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

Nathan H. Miller†
Georgetown University

Gloria Sheu‡
US Department of Justice

Matthew C. Weinberg§
The Ohio State University

February 15, 2019

Abstract

We examine an infinitely-repeated game of oligopoly price leadership in which each period one firm, the market leader, proposes a supermarkup over Nash-Bertrand prices. The supermarkup is chosen to maximize the leader's profit subject to all firms' incentive compatibility constraints (ICCs). We provide conditions under which the equilibrium supermarkup can be recovered from aggregate data on price and quantities. 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 2008 Miller/Coors merger. Counterfactual simulations demonstrate an ICC binds, as profit is greater with even higher supermarkups. We use the implied equality constraint to jointly identify a discount factor and the antitrust risk, the remaining structural parameters. We then explore the coordinated effects of ABI's acquisition of Grupo Modelo. Without divestitures, the merger would have relaxed ICCs, resulting in substantially higher prices. Finally, we return to the Miller/Coors merger. For many parameter values, no supermarkup satisfies ICCs without the merger. Thus, the merger may have be pivotal in generating the observed price leadership behavior.

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

*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 US Department of Justice.

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

‡US 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 industries ranging from beer, pharmacies, and gasoline (Miller and Weinberg (2017), Lemus and Luco (2018), Byrne and de Roos (2019)). Anecdotal examples are discussed in Scherer (1980) and go back to at least Bain (1960), who described the practice as *oligopolistic price leadership*.[1]

One way to understand price leadership is as the outcome of a repeated pricing game. Repeated oligopoly is the canonical framework for understanding how firms may set prices in excess of competitive levels, but there are typically many equilibria in these models. This raises the question of which equilibrium firms will select. We argue that price leadership provides a natural way to solve this problem–it provides a focal point for price coordination. Furthermore, it does so in a manner that avoids the explicit agreements frequently targeted by antitrust authorities.

This paper presents an empirical model of oligopolistic price leadership that we estimate with data on the U.S. beer industry. We use the model to discern the price effects of mergers. The model features an infinitely repeated pricing game in which trigger strategies may elevate prices above the static Nash-Bertrand equilibrium (hereafter, "Bertrand" equilibrium). Each period, the leader announces a "supermarkup" above Bertrand prices. On the equilibrium path, a set of coalition firms, comprised of the leader and its followers, accept the supermarkup. The leader chooses the supermarkup to maximize its profit, subject to the incentive compatibility constraints (ICCs) of coalition firms and taking as given the competitive response of fringe firms. All firms have perfect information. Thus, price leadership plays the role of equilibrium selection. The leader's announcement resolves the coordination problem that arises due to the infinitely many equilibrium outcomes available in unrestricted repeated pricing games (Whinston (2006)).

We use the model to understand the economic consequences of consolidation in the U.S. brewing industry, a market for which the assumption of static Nash competition may be inappropriate (Miller and Weinberg (2017)). Adopting standard antitrust parlance, we

[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.

model the *coordinated effects* of mergers as arising due a shift in equilibrium, either from a Bertrand equilibrium to a supracompetitive equilibrium, or from one supracompetitive equilibrium to another. Supporting this approach is documentary evidence in the public record that suggests the two largest brewers, ABI and MillerCoors, engage in leader/follower pricing. We analyze the ABI acquisition of Grupo Modelo in 2013 and the 2008 joint venture of Miller and Coors, reviewed by DOJ as a merger (forming MillerCoors). For the first we examine the allegation that Modelo constrained coordinated pricing, and for the second we evaluate whether the merger was pivotal for coordinated pricing. Our research is among the first attempts to formally model coordinated effects in real-world markets.[2]

The model of competition is estimated in two separate steps. We first show that within each market there is a distinct, unique vector of unobserved marginal costs implied by each candidate supermarkup chosen by the price leader. The assumption that ABI's marginal costs do not systematically increase relative to fringe firms' marginal costs provides a moment condition used to estimate the supermarkup. The identification result and the estimate do not depend on any particular assumption on strategies off the equilibrium path. We estimate that the joint venture increased prices by \$0.59 above Nash prices. Next, we use the estimated demand system, implied marginal costs, and counterfactual simulations to test whether binding incentive compatibility constraints limited the magnitude of the equilibrium supermarkup. We find that ABI would have chosen a supermarkup substantially higher than our estimate if there were no ICCs.

The second part of our estimation procedure exploits the inferred equality between the present value of coordination and the present value of deviating at the estimated supermarkup. Under the assumption that firms play grim trigger strategies with Nash reversion, we show that for many discount factors the value of price leadership strictly exceeds the value of deviating and then triggering competitive pricing. This implies the existence of a cost of coordinating. Because our estimates imply that the gap between the value of price leadership and the value of deviating grows larger as the supermarkup increases from its point estimate, the coordination cost must be increasing in the supermarkup. We calibrate different implied costs of coordination that are linearly increasing in the supermarkup for several values of the discount factor.

We then use the model to study ABI's acquisition of Grupo Modelo. The Department of Justice approved the merger only after Grupo Modelo's brands were divested to a third party

[2]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).

that at the time produced no beer, in principle preserving the pre-merger market structure. The DOJ's complaint describes Grupo Modelo as disrupting the market by not matching annual price increases initiated by ABI and typically followed by MillerCoors. In terms of the FTC/DOJ merger guidelines, the complaint portrays Grupo Modelo as a "maverick", described as "a firm that has often resisted otherwise prevailing industry norms to cooperate on price setting or other terms of competition." Mavericks are difficult to understand in the canonical models used to analyze mergers and prices based on static competition, but they are naturally modeled as firms outside of the set of leader-follower firms in our framework. We show that had the Grupo Modelo brands been acquired by ABI, coalition brand prices would have increased by over 10 percent of the average price of a 12 pack.

We next use the model to study the role of the MillerCoors joint venture in facilitating coordination. We do this by assigning Miller and Coors brands to separate firms with the pre-merger cost structure. We then examine whether the present value of price leadership profits exceeds the value of deviating at a range of supermarkups. We show that for any supermarkup and many combinations of the discount factor and antitrust risk, Coors would prefer to deviate rather than match ABI's proposed price increase, and also that no coalition involving only Miller and ABI was feasible in the absence of the joint venture.

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).[3] As their model incorporates asymmetric information, 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.[4] Finally, a number of articles consider partial coalitions (e.g., d'Aspremont et al. (1983), Donsimoni et al. (1986), and Bos and Harrington (2010)), albeit in homogeneous product settings.

From a methodological standpoint, our research is most similar to that of Igami and

[3]In an 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)).

[4]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 of price ceilings as focal points in the credit card industry.

Sugaya (2018) on the vitamins cartels of the 1990s. Both papers estimate the structural parameters underlying a supergame in which trigger strategies sustain supracompetitive prices, relying on counterfactual simulations to recover the profit terms that enter ICCs. There are also important points of departure. Igami and Sugaya assume that all firms either engage in maximal collusion (if ICCs are satisfied) or revert to Nash-Courot equilibrium (otherwise). Thus, some interesting aspects of our model, such as partial coalitions and the leader's ability to adjust the supermarkup to satisfy ICCs, are not present in their setup.

Also related is the contemporaneous empirical work of Eizenberg and Shilian (2019) and Fan and Sullivan (2018). The former paper tests whether competitive Nash-Bertrand pricing can explain price and quantity data across 40 categories of food. In markets where competition is not rejected, marginal costs are recovered from static pricing first-order conditions and critical discount factors necessary to sustain joint monopoly pricing under grim trigger strategies are computed. Mergers are not considered, and furthermore it is not clear whether they would increase incentives to tacitly collude in this framework. This is because in the absence of efficiencies, mergers do not change the benefit of collusion, but they increase incentives to deviate by raising competitive Nash-Bertrand profits and immediate deviation payoffs (see Davis and Huse (2010)). Fan and Sullivan (2018) show that marginal costs can be recovered from aggregate price and quantity data for coalition pricing games that satisfy Pareto efficiency for participating firms. Price leadership obtains Pareto efficiency only under strong symmetry conditions so the models are mostly applicable in different settings.

The paper proceeds as follows. Section 2 describes the industry background and qualitative evidence that the price leadership model is a good fit for the industry. It also briefly describes the model of consumer demand estimated in Miller and Weinberg (2017). Section 3 defines the elements of the repeated game. Section 4 discusses identification and estimation of the model of competition, and the results appear in Section 5. Section 6 presents counterfactuals that help understand the price effects of two recent mergers, the proposed acquisition of Grupo Modelo brands by ABI and the MillerCoors joint venture. Section 7 concludes.

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

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.

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.[5] While retail price maintenance is technically illegal in many states, in practice, distributors are often induced to sell at wholesale prices set by brewers (Asker (2016)).

Table 1 summarizes the revenues 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.[6] 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 saving 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 ac-

[5]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).

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

quired 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 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.[7]

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."[8]

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 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

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

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

sales (McClain (2012)).[9] 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 shares based on 144-ounce equivalent units, the size of a 12-pack, and measure price as the ratio of revenue to equivalent unit sales.[10]

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 is 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 show 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 marking 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 cost or demand factors. Miller and Weinberg (2017) provide econometric evidence that visual break apparent in the figure is not due artificial sample selection, trends, or omitted variables. Further, qualitative evidence culled from regulatory filings are consistent with

[9]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.

[10]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 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.

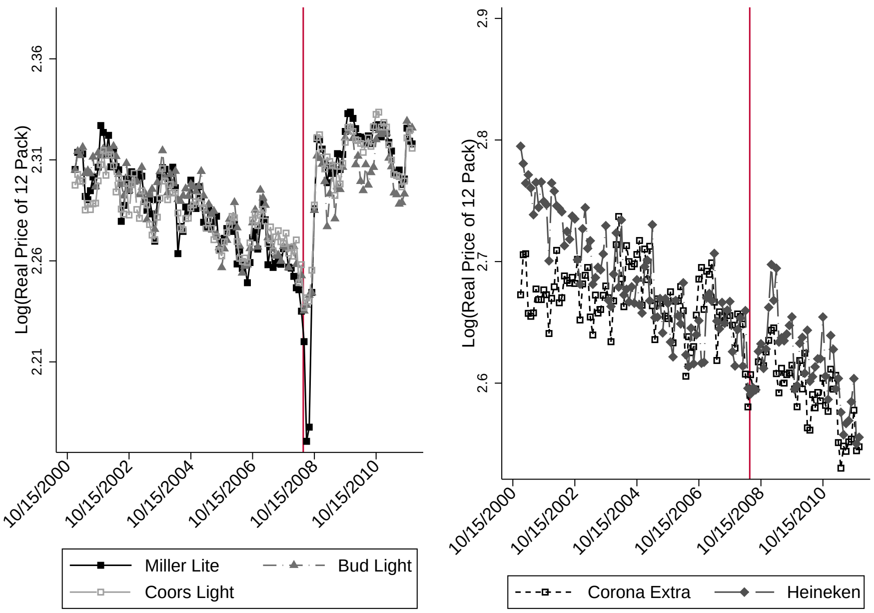

Figure 1: Average Retail Prices of Flagship Brand 12-Packs

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).

softened competition among the domestic brewers in the post-merger period.[11]

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 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.

[11]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 (pp. 29), and 2011 (p. 28) and Anheuser-Busch's Annual Report in 2005 (p. 5).

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 $\mu_{ijrt} = [p_{jrt}, x_j]' * \Pi D_i$.[12] 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).

[12] 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 $I_{irt} = \log(1 + \exp I_{i1rt})$.

Table 2 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) accomplish a similar effect by allowing 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 due to magnitude of the nesting parameter, and indicate that most substitution occurs among the inside goods, rather than between the inside goods and the outside good.[13][14]

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:

(i) The leader announces a non-binding supermarkup, $m_t \geq 0$, above 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.

[13] 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.

[14] **NOTE** add firm-level elasticities

Table 2: 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).

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.[15] 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.

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

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

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 Bertrand prices, $p^B_{ft}(\Psi_t)$, obtain. We assume a unique solution exists, which can be verified in our setting with nested logit demand (Mizuno (2003)) but need not hold with the RCNL demand model. Coalition firms 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 6.1.

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.[16] 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.

[16] 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 $\sigma_1 : \mathbb{H} \rightarrow \mathcal{M} \times \mathcal{R}^{J_1}$, where $\mathbb{H}$ is the set of histories, $\mathcal{M}$ is the set of possible supermarkups, and J_1 is the number of products controlled by the leader. The strategies of firms $f = 2, \ldots, F$ are $\sigma_f : \mathcal{M} \times \mathbb{H} \rightarrow \mathcal{R}^{J_f}$. We obtain the strategies that constitute the PLE, starting with the pricing stages, continuing with the announcement stages, and then finishing with the coalition selection at ($t = 0$). We then discuss the equilibrium and describe some of its characteristics.

Consider the pricing stage in some arbitrary period t. Each coalition firm $f \in \mathbb{C}$ "accepts" the leader's proposed supermarkup m_t if it prices according to $p_{ft}^{PL}(m_t; \Psi_t) = p_{ft}^{B}(\Psi_t) + m_t$. Fringe firms accept simply by pricing on their best reactions functions. Thus, let $p_{ft}^{PL}(m_t; \Psi_t)$ 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 $p_t^{D,f}(m_t; \Psi_t)$ collect the prices that arise if firm f solves equation (5) with the (correct) anticipation that other firms accept. Second, let the vector $p_t^B(\Psi_t)$ collect the Bertrand prices that solve equation (5) for all firms. We refer to $p_t^{D,f}(\cdot)$ and $p_t^B(\cdot)$ as deviation and punishment prices, respectively.

One additional definition is necessary. 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

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

where $\delta \in (0, 1)$ denotes a common discount factor and m^* is a scalar that we define below.[17] In the PLE, the inequalities $s_{ft}(m_t; \Psi_t) \geq 0$ play the role of the incentive compatibility

[17] The slack function of fringe firms does not include antitrust risk $R(m^*(\Psi))$ and is therefore guaranteed to be at least weakly positive, as there is no benefit from deviation by construction.

constraints (ICCs). 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 $p_t^B(\Psi)$. 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 in any ICC, then this is known by all firms. Deviation prices are never realized in the pricing subgame as play shifts immediately to Bertrand prices.

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

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

As the slack function equal 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 the ICCs of coalition firms, even if this implies 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 period. In numerical experiments, we have confirmed that partial coalitions can be optimal for the leader. Typically this occurs if there is substantial heterogeneity in the slack functions, which can allow for higher supermarkups with a partial coalition as ICCs 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

[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.

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 supermarkup $m^* > 0$ satisfies*

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

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 Bertrand level satisfies ICCs 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 ICCs. Thus, antitrust risk is important insofar as it creates the theoretical possibility that some market structures cannot support positive supermarkups.

The PLE is not generally Pareto optimal for the coalition firms because the leader acts in its own interest and side payments are not permitted.[19] Otherwise, the model closely resembles the canonical Rotemberg and Saloner (1986) model of collusion. Figure 2 provides a numerical example involving two coalition firms and logit demand.[20] The leader selects among supermarkups that raise both firms' prices by the same amount above Bertrand equilibrium. The supermarkup that maximizes the leader's profit (the "Unconstrained PLE") violates ICCs, so the PLE features a smaller supermarkup. The figure also may help readers visualize how PLE outcomes could be reached in more standard repeated pricing games with particular assumptions on equilibrium selection.

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

[20] 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 parameterization $\beta_1 = \beta_2 = 3$, $\beta_3 = 1$, $\alpha = 1.5$. Marginal costs are $mc_1 = mc_2 = 0$ and $mc_3 = 2$, the discount factor is $\delta = 0.4$, and antitrust risk is $R(m) = 0$.

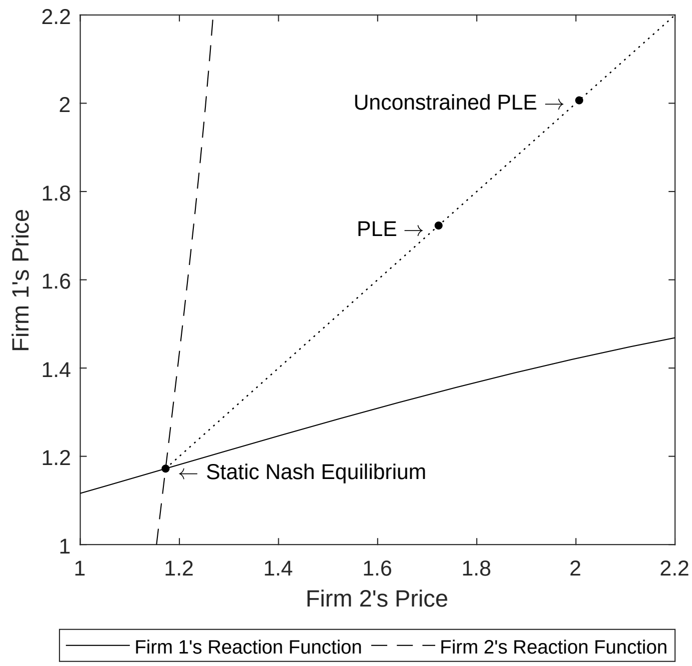

Figure 2: Illustration of the Price Leadership Equilibrium

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 the ICCs. Nonetheless, with the supermarkups in hand, one can test whether an ICCs 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 6.1.

4.1 Identification of Marginal Costs

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

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

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^B). 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 $p_k^B = p_k - m$ for each coalition firm $k \in \mathbb{C}$.

3. Compute p_j^B 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., $p_k = p_k^B$ 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^B 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:

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

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.[21]

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

[21] 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.

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).[22]

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 super-markets absent additional restrictions on the model.

We make progress by assuming Bertrand pricing ($m_t = 0$) in periods predating the Miller/Coors merger, which resolves the otherwise intractable under-identification problem.[23] 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

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

[23]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.

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.[24] 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

Table 3 contains supply side estimates. In addition to the trucking distance between each retail location and the nearest brewery interacted with diesel prices, the marginal cost functions incorporate time fixed effects, region fixed effects, and brand/package size fixed effects. Marginal costs change with the joint venture in two ways: through an intercept shift that is allowed to differ for Miller and Coors products and through a reduction in shipping distances. The different columns correspond to the four different specifications for the random coefficient nested logit model of demand. We use the specification in column two in each of the following counterfactuals, though each estimate is fairly stable across demand specifications.

The supermarkup is \$0.60 above Nash prices, or about 6% of the total average price of a 12 pack of beer, and is precisely estimated. The marginal cost intercept of Miller and Coors brands dropped after the joint venture by \$0.53 and \$0.83 per 12 pack, respectively. Marginal costs increase in shipping distances, implying a second source of efficiencies from the MillerCoors joint venture as production of Coors and, to a lesser extent Miller brands are moved to breweries closer to each retail location.

We explore whether an ICC binds by computing the unconstrained supermarkup ABI would choose to maximize its profits, given estimated demand and inferred marginal costs. Each period the fringe maximizes profits conditional on the coalition implementing the supermarkup as before. Column 2 shows that the unconstrained supermarkup is \$2.57 above Nash prices, over four times higher than the estimated supermarkup. The binding constraint that limits ABI's choice of the supermarkup is MillerCoors ICC, an issue we explore further in section 6.[25]

Figure 3 and appendix Figure B.1 explore MillerCoors and ABI profits, prices, and

[24] 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.

[25] In the next draft we will implement a formal test of equality between the constrained and unconstrained supermarkup, in addition to the supermarkup for demand specifications in columns 3 and 4.

Table 3: Baseline Supply Estimates

	Parameter	RCNL-1	RCNL-2	RCNL-3	RCNL-4
Estimation Results					
Supermarkup	m	0.643	0.596	0.738	0.709
		(0.025)	(0.027)	(0.034)	(0.033)
Miller×Post Merger	γ_1	-0.540	-0.533	-0.583	-0.416
		(0.007)	(0.007)	(0.005)	(0.002)
Coors×Post Merger	γ_2	-0.826	-0.831	-0.914	-0.666
		(0.009)	(0.009)	(0.006)	(0.004)
Distance	γ_3	0.168	0.164	0.172	0.153
		(0.001)	(0.001)	(0.001)	(0.001)
Supplementary Results					
Unconstrained supermarkup		2.69	2.57	3.25	2.56

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. Standard errors clustered by region and shown in parentheses. Bootstrapped 95% confidence intervals, shown in brackets, are provided for the unconstrained supermarkups.

shares at different supermarkups. The left panel of Figure 3 shows FY2011 MillerCoors price leadership and deviation profits, relative to Nash Bertrand, over a range of supermarkups. The price leadership profit is an inverted-U in the supermarkup, with a maximum just over $2.50. In contrast, deviation profits increase monotonically in the supermarkup because higher supermarkups correspond to higher ABI prices, and the static best-response profits of MillerCoors increase in ABI prices. At the estimated supermarkup, MillerCoors profits are about 7 percent higher than Nash-Bertrand. Deviation profits are only about 8 percent higher, which has important implications for ICC constraints governing the viability of coordination.

The right panel shows that at the estimated supermarkup MillerCoors prices are about 8 percent higher than Nash Bertrand, and shares are about 8 percent lower. The figure also demonstrates that prices are strategic compliments in our model, as higher supermarkups correspond to higher deviation prices. The market share that MillerCoors captures with deviation also increases in the supermarkup. Together, this shows that the incentive to deviate increases with the level of the supermarkup. Figure B.1, which contains the same information for ABI, is essentially the same.

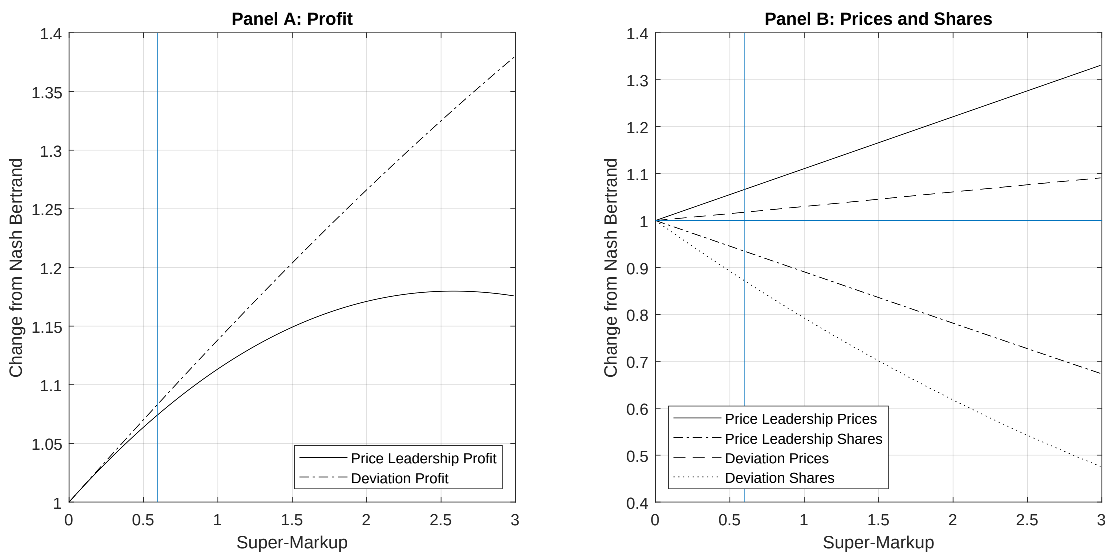

Figure 3: Profit, Prices and Shares of MillerCoors with Price Leadership and Deviation

Notes: The figure provides the profit (Panel A) and average prices and market share (Panel B) for MillerCoors 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.

6 Counterfactual Analysis of Mergers

We now examine how consolidation affects market outcomes given the possibility of price leadership. The slack functions defined in Section 3 play a critical role in determining the extent to which positive super-markets can be sustained in equilibrium. We therefore begin by explaining how we calibrate the discount factor and the antitrust risk. We then examine an ABI/Modelo merger that was proposed in 2012 and the consummated Miller/Coors merger.

6.1 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, corresponding to single ICC for each coalition firm.[26]

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)). We apply a simple parameterization, $R(m_t;\phi) = \phi m_t$, and 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 ICCs.[27] Therefore we assume that least one ICC binds. The parameters (δ, ϕ) are jointly identified.

Before proceeding, we recall that 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 the plaintiffs.[28] Further, historical evidence of pricing coordination sometimes is cited by antitrust authorities as contributing to a decision to challenge a merger.[29] Our parameterization of the antitrust risk is meant as a simple reduced-form way to capture these influences.

Figure 4 plots the values that balance the MillerCoors ICCs in 2011:Q4. As shown, with $\phi = 0$, an annualized discount factor of 0.11 balances the ICC, 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 ICCs 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.[30]

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

[27] The econometric tests, will be provided in the next version of the draft.

[28] Examples involve firms involved in text messaging (*Re: Text Messaging Antitrust Litig.*, 782 F.3d 867 (7th Cir 2015)), titanium dioxide (**CITE**), and chocolate (**CITE**).

[29] Interestingly, a prime example is ABI's attempted acquisition of Modelo in 2012-2013, which the DOJ challenged in part due to a concern that it would eliminate a constraint on coordinated ABI and MillerCoors price increases. We return to the economic effects of the proposed ABI/Modelo merger in Section 6.2.

[30] 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. 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 of Figures 3 and B.1: gap between price leadership and Bertrand per-period profit is much larger than the gap between deviation and price

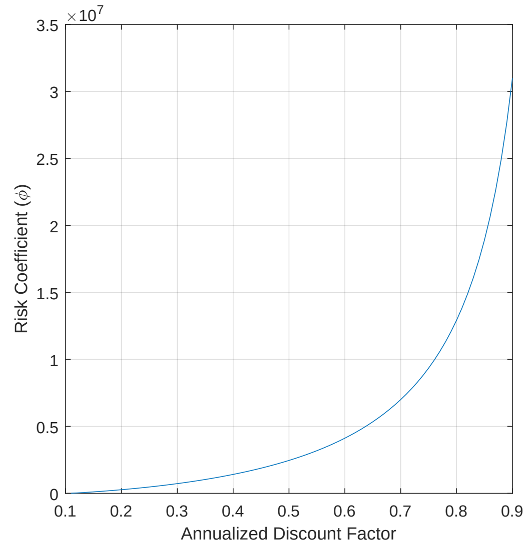

Figure 4: Joint Identification of Antitrust Risk and the Discount Factor

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

Figure 5 plots the slack in the ICCs of MillerCoors (Panel A) and ABI (Panel B) over the range of supermarkups $m \in [0, 0.8]$. Four alternative assumptions are used to calibrate the ICCs: $\delta = 0.7$, $\delta = 0.5$, $\delta = 0.3$, and $\gamma = 0$. In each case, we select the free parameter such that the ICC of MillerCoors binds at the estimated supermarkup of 0.596 (Figure 4 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 ICCs for any supermarkup less than 0.596. MillerCoors would prefer to deviate for any higher supermarkup. ABI, by contrast, still has slack in its ICC at $m = 0.596$. Thus we conclude that MillerCoors constrains coalition pricing in the observed equilibrium.[31]

leadership per-period profit.

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

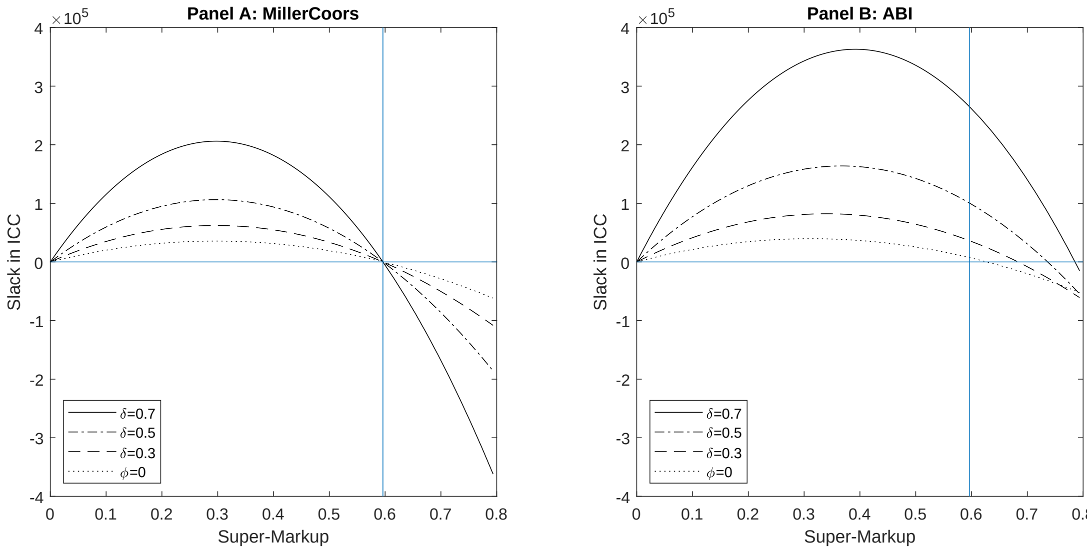

Figure 5: Slack Functions Given Observed Post-Merger Market Structure

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

6.2 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.[32] 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.[33] In 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.

[32] 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.*

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

> As the two largest brewers, ABI and MillerCoors often find it 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.[34]

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.[35] 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 ICCs 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 6 graphs the counterfactual slack functions of MillerCoors (Panel A) and ABI (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 ICCs 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 $p_t = p_t^B + m$ for coalition firms, the total change in price also reflects a shift in the Bertrand equilibrium, as ABI and Modelo internalize their pricing externalities. Applying standard antitrust terminology, we refer to the shift in Bertrand prices as the *unilateral effect* of the merger.

Table 4 provides greater detail on our estimates for the unilateral ("Δ Bertrand Price") and coordinated ("Δ Supermarkup") effects of the merger. As shown, the 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

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

[35] Results are surprisingly similar if we assume that Modelo brands are produced in ABI breweries (reducing 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 the ICC does not bind.

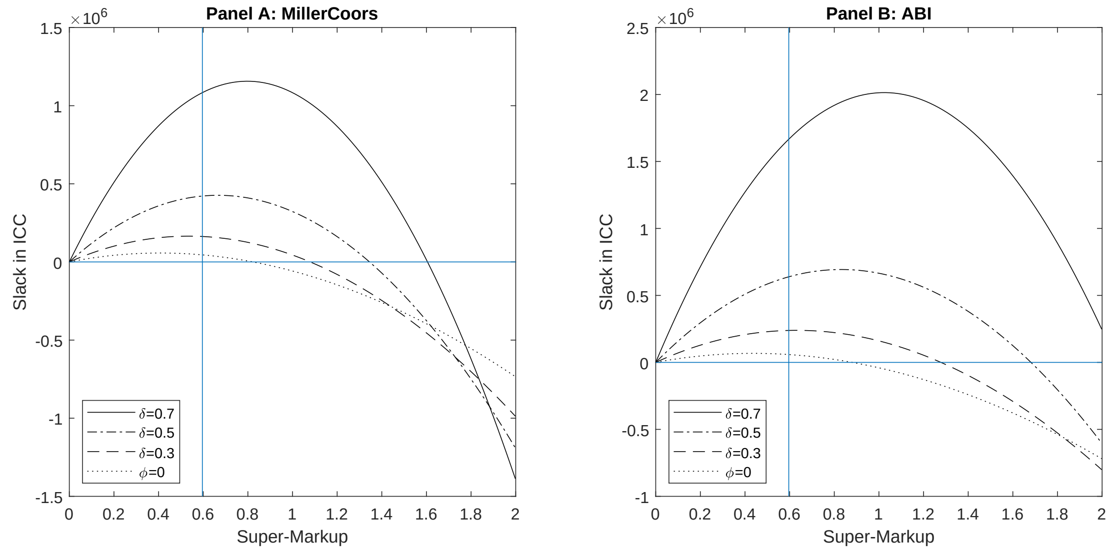

Figure 6: Slack Functions with an ABI/Modelo Merger

Notes: The figure provides the slack functions in 2011:Q4 ICC for MillerCoors (Panel A) and ABI (Panel B) and with supermarkups $m \in [0, 0.8]$. The ICCs are 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.

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.[36]

6.3 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

[36]The results for Heineken are interesting. Its Bertrand prices increase by \$0.01, reflecting a small degree of strategic complementarity in prices. However, it responds to the (large) supermarkups in the post-merger 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.

Table 4: 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.

as other demand and cost factors are unchanged.[37] 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 ICCs of each firm under the calibrated (δ, ϕ) combinations.

Figure 7 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, the ICCs of both firms are violated for any positive supermarkup. In Panel C, the slack function of Miller is positive for $m \leq 0.05$ but the Coors ICC is violated for any positive supermarkup. Thus,

[37]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.

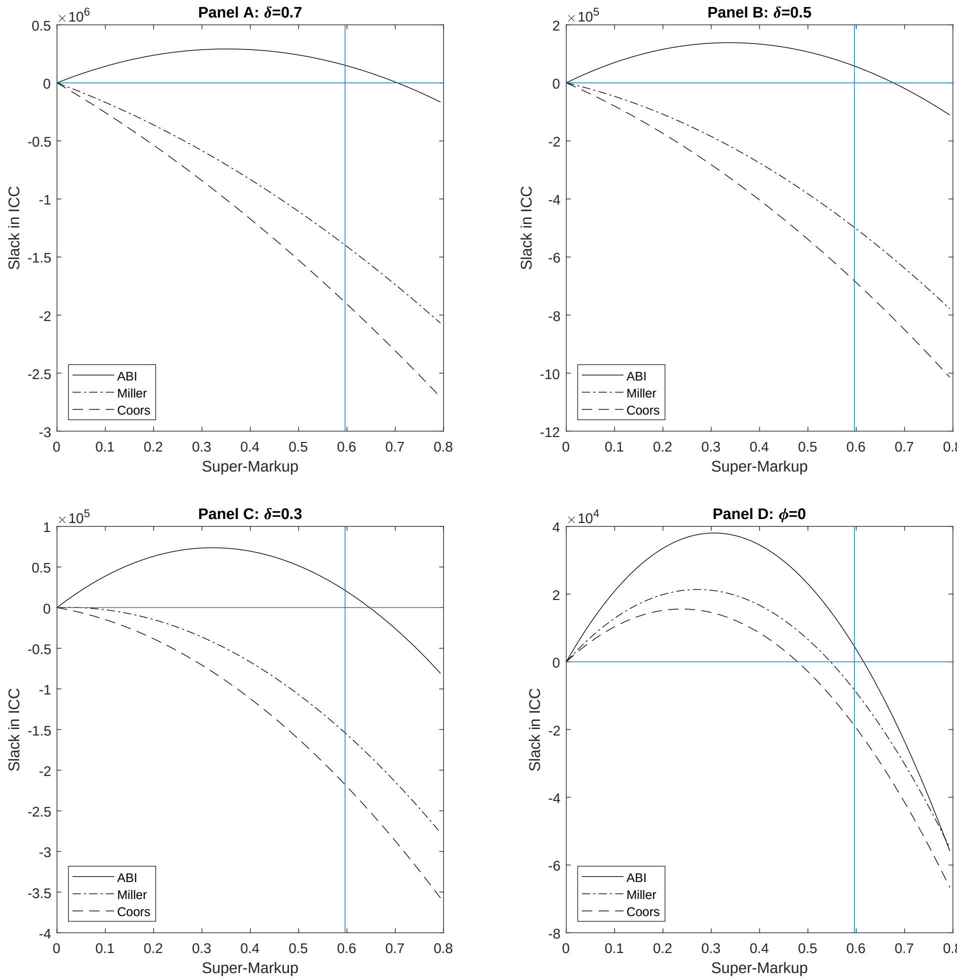

Figure 7: Slack Functions with an ABI/Miller/Coors Coalition

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. The ICCs are 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.

in these first three panels, any ABI/Miller/Coors coalition is unsustainable. In Panel D, all ICCs 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. 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 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.[38] As Modelo/Heineken provide the control 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 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 5 summarizes the regression results. As shown, the average price effect in the raw data is about 6.7%. A transition from 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.[39]

To complete the analysis, we also explore a coalition structure which leaves Coors in the fringe. The ICCs of ABI and Miller shift down substantially (see Appendix Figure B.2). With a discount factor of $\delta = 0.3$, for example, both ICCs are violated for 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.

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

[39]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.

Table 5: 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.

7 Conclusion

There is a longstanding concern that horizontal mergers could facilitate tacit collusion. Despite this concern, there is little empirical work that estimates models where tacit collusion can arise. This paper makes progress on this issue by estimating a repeated oligopoly model of the U.S. beer industry. We justify assumptions on firms strategies on the equilibrium path with antitrust documents describing ABI as a price leader. We estimate that the major 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 ABI's 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 parameter values the MillerCoors joint venture was pivotal in facilitating supracompetitive pricing.

The results are subject to several caveats. Estimating a repeated game requires assumptions on firms' strategies, not only on the equilibrium path but off it as well. We view industry documents as providing some support for our assumptions on the equilibrium path and make the simple assumption that firms would revert to static Nash Bertrand pricing after any deviation, but other assumptions could be explored. We also require an assumption on firms' beliefs about future demand and cost conditions, which we take as known with perfect foresight. Another assumption is that each firm evaluates streams of profits with a common discount factor that is invariant to changes in ownership, but as each firm in our sample is publicly traded it may be possible to use asset pricing models to infer the component of firm-specific discount factors related to time preferences.

References

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.*

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.

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*.

Fan, Y. and C. Sullivan (2018). Estimating markups with a flexible model of supply. Presentation at the HEC Montreal-RIIB Conference on Industrial Organization, July 11-12.

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.

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. 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`.

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.

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.

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. 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.

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.

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.

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 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

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

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.

B Additional figures and tables

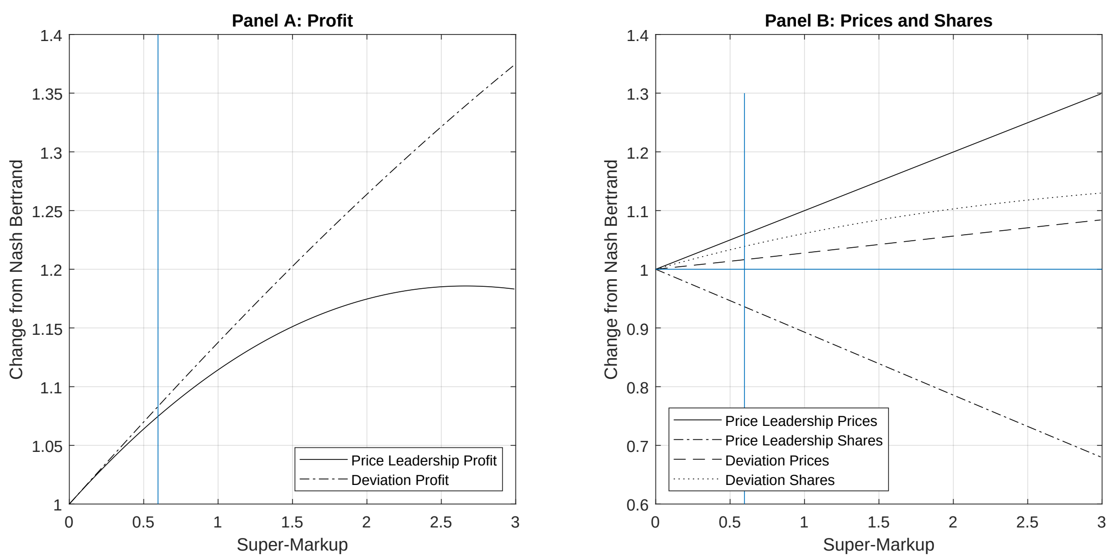

Figure B.1: Profit, Prices and Shares of ABI with Price Leadership and Deviation

Notes: The figure provides the profit (Panel A) and average prices and market share (Panel B) for ABI 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.

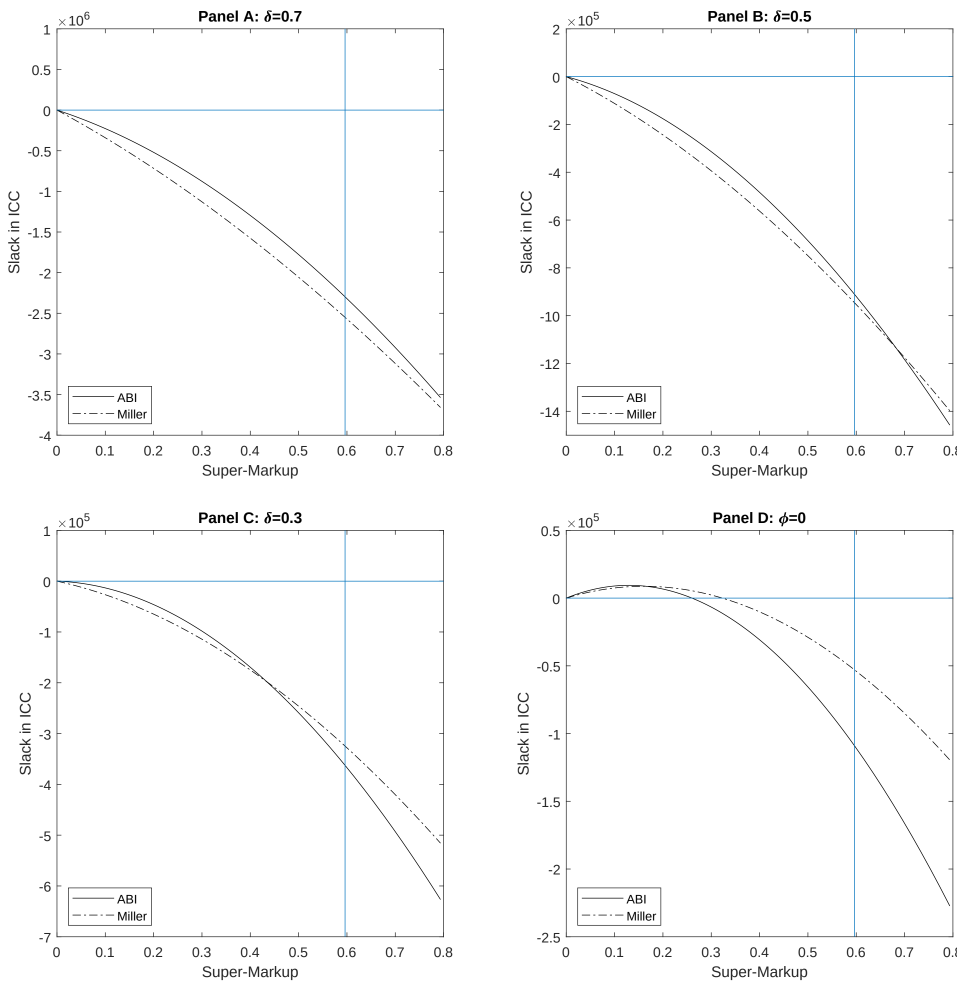

Figure B.2: ICC Slack with an ABI/Miller Coalition

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. The ICCs are 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.