that $\lim_{\delta \to 1} s_{ft}(m) = +\infty$ given the sign of expectation, because the term labeled "Expected Future Net Benefit of Price Leadership" converges to infinity as $\delta \to 1$, while the term labeled "Immediate Net Benefit of Deviation" is unaffected by the discount factor. For the second claim, we consider single-product firms for simplicity. We have

$$E_\Psi \left[\pi_f \left(p^{PL}(m^*) \right) - \pi_f \left(p^{NB} \right) \right] = c_f > 0$$

Defining $\varepsilon \equiv \frac{\delta}{1-\delta} c_f > 0$ and $t_f(m) \equiv \pi_f(p^{D,f}(m)) - \pi_f(p^{PL}(m))$ the slack function is $s_f(m) = \varepsilon - t_f(m)$. The proposition states that for any $\varepsilon > 0$ there exists some $m^*(\varepsilon) > 0$ such that $\varepsilon > t_f(m)$ if $m = m^*(\varepsilon)$. This is guaranteed if $t_f(m) \geq 0$, $t_f(0) = 0$ and $\left. \frac{\partial t_f(m)}{\partial m} \right|_{m=0} = 0$. These first two of these properties hold in the model by construction. The third property is easily verified. The derivative of deviation profit function is given by

$$\left. \frac{\partial \pi_f(p)}{\partial m} \right|_{p_f = p_f^{D,f}(m);\, p_{-f} = p_{-f}^{PL}(m)} = \left. \left(\frac{\partial \pi_f(p)}{\partial p_f} + \sum_{k \neq f} \frac{\partial \pi_f(p)}{\partial p_k} \right) \right|_{p_f = p_f^{D,f}(m);\, p_{-f} = p_{-f}^{PL}(m)} \tag{A.1}$$

The first term inside the parentheses equals zero if evaluated at $m = 0$ by the envelop theorem, and the remaining terms are positive. The derivative of the price leadership profit function is given by

$$\left. \frac{\partial \pi_f(p)}{\partial m} \right|_{p = p^{PL}(m)} = \left. \left(\frac{\partial \pi_f(p)}{\partial p_f} + \sum_{k \neq f} \frac{\partial \pi_f(p)}{\partial p_k} \right) \right|_{p = p^{PL}(m)} \tag{A.2}$$

By inspection, the right hand side of equations (A.1) and (A.2) are identical for $m = 0$ (see also Figure 3 for a graphical illustration). QED.

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

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

$$\frac{1}{1-\delta}\pi^{PL} \geq \pi^{D} + \frac{\delta}{1-\delta}\pi^{NB}, \tag{A.3}$$

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

$$\sum_{t=0}^{\infty}\eta^{t}\pi^{PL} \geq \pi^{D} + \sum_{t=1}^{n}\eta^{t}\pi^{NB} + \sum_{t=n+1}^{\infty}\eta^{t}\pi^{PL}, \tag{A.4}$$

with a discount factor of η. Rearranging equation (A.4) and applying rules for geometric series yields

$$\frac{1-\eta^{n+1}}{1-\eta}\pi^{PL} \geq \pi^{D} + \frac{\eta(1-\eta^{n})}{1-\eta}\pi^{NB}. \tag{A.5}$$

By inspection, equations (A.3) and (A.5) are equivalent if and only if

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

These conditions are satisfied for

$$\delta = \frac{\eta(1-\eta^{n})}{1-\eta^{n+1}}. \tag{A.6}$$

Punishment for n periods at a discount factor of η is equivalent to grim trigger punishment at a discount factor of δ, provided equation (A.6) holds. Further, by inspection, for a given η, decreasing n will decrease δ. Thus, a model with a low discount factor and lengthy punishment is equivalent to a model with a high discount factor and short punishment.

B Additional Figures and Tables

Table B.1: Product-Specific Elasticities for 12 Packs

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

Notes: This table provides the mean elasticities of demand for 12 packs based on the RCNL-1 specification (column (i) of Table 3). The cell in row i and column j is the percentage change in the quantity of product i with respect to the price of product j. Means are calculated across year–month–region combinations. The category cross-elasticities are the percentage change in the combined shares of products in the category due to a 1 percent change in the price of the product in question. Letting the category be defined by the set B, we calculate $\left(\sum_{j\in B, j\neq k} \frac{\partial s_j(p)}{\partial p_k}\right) \frac{p_k}{\sum_{j\in B, j\neq k} s_j(p)}$. The categories exclude the product in question. Thus, for example, the table shows that a 1 percent change in the price of a Bud Light 12 pack increases the sales of other 12 packs by 0.320 percent. Reproduced from Miller and Weinberg (2017).

Table B.2: Firm-Specific Elasticities

Panel A: Mean Elasticities in 2007

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

Panel B: Mean Elasticities in 2011

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

Notes: This table provides the mean firm-specific elasticities of demand in 2007 and 2011 based on the RCNL-1 specification (column (i) of Table 3). The cell in row i and column j is the percentage change in the quantity of firm i with respect to the prices of firm j. The elasticity of demand for products in set A with respect to prices of products in set B is defined as: $\left(\sum_{n\in A}\sum_{k\in B}\frac{\partial q_n}{\partial p_k}\right)\frac{\overline{p}_B}{\sum_{n\in A} q_n}$. Means are calculated across month–region combinations.

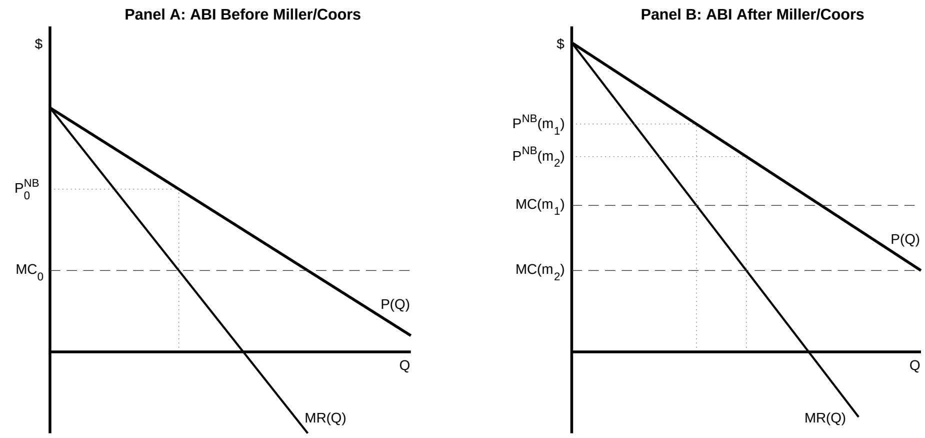

Figure B.1: Illustration of the Identification Strategy

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

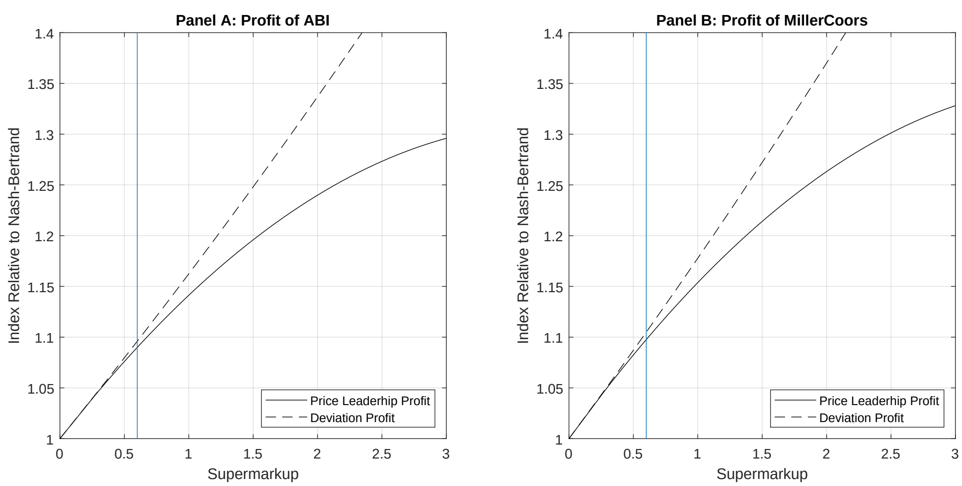

Figure B.2: Profit with Price Leadership and Deviation under a Linear Demand System

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

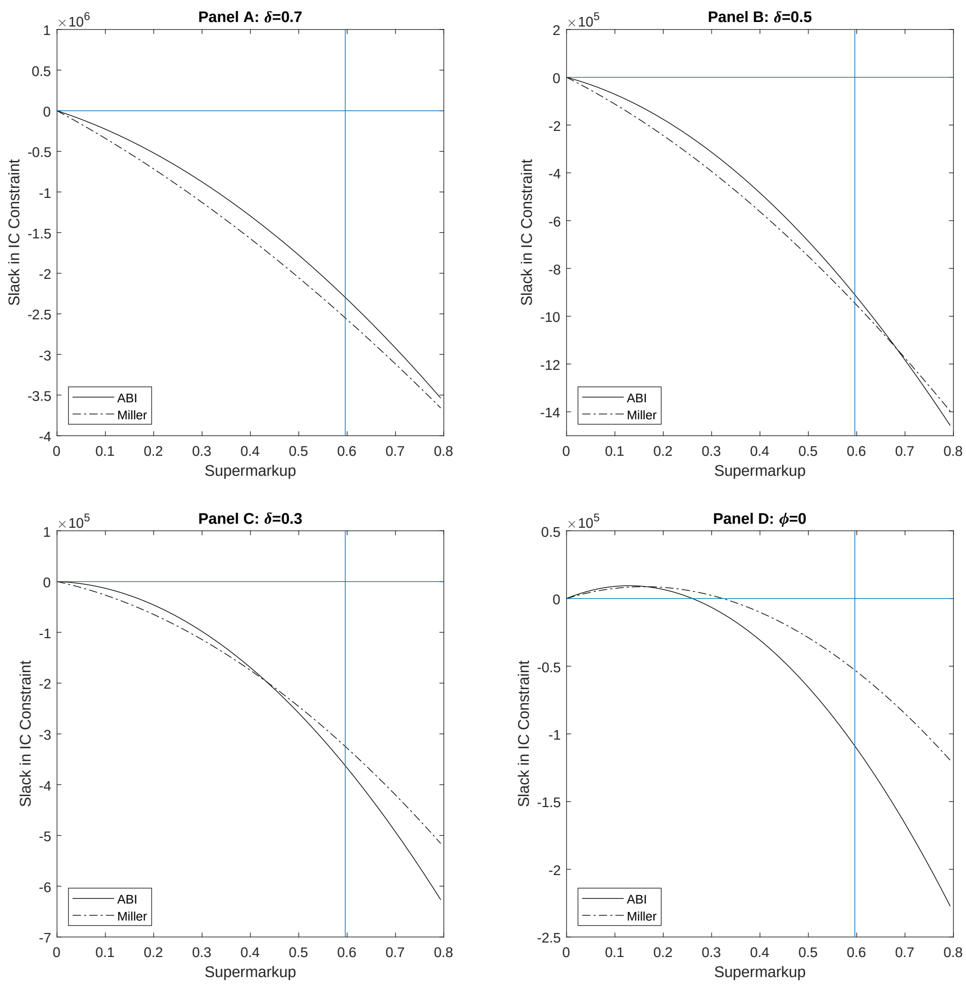

Figure B.3: Slack Functions with an ABI/Miller Coalition

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