Oligopolistic Price Leadership: 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

Very Preliminary
December 10, 2018

Abstract

We examine an infinitely-repeated game of oligopoly price leadership in which one firm, the market leader, announces a super-markup over static Nash prices that serves as an endogenous focal point for other firms. Two identification results make the model suitable for empirical analysis. First, marginal costs and the equilibrium super-markup can be recovered from aggregate data on price and quantities. Second, counterfactual simulations can be used to test whether incentive compatibility constraints bind and, in the affirmative case, recover the discount factor. We apply the model to the U.S. beer industry over 2005-2011. Estimation results indicate that price leadership increased prices by \$0.87 above static Nash levels after the Miller/Coors merger. Incentive compatibility constraints appear to bind. In future drafts, we will explore a number of novel counterfactual scenarios, including one in which we unwind the Miller/Coors merger and examine the implications for pricing equilibria.

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 are sometimes observed changing prices together in quick succession and by similar magnitudes. If these price changes tend to be initiated by a single firm then this pattern of behavior can reasonably be termed *oligopoly price leadership*. Stigler (1947) emphasizes that oligopoly price leadership can arise if one firm is particularly well informed about market conditions and its competitors free-ride on its knowledge. Another possibility—the focus of our research—is that oligopoly price leadership constitutes an equilibrium outcome of a repeated game in which profits exceed static Nash levels. The "Folk Theorem" asserts that, for sufficiently low discount rates, price leadership is only one of infinitely many equilibrium outcomes.[1] However, this myriad of possibilities also points to the practical benefit of price leadership: it helps firms solve their coordination problem by providing an (endogenous) focal point for prices. Further, it does so in a manner that avoids the explicit agreements often deemed necessary for violations of the antitrust statutes.

In this paper, we provide a simple model of oligopoly price leadership and show that the underlying structural parameters, including the discount factor, are identifiable with the aggregate price and quantity data typically used in empirical industrial organization. In our empirical application, we focus on the United States beer industry, which is highly concentrated and allegedly features price leadership behavior. We select the beer industry because the institutional details align with the empirical model and because high-quality scanner data are available. However, the modeling framework is more broadly applicable: Borenstein (1998), Byrne and de Roos (2017), and Chilet (2018) provide analyses of price leadership in airlines, retail gasoline, and pharmacies, respectfully, and Lanzillotti (2017) and Harrington and Harker (2017) report descriptive evidence of price leadership in industries ranging from automobile parts to text messages and titanium dioxide.[2]

The model is an infinitely repeated pricing game of perfect information. In some preliminary period, the leader selects followers to join it in a pricing coalition. Each subsequent period features two stages: (i) the leader announces some super-markup over static Nash prices, and (ii) all firms set prices and earn profit. We define a *price leadership equilibrium* as being constituted by a set of strategies in which all coalition firms price according the announced super-markup and the leader selects a super-markup to maximize its profit subject

[1] See Fudenberg and Tirole (1991) for a discussion of the folk theorem.

[2] The popular press is replete with examples of price leadership behavior. For examples, 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.

to incentive compatibility for all coalition firms. Deviations are punished by infinite Nash reversion. Fringe firms not in the coalition set prices to maximize profit conditional on all other prices. The price leadership equilibrium can be conceptualized as incorporating two refinements that collapse the infinite set of sustainable subgame perfect outcomes to a single point: all coalition firms raise price above static Nash levels by the same super-markup, and the leader determines the magnitude of the super-markup.

We provide two novel identification results that allow for estimation. First, we provide a constructive proof that, for any candidate super-markup (an object of estimation) and given some demand system, the marginal costs that rationalize prices can be recovered using the static Nash first order conditions. Four numerical steps are required for this inference, rather than the single step possible with many static oligopoly models (e.g., as in Rosse (1970); Berry et al. (1995)). The econometrician must be able to distinguish the coalition firms from the fringe firms, if any, based on prior knowledge. The equilibrium super-markup and marginal cost parameters then can be estimated given instruments that provide exogenous variation in residual demand, such as competitors' costs and characteristics (e.g., Berry et al. (1995); Berry and Haile (2014); Gandhi and Houde (2016)). Implementation can be computationally intensive but does not require any assumptions about the precise form of punishment off the equilibrium path.

The second identification result is that the econometrician can conduct a series of counterfactual experiments to test whether an incentive compatibility constraint (ICC) binds and, in the affirmative case, to recover the discount factor. The first experiment recovers the super-markup that maximizes the leader's stage-game profit; if this differs from the equilibrium super-markup then the econometrician can infer that an ICC binds. This provides an addition equation that can be exploited, along the lines of

$$\pi^D - \pi^C = \frac{\delta}{1-\delta}(\pi^C - \pi^N)$$

where (π^D, π^C, π^N) are stage-game profit under deviation, price leadership, and static Nash, respectively, and δ is a discount factor. Price leadership profit can be calculated given the data and marginal costs, and deviation and static Nash profits can be computed numerically with counter-factual simulations. The discount factor that rationalizes the equilibrium super-markup then can be recovered. Implementation requires the econometrician to specify the unobserved off-equilibrium paths, such as punishment via infinite Nash reversion. These assumptions are not overly strong, as the discount factor can be interpreted somewhat loosely as scaling the severity of punishment (Rotemberg and Saloner (1986)) or incorporating the continuation probability in a repeated game of finite length.

We apply the model to data on the U.S. beer industry over 2005-2011. We focus especially on the MillerCoors joint venture, which was formed in 2008 as a merger of SAB Miller and Molson Coors. After the merger, MillerCoors and Anheuser-Busch InBev (ABI), account for nearly two-thirds of revenues in the US beer industry. Retail prices increase after the merger in a manner that is difficult to reconcile with static Nash price competition (Miller and Weinberg (2017)). Indeed, qualitative evidence that we summarize in the body of the paper suggests that pricing strategy in the industry takes to form of price leadership: ABI proposes annual price increases that are matched by MillerCoors, while other competitors compete in price more aggressively.[3]

We estimate a version of the price leadership model in which ABI is the leader, MillerCoors is the (only) follower in the pricing coalition, and other firms compose the fringe. To estimate the equilibrium super-markups, we make the identifying assumptions that (i) competition is static Nash before the Miller/Coors merger, and (ii) that the unobserved marginal costs of ABI do not change on average relative to fringe firms before vs. after the merger. This estimation strategy, which leans heavily on the qualitative evidence, eliminates the need to observe exogenous variation in residual demands. Throughout we rely on the demand estimates of Miller and Weinberg (2017). Our estimation results indicate equilibrium super-markups of \$0.87 by the end of 2011.

The magnitude of equilibrium super-markups appear to be limited by binding incentive compatibility constraints, as a counterfactual simulation reveals that higher super-markups produce greater stage-game profit for ABI. We incorporate a "friction" into the ICCs to capture some preference for static Nash competition, which could be present due to antitrust litigation risk, monitoring costs, or other factors outside the model. We then jointly calculate the discount factors and frictions that rationalize the equilibrium super-markups under the assumption that punishment takes the forms of infinite Nash reversion. In future drafts of the paper, we plan to treat these discount factors and frictions as structural parameters and evaluate a number of novel counterfactuals, including a scenario in which we explore whether price leadership survives an unwinding of the MillerCoors joint venture.

Our research connects to a number of literatures. As best we can ascertain, contri-

[3]The DOJ closed its investigation into the Miller/Coors merger in June 2008 after reviewing evidence that the transaction would create substantial cost efficiencies. Those efficiencies appear to have been at least partially realized, as discussed in Ashenfelter et al. (2015). See also "Statement of the Department of Justice's Antitrust Division on Its Decision to Close Its Investigation of the Joint Venture Between SABMiller PLC and Molson Coors Brewing Company," press release, 5 June 2008. Sweeting and Tao (2017) provides an alternative explanation for the post-merger price increases based on a model of signaling.

butions on oligopoly price leadership itself are relatively few. Articles that provide description and/or atheortical empirical analysis include Stigler (1947), Nicholls (1949), Borenstein (1998), Byrne and de Roos (2017), Lanzillotti (2017), Harrington and Harker (2017), and Chilet (2018). Another uses experimental techniques (Harrington et al. (2016)). As previously mentioned, the leader's price announcement can be interpreted as providing an endogenous focal point for coalition firms. 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 in the credit card industry.

We draw on a number of theoretical articles in building the model of oligopoly price leadership. In some sense, most similar is the canonical Rotemberg and Saloner (1986) model of collusion: because we assume the leader can adjust super-markups to satisfy incentive compatibility, deviation does not occur in equilibrium and conditions exist under which prices are counter-cyclical. As in d'Aspremont et al. (1983), Donsimoni et al. (1986), and Bos and Harrington (2010), we incorporate the possibility that not all firms participate in the pricing coalition. However, these papers focus on homogeneous goods markets, whereas we study a differentiated products, price-setting environment.

A recent empirical literature develops econometric tests for static Nash competition using identification strategies that are robust to the infamous Corts (1999) critique (e.g., Ciliberto and Williams (2014), Igami (2015), Michel (2016), Sullivan (2016), and Miller and Weinberg (2017)).[4] Rather than focusing on reduced-form coefficients to summarize the impact of coordination, we posit and estimate an explicit dynamic game. This allows for a greater range of insights about the composition of pricing coalitions and the economic conditions under which supra-competitive pricing may emerge.

From a methodological standpoint, our research is most similar to Igami and Sugaya (2017) and the contemporaneous work of Fan and Sullivan (2018). The former paper studies the 1990s vitamin cartels. It estimates an explicit collusive super-game and uses counterfactual simulations to explain the persistence of collusion for some vitamins and the failure of collusion for others. Some interesting aspects of the price leadership model, including binding incentive compatibility constraints and the possibility of partial coalitions, are not examined. 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

[4]These articles use identification strategies that are similar to the seminal research of Porter (1983).

conditions so the models are mostly applicable in different settings.

The paper proceeds as follows. Section 2 sketches the price leadership model. Section 3 provides background on the beer industry and presents qualitative evidence that the price leadership model is a good fit for the industry. Section 4 develops the supply-side estimates and counterfactual results that we obtain using the price leadership model.

2 Price Leadership Model

2.1 Primitives

We now develop our 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)). In our application, there is seemingly little empirical variation in the leadership role, so by necessity we take it as exogenously determined.

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 X_t is realized and observed by all firms. Competition then plays out in two stages:

(i) The leader announces a super-markup, $m_t \geq 0$, above static Nash prices (to be defined), given history h_t (also to be defined).

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

The timing of the game mimics a frequently observed practice in which one firm announces a price change before the new prices become available to consumers, possibly in order to signal competitors.[5] In the model, the super-markup is not binding on any firm, including the leader, in the pricing stage. Nonetheless, firm beliefs create a important distinction between followers and the fringe firms.

[5]For examples, see Borenstein (1998) on the U.S. airlines industry, Section 3 on the beer industry, or "Your Chocolate Addiction is Only Going to Get More (and More, and More) Expensive" by Roberto A. Ferdman, Washington Post, 18 July 2014.

We use a standard profit function to characterize payoffs. In particular, the profit of firm f in each period $t = 1 \ldots, \infty$ is given by

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

where $p_t = (p_{1t} \ p_{2t} \ \ldots p_{Jt})'$ is a vector of prices realized in the second stage of period t, $c_j(X_t)$ is a constant marginal cost function, and $q_j(p_t, X_t)$ is taken from a demand function satisfying standard normalcy conditions.

We assume the cost and demand functions are common knowledge, and that each firm f observes prices and quantities each period. Different assumptions regarding the economic states are possible. The assumption that X_t is stochastic and iid across periods yields the history

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

This treatment is theoretically appealing because it avoids certain scenarios in which price leadership unravels due to an adverse realization of X_t. Finally, we assume that firm actions do not affect the economic states.[6]

2.2 Equilibrium

We assume the firms are playing a *price leadership equilibrium* in which (i) coalition firms set prices that exceed the static Nash level by some amount, m_t, determined by the leader, and (ii) each fringe firm maximizes its profit conditional on all other prices. Some mathematical preliminaries are helpful in formalizing this equilibrium. Consider some firm f that maximizes its profit in period t conditional on the prices of other firms. The prices p_{jt} for $j \in \mathbb{J}_f$ satisfy standard first order conditions:

$$p_{jt} = c_j(X_t) - \sum_{k \in \mathbb{J}_f} \left[\frac{\partial q_k(p_t, X_t)}{\partial p_{jt}} \right]^{-1} q_k(p_t, X_t) \tag{2}$$

If all prices satisfy the first order conditions then economic outcomes correspond to the static Nash equilibrium; we label this price vector p_t^N. As we develop below, fringe firms always price in accordance with equation (2) in equilibrium, but coalition firms may not.

[6] In the empirical implementation, we assume that firms know the entire sequence of past and present economic states $(X_\tau)_{\tau=1}^{\infty}$, which avoids having to specify a data generating process for the observables. This alternative assumption is plausible in the U.S. beer industry have strong incentives to be informed about future economic conditions for a variety of reasons. We do not observe any such instances in the data.

The first order conditions also characterize optimal deviations from the price leadership equilibrium. Denote the vector of equilibrium prices as $p_t^C(m_t)$. The elements of this vector are $p_{jt}^C(m_t) = p_{jt}^N + m_t$ for products in the coalition and given by equation (2) for fringe firms. By construction, for any $m_t > 0$, the price vector p_t^C does not maximize the single-period profit of any coalition firm conditional on others' prices. If some firm f were to deviate, it would maximize its profit by solving equation (2), holding fixed competitor prices. Let the $J \times 1$ vector, $p_t^{D,f}(m_t)$ contain the prices that arise in the event of deviation by firm f, including elements $p_{kt}^C(m_t)$ for products $k \notin \mathbb{J}_f$. There are $f = 1, \ldots, F$ such vectors corresponding to each possible single-firm defection. With fringe firms, $f \notin \mathbb{C}$, we have $p_t^{D,f}(m_t) = p_t^C(m_t)$ because there is no unilaterally profitable deviation.

The equilibrium concept is subgame perfection. If markups that exceed the static Nash level are to be sustained, then strategies must incorporate some form of punishment for deviations from p_t^C. Our starting assumption is that all firms revert to the static Nash equilibrium in every period following a deviation. The corresponding incentive compatibility constraint (ICC) for each firm f takes the form:

$$\begin{aligned} \sum_{j \in \mathbb{J}_f} \pi_j(p_t^{D,f}(m_t), X_t) - [\pi_j(p_t^C(m_t), X_t) - F(m_t)] \leq \\ \frac{\delta}{1-\delta} \sum_{j \in \mathbb{J}_f} E\left[\pi_j(p^C(m), X) - F(m) - \pi_j(p^N, X)\right] \end{aligned} \tag{3}$$

where $\delta \in (0,1)$ is a discount factor, and $F(m_t)$ is a "friction" that can be interpreting as capturing some per-period preference for static Nash competition, due to antitrust litigation risk, monitoring costs, or other considerations outside the model. The left-hand side (LHS) captures the increase in profit due to deviation while the right-hand side (RHS) captures the stream of future profit that is lost in the event of deviation.[7] Note that the ICCs hold trivially for fringe firms because the LHS equals zero and the RHS is weakly positive for any super-markup.

With these preliminaries in hand, we can define our price leadership equilibrium:

Price Leadership Equilibrium (Definition). *The equilibrium is constituted by the following strategies:*

1. *In the second stage of each period $t > 0$, firms price according to $p_t^C(m_t)$ if the incentive compatibility constraints of equation* (3) *hold for all firms and all firms have priced*

[7] In a slight abuse of notation, we remove time subscripts from the economic state X and the price vectors on the RHS because their expected values are time invariant under histories h_t.

according to $p^C_{t-s}(m_{t-s})$ for all $s = 1, \ldots, t-1$. Otherwise, firms price according to p^N_t.

2. *In the first stage of each period $t > 0$, the leader announces a super-markup, m_t, that maximizes its profit subject to the incentive compatibility constraints of equation* (3) *holding for all firms.*

3. *In period $t = 0$, the leader proposes a coalition that maximizes the present value of its profit, taking as given the strategies outlined above for $t > 1$.*

Thus, price leadership is a form of collusion that may raise prices beyond those in a static Nash equilibrium.

Given this formulation, we can show that any super-markup that raises profits above the static Nash level can be sustained by some discount factor (as in Fudenberg and Tirole (1991)). This gives the following proposition:

Proposition 1A: *For any m^* where price leadership is more profitable than static Nash in expectation, i.e. where*

$$E\left[\pi_j\left(p^C(m^*), X\right) - F(m^*) - \pi_j\left(p^N, X\right)\right] > 0,$$

there exists some $\widetilde{\delta}(m^*)$ such that if $\delta > \widetilde{\delta}(m^*)$ then m^* can be sustained in a SPE.

Proof: The RHS of the ICC $\to \infty$ as $\delta \to 1$ so the ICC is non-binding:

$$\pi_j\left(p^{D,b}_t(m_t), X_t\right) - \left[\pi_j\left(p^C_t(m_t), X_t\right) - F(m_t)\right] \leq \frac{\delta}{1-\delta} E\left[\pi_j\left(p^C(m_t), X\right) - F(m_t) - \pi_j\left(p^N, X\right)\right].$$

We also have a related result, that if the friction is zero, some positive super-markup can be sustained in equilibrium:

Proposition 1B: *If $F(m_t) = 0 \;\; \forall \;\; m_t$ then, for any $\delta \in (0, 1)$, there exists some $m_t > 0$ that can be sustained in SPE.*

Proof: The full proof is excluded for brevity, but the result can be seen from the ICC when $F(m_t) = 0$:

$$\pi_j\left(p^{D,b}_t(m_t), X_t\right) - \pi_j\left(p^C_t(m_t), X_t\right) \leq \frac{\delta}{1-\delta} E\left[\pi_j\left(p^C(m_t), X\right) - \pi_j\left(p^N, X\right)\right].$$

Thus, the inclusion of a friction is key to fitting a dataset where one believes that there were periods where collusion with a positive super-markup did not occur.

By inspection, the equilibrium also is unique if, for any set of firms that seek to maximize per-period profit conditional on the prices of all other firms, there is a unique vector of prices that satisfy the first order conditions of equation (2). Outside of the symmetric case, the equilibrium is not generally Pareto optimal for the coalition firms, as the leader acts in its own interest and side payments are not permitted. The comparative statics of the model somewhat resemble the canonical Rotemberg and Saloner (1986) model of collusion. If some ICCs bind, the leader may find it necessary to choose a smaller super-markup in prosperous periods so as to deter deviation.

2.3 A Logit Numerical Example

In order to build intuition, here we discuss a numerical example using a logit demand function. The logit, although an admittedly restrictive demand form, serves to illustrate many of the features of our price leadership model while remaining transparent and mathematically tractable. Although we use a much richer demand system in our empirical application, a number of the effects we see there also appear in this simplified version.

In this section, we abstract from the economic shocks in X_t and assume single product firms. We also ignore the friction $F(m_t)$. Thus, the index j is equivalent to f, and all market conditions are the same across periods (meaning we can drop the subscript t). We parameterize demand as $q_j(p) = s_j M$, where M is the market size, and the market share is given by

$$s_j(p) = \frac{\exp(\beta_j - \alpha p_j)}{1 + \sum_{k \in \mathbb{J}} \exp(\beta_k - \alpha p_k)}.$$

The $\mathbb{J}$ is the set of firms (or products) in the market. This expression assumes that there is an outside good whose mean valulation has been normalized to zero.

The task for the leader is to choose a super-markup m that maximizes its profits, subject to the ICCs for all firms. Plugging the logit demand into the leader's profit function, we find that the leader has a well-defined maximization problem. For discussion purposes, assume first that none of the ICCs bind. Then solving for the first order condition gives

$$p_1^N + m - c_1 = \frac{1}{\alpha \left[1 - \sum_{j \in \mathbb{C}} s_j(p^C) - \left(\sum_{j \in \mathbb{C}} s_j(p^C)\right) \left(\sum_{k \in \mathbb{J} \setminus \mathbb{C}} \frac{(s_k(p^C))^2}{1 - s_k(p^C)}\right)\right]},$$

for some coalition $\mathbb{C}$. Here $s_j(p^C)$ is the market share for firm j when all firms are pricing according to the vector p^C. If maintaining incentive compatiblity was no issue, this expression would determine the leader's choice of super-markup.

Recall that the price first order condition for firm 1 in a standard static Nash game is given by $p_1^N - c_1 = 1/\alpha(1-s_1)$. This expression balances the firm's desire to raise price against the consumer propensity to substitute away in response, as measured by the price coefficient α times $1-s_1$.[8] Note that when all firms are in the coalition, the price leadership first order condition simplifies to

$$p_1^N + m - c_1 = \frac{1}{\alpha\left[1 - \sum_{j\in\mathbb{C}} s_j(p^C)\right]}.$$

This expression is highly similar to the static Nash condition. However, in the price leadership equilibrium, the leader maximizes profits accounting for substitition outside the entire coaltion (as reflected in $1-\sum_{j\in\mathbb{C}} s_j$), instead of beyond just itself (as reflected in $1-s_1$).

When not all firms are in the coalition, the first order condition has an additional term in the denominator, $\sum_{j\in\mathbb{C}} s_j \sum_{k\in\mathbb{J}\setminus\mathbb{C}} s_k^2/(1-s_k)$. This expression accounts for the effect that raising the super-markup has on the prices of fring firms. By applying the implicit function theorem to the first order condition for non-coalition firms, we find that the derivative $\partial p_k/\partial\theta$ is given by $(s_k/(1-s_k))\sum_{j\in\mathbb{C}} s_j$ for all $k \notin \mathbb{C}$. Therefore, the additional term in the first order condition of the leader is equal to the weighted average derivative of fringe prices with respect to the super-markup. The leader takes the anticipated reaction of firms outside the coalition into account when choosing which markup to announce.

In order to examine the effect that potentially binding ICCs have on this problem, it is helpful to look at a concrete numerical example. Assume there are three firms in the market (not counting the outside good). The first two firms are both high quality and low cost, while the third is low quality and high cost.[9] There are three potential coalitions that the leader could form: it could partner either with firm 2 or firm 3, or all three firms could join one coalition together.

Imagine that firm 1 is considering forming a coalition between it and both firms 2 and 3. As explained in the previous subsection, price leadership allows firms to raise prices higher than they otherwise would in a static setting, subject to the ICCs. This is illustrated in Figure 1. Here, we graph the best-response functions for firm 1 and firm 2, which show the optimal static prices that each firm should choose when faced with a given price from the

[8] Remember that the logit own-price elasticity is given by $\alpha(1-s_1)p_1$.

[9] Specifically, we assume that $\beta_1 = \beta_2 = 3$, $c_1 = c_2 = 0$, $\beta_3 = 1$, and $c_3 = 2$. The price coefficient α is set at 1.5, the discount factor δ is 0.4, and the market size M is 20.

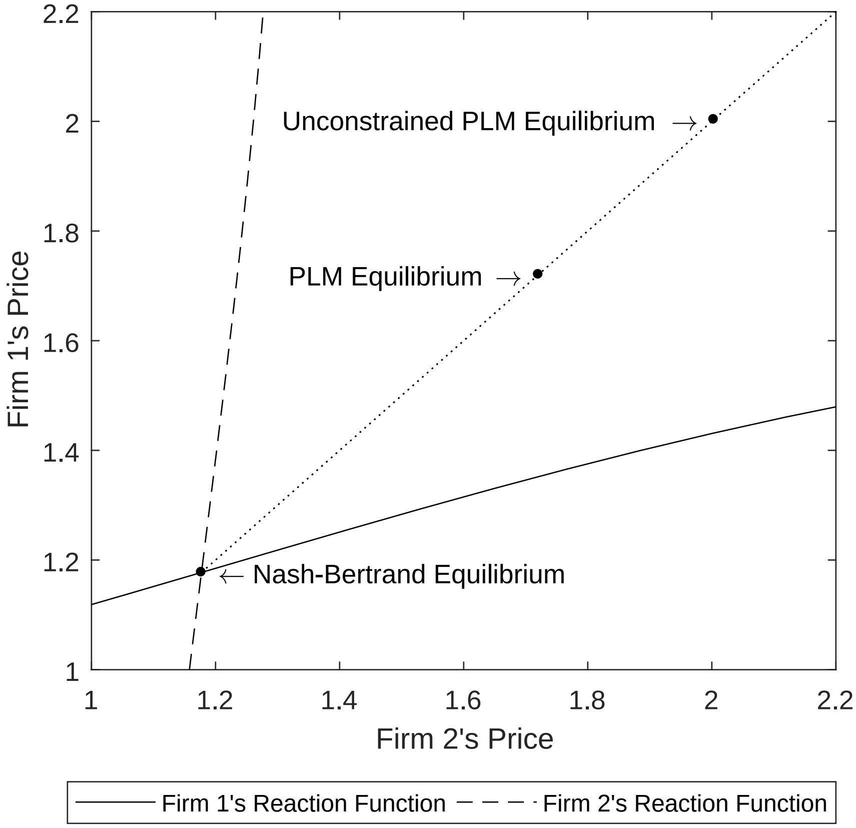

Figure 1: Logit Best Response Functions

Notes: This figure depicts a market with three firms, where $\beta_1 = \beta_2 = 3$, $\beta_3 = 1$, $c_1 = c_2 = 0$, and $c_3 = 2$. Furthermore, $\alpha = 1.5$, $\delta = 0.4$, and $M = 20$. The best-response function for firm 1 graphs the single-period profit maximizing price that firm 1 should choose when faced with a given price from firm 2. The best-response function for firm 2 is defined analogously relative to the price of firm 1. In all cases, firm 3 is allowed to respond with its profit maximizing price. The point where these reaction functions cross is the Nash-Bertrand equilibrium. The point labeled "Unconstrained PLM Equilibrium" denotes the super-markup that firm 1 (the leader) would choose for a coalition of it, firm 2, and firm 3, ignoring any possible ICCs. The point labeled "PLM Equilibrium" denotes the super-markup that firm 1 would choose for the same coalition, taking into account all ICCs.

other. The static Nash price levels are given by the point where these functions cross. The point labeled "Unconstrained PLM Equilibrium" is where firm 1 would like to raise prices if ICCs were no object. The actual price leadership equilibrium lies on the dotted line between these two points.

In this case, firm 1 can only achieve part of the unconstrained price leadership increase, because the ICC for firm 3 binds. The constrained super-markup (which is 0.54) is approx-

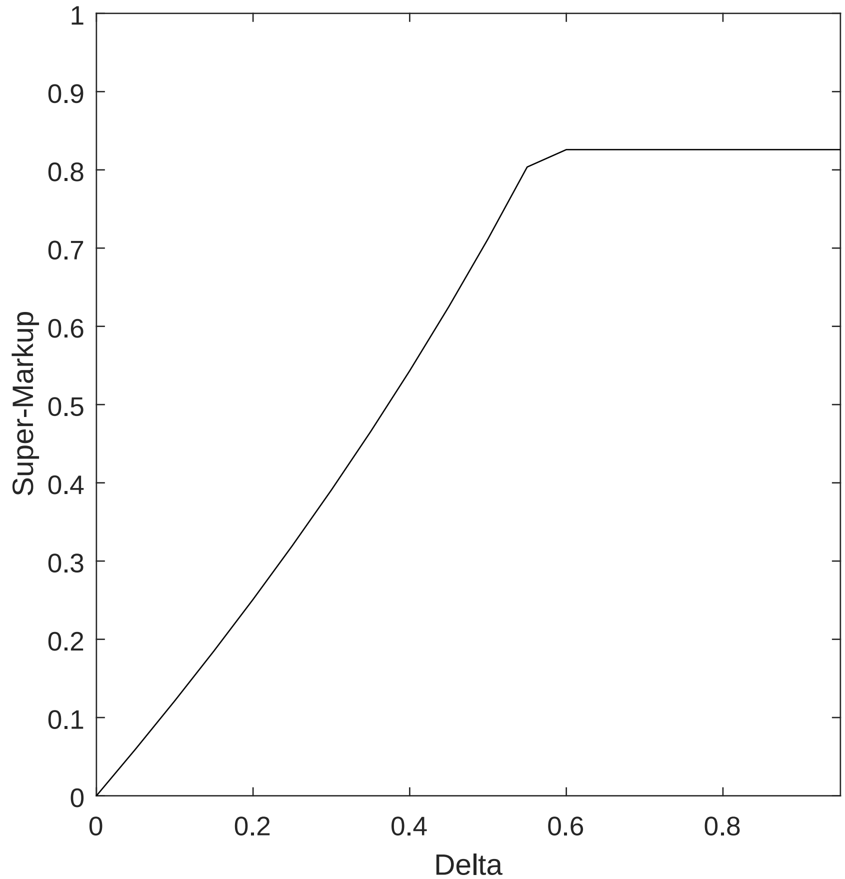

Figure 2: Logit Super-Markups as Delta Varies

Notes: This figure depicts a market with three firms, where $\beta_1 = \beta_2 = 3$, $\beta_3 = 1$, $c_1 = c_2 = 0$, and $c_3 = 2$. Furthermore, $\alpha = 1.5$ and $M = 20$. The discount factor δ is allowed to increase from zero towards one. Super-markups are calculated for a coalition containing all three firms.

imately 66% of the unconstrained super-markup. In part, this is due to the relatively low discount factor, $\delta = 0.4$. As the discount factor increases, each firm becomes more patient and less willing to forego future coalition profits in order to deviate. This relationship is pictured in Figure 2, which graphs the equilibrium super-markup for different levels of δ. As the discount factor increases, the super-markup rises towards the optimal unconstrained level, which is reached around $\delta = 0.6$. Thus, consistent with Proposition 1A, any super-markup up to the unconstrained optimal level can be sustained by some discount factor.

The differences between firm 3 versus firm 1 and firm 2 are another key factor. As a lower quality, higher cost producer, firm 3 finds it more attractive to undercut its coalition partners and steal sales as the super-markup increases. Figure 3 illustrates these various

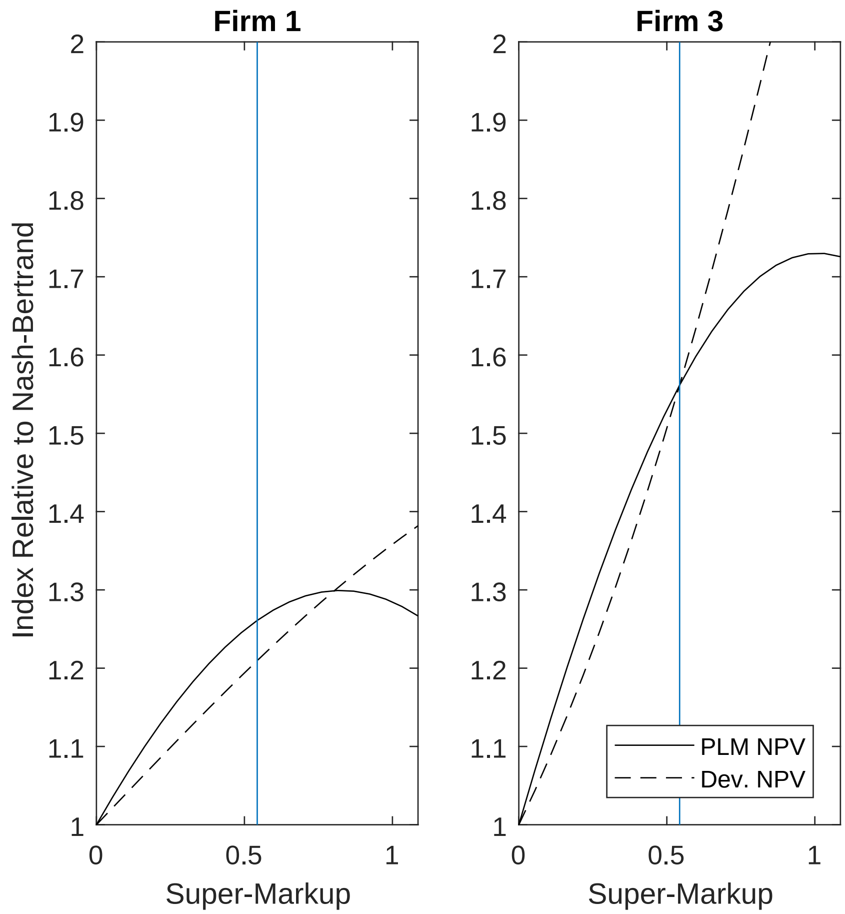

Figure 3: Logit Net Present Values for Dissimilar Firms

Notes: This figure depicts a market with three firms, where $\beta_1 = \beta_2 = 3$, $\beta_3 = 1$, $c_1 = c_2 = 0$, and $c_3 = 2$. Furthermore, $\alpha = 1.5$, $\delta = 0.4$, and $M = 20$. Calculations are done for a coalition containing all three firms. The left panel is for firm 1 (the leader), and the right panel is for firm 3 (a follower). The solid curve labeled "PLM NPV" shows the net present value of profits if the firm adheres to the price leadership equilibrium and raises price by the super-markup in every period. The dashed curve labeled "Dev. NPV" shows the net present value of profits should the firm deviate in the current period and then earn its Nash-Bertrand profit in all subsequent periods. All curves have been normalized via dividing through by the net present value of earning Nash-Bertrand profits in every period. The vertical line marks the equilibrium price leadership super-markup.

forces that must be balanced. The solid curves show the present value of profits (relative to earning only the Nash-Bertrand level) if each firm charges the price leadership super-markup in every period. These curves increase at first and then fall off, which is consitent with the price leadership problem having a well-defined unconstrained maximum, as was shown above. The dashed curves show the present value of profits (also relative to earning only the Nash-Bertrand level) if a firm deviates in the current period, and then earns Nash-Bertrand profits in every following period. These curves rise nearly linearly in the super-markup, as an increasing coalition markup provides increasing profit opportunities to deviating.[10] Eventually the price leadership and deviation curves cross. The left panel is for firm 1 (firm 2 would be identical), while the right panel is for firm 3. As indicated by the vertical lines, the two curves cross at a lower super-markup for firm 3 than for firm 1, meaning that the highest super-markup that can be sustained is that which causes the ICC for firm 3 to bind. We see that although increasing the super-markup leads to relatively higher gains to firm 3 from joining the coalition (its price leadership curve is steeper than than for firm 1), it has an even larger impact on the gains to deviating (its deviation curve is even steeper).

Figure 4 illustrates what happens when the difference in marginal costs between firm 3 versus firm 1 and firm 2 changes. The dashed curves show each firm's single-period profits in the price leadership equilbrium when all three are in the coalition. The solid curves show each firm's single-period profits when only firms 1 and 2 are in the coalition (meaning firm 3 can price without regards to the announced super-markup). The top panel is for firm 1 (firm 2 would be identical), while the bottom panel is for firm 3. As shown in the bottom panel, as the marginal cost for firm 3 increases, its profits from both coalitions decrease, as would be expected. Of the two graphed coalitions, firm 3 would prefer the one containing only firms 1 and 2, since this allows firm 3 to price optimally in response without triggering punishment. However, the choice of coalition is up to the leader, firm 1. In the top panel, we see that firm 1 prefers to have all three firms in the coalition until the marginal cost for firm 3 rises to between $c_3 = 1$ and $c_3 = 1.5$. At this point, the binding nature of the ICC for firm 3 becomes too limiting relative to just excluding it from the coalition entirely. This can be seen in Figure 5, which shows that as the marginal cost for firm 3 increases, the Lagrange multiplier on its ICC also increases. Meanwhile, the equilibrium super-markup falls.

Therefore, in our numerical example where $c_3 = 2$, the leader would choose to form a coalition of only itself and firm 2. This would allow the coalition to raise its prices by more (from a markup of 0.54 to 0.78), closer to the unconstrained optimal super-markup. Indeed, as shown by Figure 6, the chosen super-markup is nearly identical to that which maximizes price leadership profits for firm 1. These graphs are analogous to those in Figure 3, except they show the net present values for firm 1 and firm 2 in a coalition that excludes firm 3. The constrained super-markup is 97% of the unconstrained level. Thus, by forming a coalition with only the other low cost/high quality firm, the leader is able to better align incentives and raise profits.[11] In turn, this suggests that a merger that increases similarity

[10] This occurse because prices are strategic complements.

[11] Although not graphed here, the coalition of firms 1 and 2 also dominates that of firms 1 and 3.

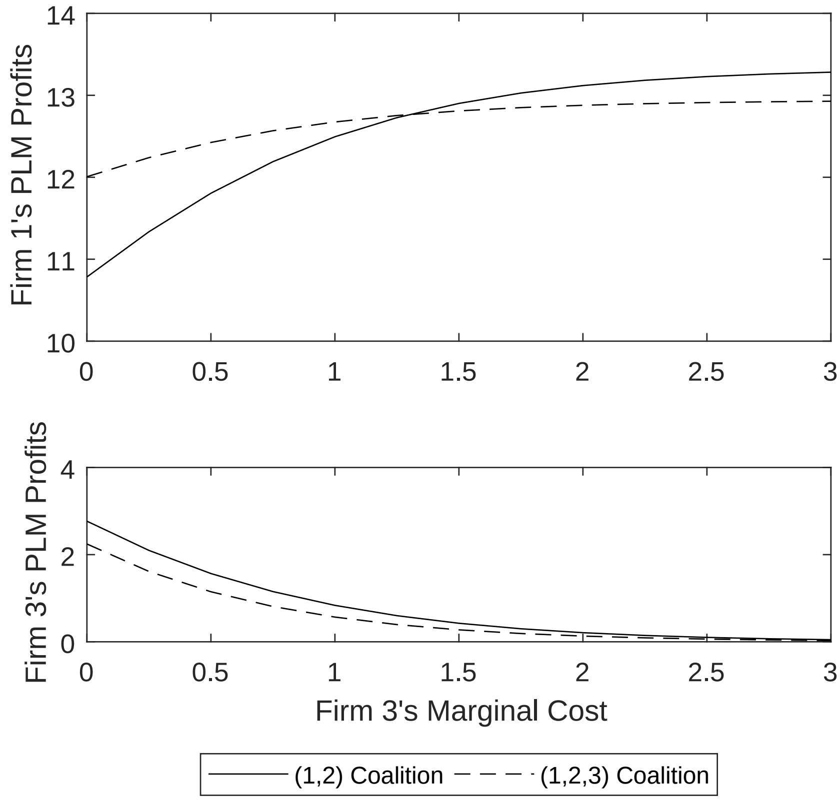

Figure 4: Logit Price Leadership Profits as Marginal Cost Varies

Notes: This figure depicts a market with three firms, where $\beta_1 = \beta_2 = 3$, $\beta_3 = 1$, and $c_1 = c_2 = 0$. Furthermore, $\alpha = 1.5$, $\delta = 0.4$, and $M = 20$. The single-period profits that a firm earns when all coalition members adhere to the price leadership super-markup are graphed against different values of marginal cost for firm 3. The solid line depicts profits when the coalition only contains firms 1 and 2 (meaning firm 3 is allowed to choose its price separately from the super-markup), while the dashed line depicts profits when the coalition contains all three firms.

between firms in terms of cost or quality has the potential to raise the likelihood of sucessful coordination.

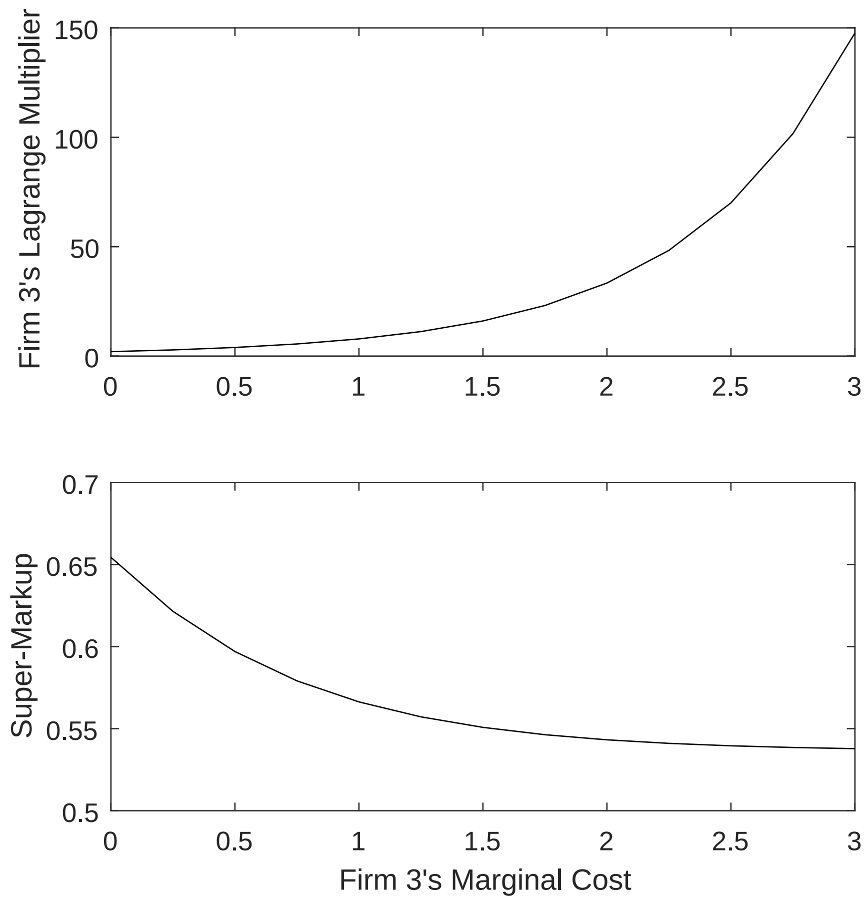

Figure 5: Logit Lagrange Multiplier and Super-Markup as Marginal Cost Varies

Notes: This figure depicts a market with three firms, where $\beta_1 = \beta_2 = 3$, $\beta_3 = 1$, and $c_1 = c_2 = 0$. Furthermore, $\alpha = 1.5$, $\delta = 0.4$, and $M = 20$. Calculations are done for a coalition containing all three firms. The top panel graphs the Lagrange multiplier on the ICC of firm 3 for different values of marginal cost for firm 3, while the bottom panel graphs the super-markup.

3 The U.S Beer Market

3.1 Market Structure and Pricing

Table 1 shows revenue-based market shares at two-year intervals over 2001-2011, based on retail scanner data that we describe later in this section. The brands of five brewers—ABI, SABMiller, Molson Coors, Grupo Modelo, and Heineken—account for approximately 80% of total retail revenue. A main focus of the paper is on the Miller/Coors merger. The new entity, MillerCoors, accounts for around 30% of revenue in 2009 and afterwards.

Qualitative evidence is consistent with softened competitive intensity coincident with

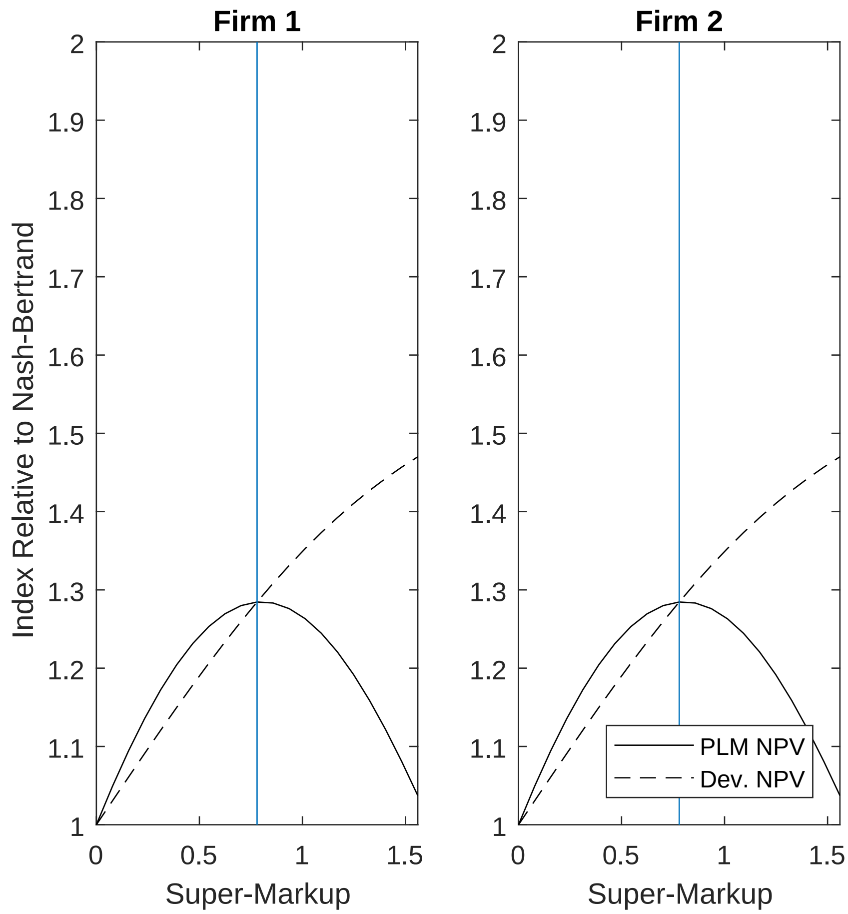

Figure 6: Logit Net Present Values for Symmetric Firms

Notes: This figure depicts a market with three firms, where $\beta_1 = \beta_2 = 3$, $\beta_3 = 1$, $c_1 = c_2 = 0$, and $c_3 = 2$. Furthermore, $\alpha = 1.5$, $\delta = 0.4$, and $M = 20$. Calculations are done for a coalition containing only firms 1 and 2. The solid curve labeled "PLM NPV" shows the net present value of profits if the firm adheres to the price leadership equilibrium and raises price by the super-markup in every period. The dashed curve labeled "Dev. NPV" shows the net present value of profits should the firm deviate in the current period and then earn its Nash-Bertrand profit in all subsequent periods. All curves have been normalized via dividing through by the net present value of earning Nash-Bertrand profits in every period. The vertical line marks the equilibrium price leadership super-markup.

the consummation of the merger. 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. The tenor of the annual reports changes around the time of the merger. In its 2009 report, SABMiller attributes increasing earnings before interest, taxes, and amortization expenses to "robust pricing" and "reduced promotions

Table 1: Revenue 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.

and discounts." In its 2010 and 2011 reports, it references "sustained price increases" and "disciplined revenue management with selected price increases."[12]

Figure 7 supports the qualitative evidence using retail scanner data that we describe below. The prices of Bud Light, Miller Lite and Coors Light increase starkly after the merger, breaking a downward pre-merger trend. The prices of more expensive Corona Extra (Modelo) and Heineken brands continue on their trend, suggesting that the change in ABI and MillerCoors pricing is not due to common cost factors. Miller and Weinberg (2017) provide rigorous econometric analyses that suggests it is difficult to reconcile these price patterns with static Nash competition but do not explore why competition softened nor the form of any specific post-merger supergame.

The Complaint filed by the DOJ to enjoin the 2013 acquisition of Grupo Modelo by ABI suggests that price leadership may have emerged in the wake in the Miller/Coors merger. In particular, the Complaint alleges that

> ABI and MillerCoors typically announce annual price increases in late summer for execution in early fall. In most local markets, ABI is the market share leader and issues its price announcement first, purposely making its price increases transparent to the market so its competitors will get in line. In the past several years, MillerCoors has followed ABI's price increases to a significant degree.[13]

The rationale behind this pricing strategy—known within ABI as the "Conduct Plan"—

[12] 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). ABI's annual reports in the post-merger years are more opaque.

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

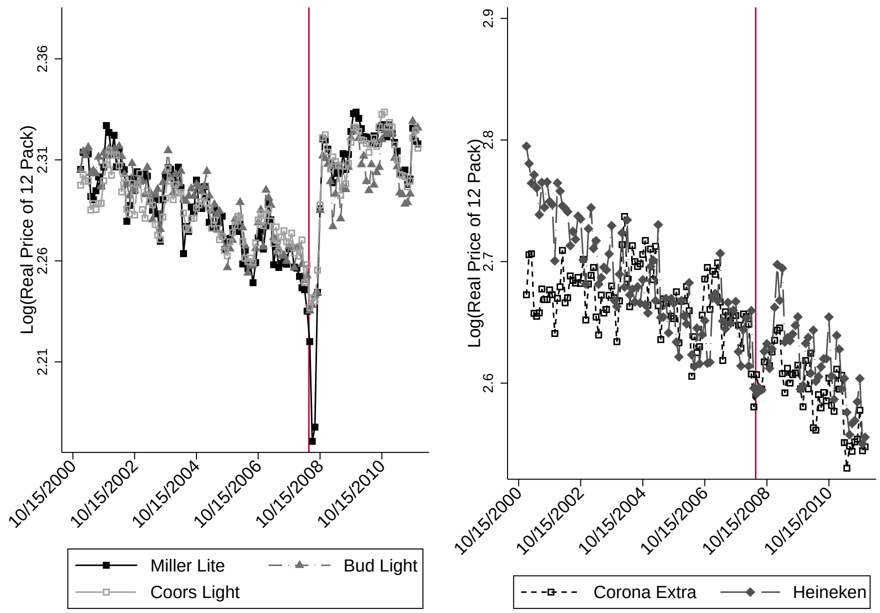

Figure 7: Average Retail Prices of Flagship Brand 12-Packs

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).

appears to be consistent with the notion that price leadership simplifies the dynamic pricing game in a way that more easily allows for profits above static Nash to be realized. The DOJ Complaint quotes from the normal course 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."

3.2 Data Sources

We use the data of Miller and Weinberg (2017) in our analysis. The main data source is the retail scanner data from the IRI Academic Database (Bronnenberg et al. (2008)). There are 167,695 observations at the product-region-month-year level, spanning 2001-2011, where products are defined as brand×package size combinations. Attention is restricted to 13 popular brands: These 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. All prices are deflated using the CPI and are reported in 2010 dollars. A number of supplementary data are employed; see Miller and Weinberg (2017) for details.

3.3 Consumer Demand

We rely the random coefficient nested logit (RCNL) model of Miller and Weinberg (2017) to characterize consumers substitution. 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

$$u_{ijrt} \quad = \quad x_j \beta_i^* - \alpha_i^* p_{jrt} + \sigma_j^D + \tau_t^D + \xi_{jrt} + \bar{\epsilon}_{ijrt} \tag{4}$$

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 an inside good) and calories, which is highly correlated with alcohol content. 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

$$\bar{\epsilon}_{ijrt} = \zeta_{igrt} + (1 - \rho)\epsilon_{ijrt} \tag{5}$$

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.

We refer readers to Miller and Weinberg (2017) for the details of implementation and a discussion of the identifying assumptions. Table 2 presents the results of the baseline model (RCNL-1) and a simpler nested logit model in which we impose the restriction $\Pi = 0$. The coefficients are precisely estimated and take the expected signs. The median own price elasticities are from -3.76 to -4.74 with the RCNL-1 and NL-1 models. The market price elasticities are much lower, indicating that most substitution occurs within the inside goods,

Table 2: Baseline Demand Estimates			
Demand Model:		NL-1	RCNL-1
Data Frequency:		monthly	monthly
Variable	Parameter	(i)	(ii)
Price	α	-0.1377	-0.0887
		(0.0933)	(0.0141)
Nesting Parameter	ρ	0.6067	0.8299
		(0.0994)	(0.0402)
Demographic Interactions			
Income×Price	Π_1		0.0007
			(0.0002)
Income×Constant	Π_2		0.0143
			(0.0051)
Income×Calories	Π_3		0.0043
			(0.0016)
Median Own Price Elasticity		-3.76	-4.74
Median Market Price Elasticity		-1.09	-0.60
Median Outside Diversion		29.86%	12.96%

Notes: The table shows the baseline demand results. Estimation is with 2SLS in column (i) and with GMM in column (ii). There are 94,656 observations at the brand-size-region-month-year level. The sample excludes the months/quarters between June 2008 and May 2009. All regressions include product (brand×size) and period (month or quarter) fixed effects. The elasticity and diversion numbers represent medians among all the brand-size-region-month/quarter-year observations. Standard errors clustered by region and shown in parentheses. Reproduced from Miller and Weinberg (2017).

rather than between the inside goods and the outside good. own-price elasticity of demand for every product for the RCNL-1 specification in 2007.

4 Supply-Side Estimation

4.1 Identification and Estimation

Marginal costs—which typically are unobserved by the econometrician—can be recovered with knowledge of demand and data on prices and quantities. Such inference is the critical step that enables counterfactual analysis in standard analyses of static Nash competition (e.g., Rosse (1970), Berry, Levinsohn, and Pakes (1995); Nevo (2001)). In those settings, the first order conditions of equation (2) allow for direct inferences on marginal costs. With price leadership, more steps are required. We start with a uniqueness assumption:

Assumption 1 (Uniqueness). *For any set of firms that seek to maximize per-period profit conditional on the prices of all other firms, there is a unique vector of prices that satisfy the first order conditions of equation* (2)*.*

Proposition 2 (Identification of Marginal Costs). *Suppose the econometrician has knowledge of the demand system, the identities of the coalition firms (i.e.,* $\mathbb{C}$*), and the super-markup. Then marginal costs are identified.*

The proposition can be verified by tracing the steps needed to recover marginal costs:

1. Infer c_b for each fringe firm $b \notin \mathbb{C}$ directly from the first order conditions of equation (2), using the standard techniques.
2. Obtain $p_b^N = p_b - m$ for each coalition firm $b \in \mathbb{C}$.
3. Compute p_b^N for each fringe firm $b \notin \mathbb{C}$ by solving the first order conditions of equation (2), given the inferred marginal costs c_b and holding the prices of coalition firms fixed at Nash-Bertrand level (i.e., $p_j = p_j^N$ for each $j \in \mathbb{C}$).
4. Evaluate the first order conditions of equation (2) at Nash prices p^N to infer the marginal costs, c_b, of each coalition firm $b \in \mathbb{C}$.

The recovery of marginal costs allows for estimation of the supply-side with the addition of a moment condition. Suppose that one specifies a linear marginal cost function:

$$c_{bt} = w_{bt}'\gamma + \eta_{bt}, \tag{6}$$

where w_{bt} is a vector of observed cost shifters, γ is a vector of parameters, and η_{bt} is an unobserved cost factor that we treat as the structural error term. Let $\theta^S = (m_1, \ldots, m_T, \gamma)'$ be the vector of supply-side parameters to be estimated, where T is the final period observed in the data. Then, under the population moment condition $E[Z'\eta] = 0$, a method-of-moments estimator can be constructed. Let $\eta^*(\tilde{\theta}^S; \hat{\theta}^D)$ be the vector of unobserved costs implied by the candidate parameters $\tilde{\theta}^S$. The estimator is given by

$$\hat{\theta}^S = \arg\min_\theta \eta^*(\theta; \hat{\theta}^D)' Z Z' \eta^*(\theta; \hat{\theta}^D) \tag{7}$$

Excluded instruments must affect markups and be orthogonal to unobserved costs. In implementation, we will use indicator variables for ABI products in each period after the Miller/Coors merger of 2008. With a specification of marginal costs that incorporates (i) an

Table 3: Baseline Supply Estimates

	Parameter	RCNL-1
Super-Markup	m	0.588
		(0.076)
Miller×Post Merger	γ_1	-0.504
		(0.071)
Coors×Post Merger	γ_2	-0.785
		(0.094)
Distance	γ_3	0.153
		(0.062)

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

indicator for MillerCoors post-merger to absorb cost efficiencies, and (ii) time and product fixed effects, the identifying assumption is that the unobserved costs of ABI do not change in the wake of the merger, on average, relative to those of the fringe firms. We concentrate the fixed effects and the marginal cost parameters out of the optimization problem using OLS to reduce the dimensionality of the nonlinear search. We cluster the standard errors at the region level and make an adjustment to account for the incorporation of demand-side estimates (Wooldridge (2010)).

4.2 Point Estimates

Table 3 provides a baseline set of supply-side estimates based on the RCNL-1 demand results. The marginal cost function includes product (i.e., brand×size), period (month or quarter) and region fixed effects. As shown, the super-markup of 0.588 is precisely estimated and the null of static Nash competition is easily rejected. The remaining coefficients take the expected signs. Both Miller and Coors products experience a reduction in marginal costs due to the merger, with the Coors effect being somewhat larger. Marginal costs are estimated to increase in the distance between the brewery and the region, and this provides an additional source of post-merger marginal cost reductions, as the Coors products are produced in the Miller brewing facilities.

As discussed above, identification of the super-markup is based on the magnitude of the ABI price increases that coincide with the MillerCoors merger, relative to import prices.

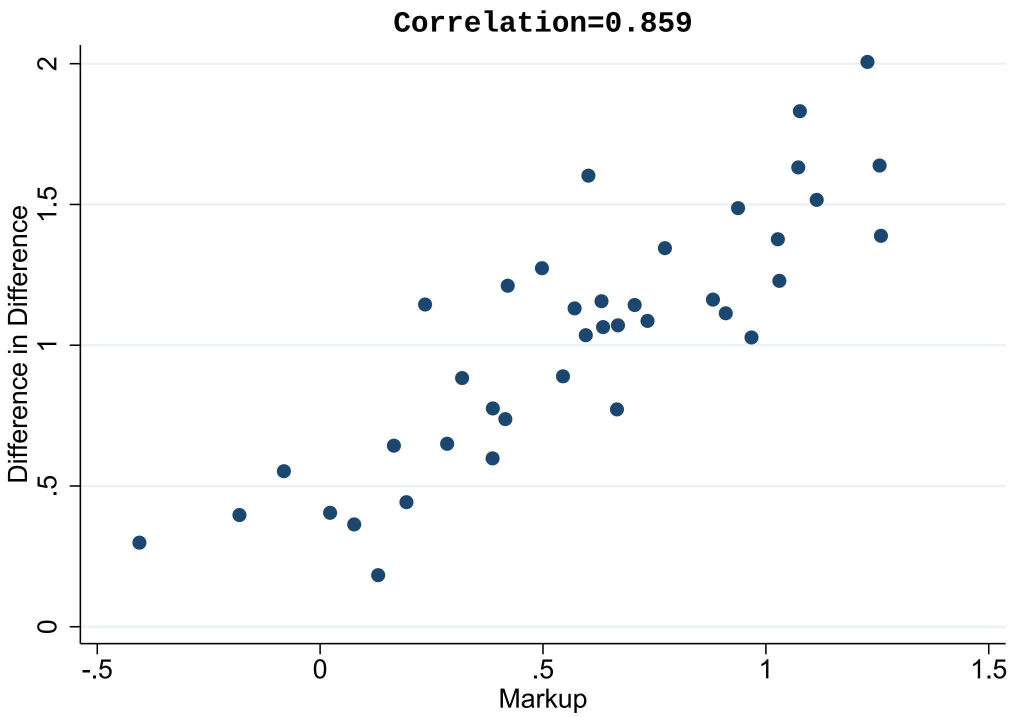

Figure 8: Estimated Super-Markups and Differences-in-Differences Coefficients

Notes: The horizontal axis convey the value of super-markups obtained with region-specific supply-side estimation. The vertical axis conveys the value of a region-specific differences-in-differences regression of ABI prices on an ABI×Post Merger indicator, period fixed effects and product fixed effects. Estimation uses quarterly observations and the supply-side estimation uses the RCNL-2 specification.

To illustrate this, we run a reduced-form regression specified according to

$$\log p_{jrt} \quad = \quad \alpha_r \mathbb{1}\{\text{ABI Post-Merger}\}_t + \phi_{jr} + \tau_{tr} + \epsilon_{jrt}$$

where each α_r is a region-specific price effect, and ϕ_{jr} and τ_{tr} are region-specific product and time fixed effects, respectively. We run the regression using only ABI and import observations, and compare the results to supply-side estimation of the price leadership model conducted on region-by-region subsamples. Figure 8 compares the reduced-form coefficients to the super-markups. There is a tight relationship between the two—regions with larger ABI price increases also have larger estimated super-markups. As shown, the correlation coefficient is 0.859. This region-specific variation is something that we may explore further in future drafts, but for now we set it aside.

Figure 7 shows the price gap between ABI and imports widens over the sample period, suggesting that super-markups may increase over time. Thus, we rerun the supply-side

Table 4: Time Variant Super-Markups

	Parameter	RCNL-1
Super-Markup	m_{2009}	0.386
		(0.064)
Super-Markup	m_{2010}	0.571
		(0.066)
Super-Markup	m_{2011}	0.737
		(0.093)
Super-Markup	m_{2012}	0.925
		(0.080)

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

allowing for time-specific super-markups. We divide the post-merger periods into four "fiscal years" that start in October and end in September, mimicking the timing of price increases suggested by the qualitative evidence.[14] Table 4 shows the results. Consistent with the descriptive price plot, estimates of the super-markup increase over time. For example, with the RCNL-1 specification, the super-markup is \$0.39 in FY2009 and \$0.93 in FY2012. This suggests that the price leadership equilibrium may not have been reached immediately post-merger, possibly due to frictions outside the model. In our remaining analyses, we take the FY2011 estimates as representing the price leadership equilibrium. The IRI dataset does not allow us to rule out that super-markups continue to increase after sample period, but supplementary data that we are examining for use in future drafts suggests that price levels stabilize around the 2011 values, so our present treatment is reasonable.

Figure 9 provides another way to visualize the results, in terms of static Nash reactions. The observed average 2011 prices Bud Light and Miller Lite are shown by the dot labeled "observed price leadership equilibrium." These prices are above the static Nash levels, which we obtain via a counterfactual simulation. We also use counterfactual simulations to construct best-response reaction functions that intersect at the static Nash equilibrium. The dashed 45-degree line that starts at the static Nash equilibrium provides the equilibrium

[14]For example, fiscal year 2009 begins in October 2009 and ends in September 2010. Because we exclude a year's worth of data after the MillerCoors merger, the regression sample includes only four months of fiscal year 2009. Similarly, because the sample ends in December of 2011, we observe only the first three months of fiscal years 2012.

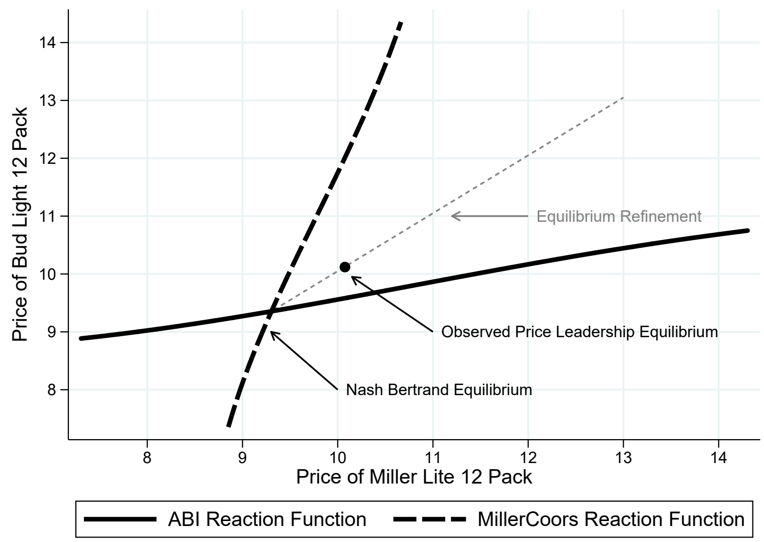

Figure 9: Prices and Best-Response Reaction Functions

Notes: Average 2011 prices are plotted as the dot labeled "observed price leadership equilibrium." The MillerCoors reaction functions is obtained with a series of counterfactual simulations in which we hold *all* ABI prices fixed at levels above or below the Nash-Bertrand equilibrium, and allow all non-ABI prices to equilibrate based on Nash principles. The ABI reaction function is analogous. The intersection of the reaction functions is the Nash-Bertrand equilibrium.

refinement of the price leadership model—ABI can propose prices that fall along this line. Counterfactual simulations can also be used to find the super-markup that maximizes period-specific profit of ABI. When we do so, we find that ABI's FY2011 profit is maximized at super-markup of \$2.78, well above the estimated super-markup on the 45-degree line.[15] This indicates that at least one ICC must bind, constraining the prices set in the market below the level that (otherwise) would be preferred by ABI.

4.3 Identification of the ICCs

Once marginal costs are recovered, a simple counterfactual simulation can be used to recover the super-markups that the leader would set if unconstrained by the need to deter deviation. If these unconstrained super-markups exceed the estimated super-markups (as they do in our application) then to rationalize the estimates it must be that, for at least one $b \in \mathbb{C}$ an

[15]This counterfactual is somewhat sensitive to the value of the nesting parameter, which helps determine the fraction of lost sales that are recaptured within the coalition. Because the instruments we use have only moderate power in the "first-stage" for that parameter (see Appendix D of Miller and Weinberg (2017)), we are exploring supplementary data that may provide more convincing exogenous variation.

incentive compatibility constraint binds:

$$\sum_{s=0}^{\infty} \delta^s \left[\pi(p_b^C, p_{-b}^C) - F(m)\right] = \pi(p_b^D, p_{-b}^C) + \sum_{s=1}^{\infty} \delta^s \pi(p_b^N, p_{-b}^N). \tag{8}$$

Note that the values of the profit functions in equation (8) can be recovered from the estimates ($\pi(p_b^C, p_{-b}^C)$) or with counterfactual simulations ($\pi(p_b^D, p_{-b}^C)$ and $\pi(p_b^N, p_{-b}^N)$). Thus, there exist a series of (δ, F) pairs that solve equation (8). While only joint identification is possible, progress can be made by considering a range of candidate discount factors, e.g., $\tilde{\delta} = (0.70, 0.80, 0.90)$ and obtaining the corresponding frictions, $F(\tilde{\delta})$, in each case.

Building toward that, the left panel of Figure 10 plots the FY2011 prices and shares of Bud Light 12-packs over a range of counterfactual super-markups, relative to Nash Bertrand, under the assumption that MillerCoors accepts the super-markup. Both price leadership equilibrium and deviation are considered. As shown, as the super-markup increases the share of Bud Light decreases roughly linearly. At the estimated super-markup prices are about 10% above static Nash levels and shares are about 10% lower than static Nash. The price that ABI would set if it deviates (from its own proposed super-markup) increases with the super-markup. This is because prices are strategic complements in the model—a higher super-markup implies higher MillerCoors prices against which ABI would compete. The share that ABI captures with defection also increases with the super-markup, for the same reason. Together this shows that the incentives for deviation increase with the level of the super-markup.[16]

The right panel of Figure 10 shows the FY2011 price leadership and deviation profit that ABI captures, relative to Nash-Bertrand, over a range of super-markups; again it is assumed that MillerCoors accepts the super-markup. The price leadership profit function is an inverted-U in the super-markup. At its maximum, which occurs at the super-markup of \$2.78, the envelope theorem applies and ABI is indifferent about arbitrarily small changes in the super-markup. The relative profitability of deviation, however, grows monotonically in the super-markup. This is because higher super-markups correspond to higher MillerCoors prices, which in turn increases the profit of ABI if ABI sets best-response prices.

Figure 11 explores the net present value (NPV) of price leadership and deviation relative to the net present value of static Nash competition, both for ABI (left panel) and MillerCoors (right panel). The deviation values are sensitive to the discount rate, so we show results for $\tilde{\phi} = (0.70, 0.80, 0.90)$.[17] For each of these discount rates the NPV of price

[16] Although we focus here on ABI, qualitatively similar results obtain for MillerCoors.

[17] These are annual discount factors. In the code we convert to monthly or quarterly discount factors as

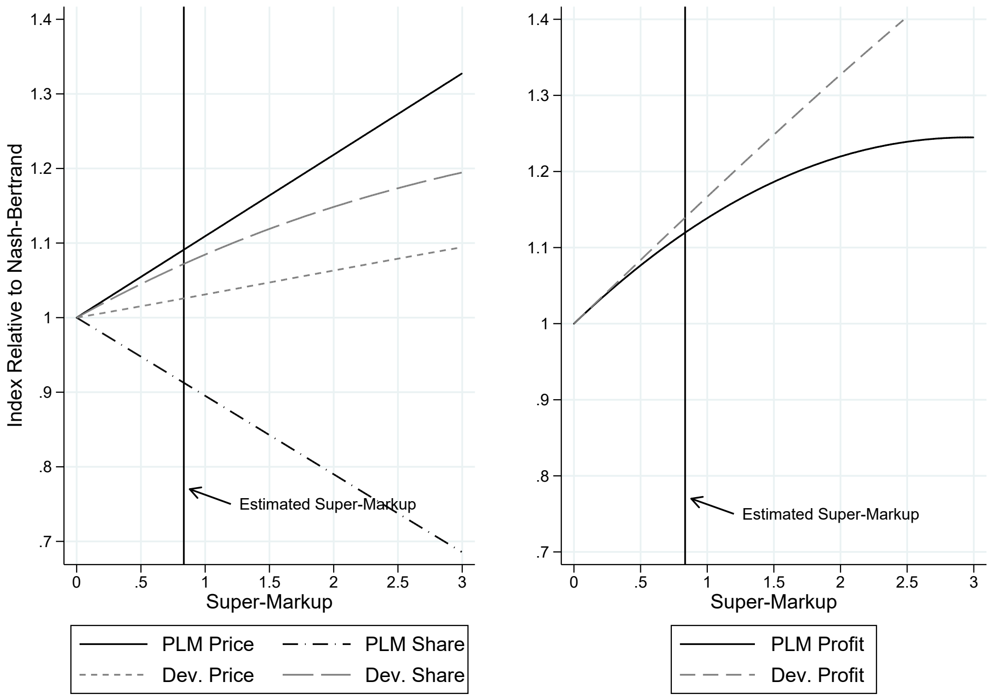

Figure 10: Price Leadership and Deviation at Alternative Super-Markups

leadership well exceeds the NPV of deviation. Thus, if the incentive compatibility constraint binds at observed prices, as in equation (8), then some friction $F(\tilde{\delta}) > 0$ is required.[18]

Another important aspect of Figure 11 is that the gap between the NPV of price leadership and deviation *increases* as the super-markup increases above the estimated level. Thus, for example, a friction that is constant in the super-markup does not rationalize the estimates because as the ICCs would be relaxed for higher super-markups, where the NPV of price leadership also is higher. Given the possible sources of the friction, however, it is plausible that indeed the friction increases with the super-markup, and to make further progress we proceed with the simplest possible specification, $F(\tilde{\delta}) = \varphi m$. We calibrate φ such that exactly one ICC binds for FY2012 (which turns out to be the MillerCoors ICC). With this approach, the discount rate and the slope parameter φ are jointly identified.

appropriate.

[18]Two details on our calculation merit discussion. First, we assume that deviation profits are earned for one year before punishment begins. There is some historical evidence that, in previous bouts of supra-competitive competition, punishment has not been immediate (Greer (1998)). Second, in order trends in consumption away from premium lagers (e.g., Bud Light) we assume that market size decreases after 2011 based on a constant "decay rate" that we estimate from the data over 2005-2011.

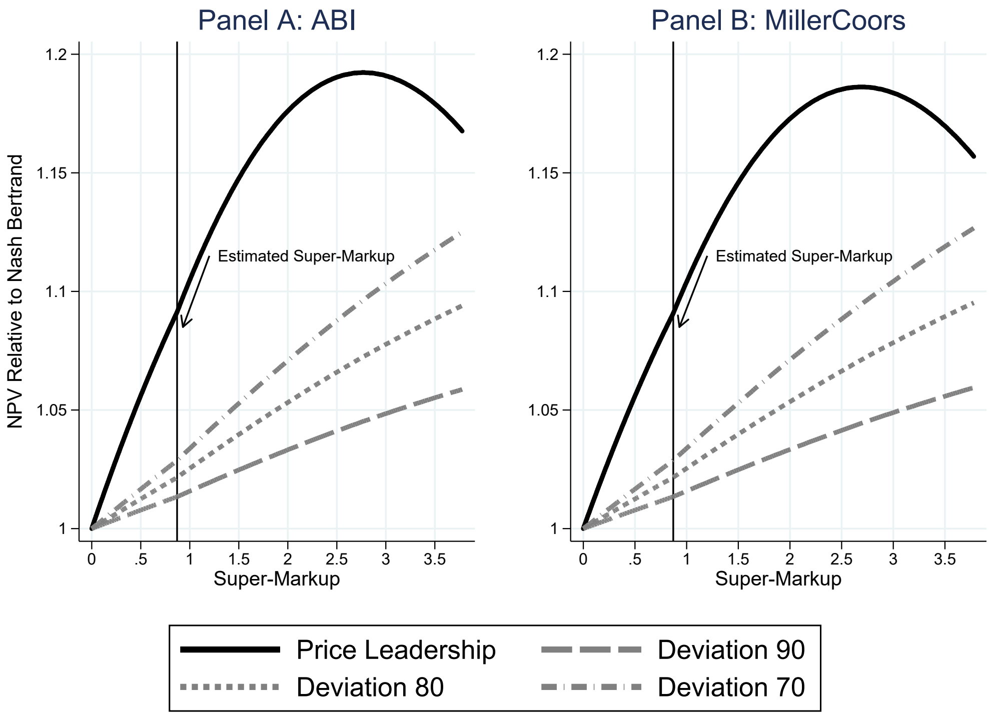

Figure 11: Net Present Value of Price Leadership and Deviation

To illustrate, we calculate for each for the NPV of price leadership less the NPV of deviation for discount factors (0.60, 0.70, 0.80, 0.90). Working from equation (8), this is given by

$$\sum_{s=0}^{\infty} \delta^s \left[\pi(p_b^C, p_{-b}^C) - F\right] - \left(\pi(p_b^D, p_{-b}^C) + \sum_{s=1}^{\infty} \delta^s \pi(p_b^N, p_{-b}^N)\right)$$

If this difference is positive for each firm in the coalition then price leadership is sustainable. Figure 12 plots the value for MillerCoors over a range of super-markups. As shown, price leadership produces greater NPV than deviation for MillerCoors for supermarkups that do not exceed 0.871, which is what we estimate the super-markup to be in FY2012. The NPV of deviation and price leadership equate for $m = 0.871$—indicating the MillerCoors's ICC is what constrains the super-markup. MillerCoors would prefer to deviate for super-markups in excess of 0.871.[19]

[19]We are investigating the source of the "kink" in the ICC at the estimated level.

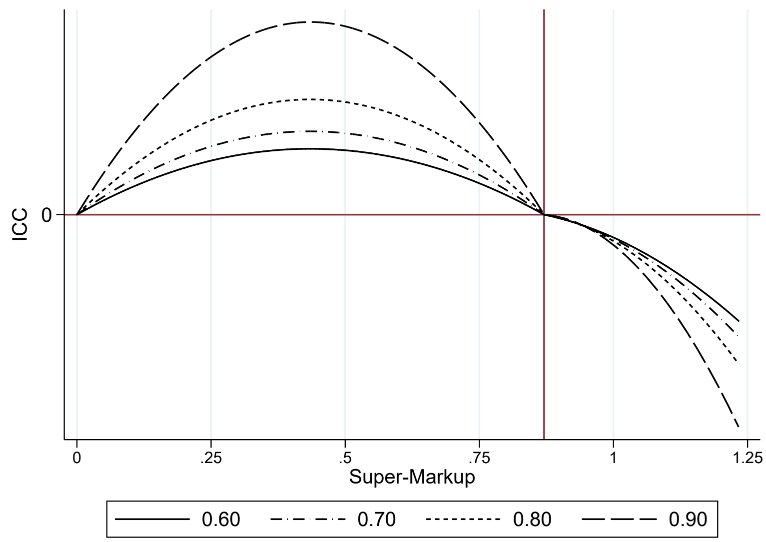

Figure 12: Analysis of MillerCoors ICC for Discount Factors of (0.60, 0.70, 0.80, 0.90)
Notes: The figure plots the value of the incentive compatibility constraints for MillerCoors across a range of counterfactual super-markups. If the value of the ICC is negative then the NPV of deviation exceeds the NPV of price leadership. The ICC equals zero at the estimated super-markup of 0.871, as shown by the intersection of the vertical and horizontal lines.

5 Conclusion

The estimation results obtained thus far unlock a number of novel counterfactuals that we plan to explore in future drafts of this paper. Chief among these is a scenario in which we unwind the MillerCoors joint venture and explore whether an equilibrium exists with prices greater than static Nash.

References

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.

Berry, S. and P. A. Haile (2014). Identification in differentiated products markets using market level data. *Econometrica 82*(5), 1749–1798.

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

Borenstein, S. (1998). Rapid Communication and Price Fixing: The Airline Tariff Publishing Company Case. In J. E. Kwoka and L. J. White (Eds.), *The Antitrust Revolution: The Role of Economics*. Oxford: Oxford University Press.

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 (2017). Learning to Coordinate: A Study in Retail Gasoline. *mimeo*.

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

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

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.

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

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.

Fudenberg, D. and J. Tirole (1991). *Game Theory.* Cambridge, MA: MIT Press.

Gandhi, A. and J.-F. Houde (2016, January). Measuring substitution patterns in differentiated products industries. *Working Paper*.

Greer, D. F. (1998). Beer: Causes of structural change. In *Industry Studies*, 2nd Edition, Larry L. Duetsch (ed), M.E. Sharpe, New York.

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

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

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

Igami, M. and T. Sugaya (2017). Measuring the Incentive to Collude: The Vitamin Cartels, 1990-1999. *mimeo*.

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.

Michel, C. (2016). Identification and estimation of intra-firm and industry competition via ownership change.

Miller, N. H. and M. C. Weinberg (2017). Understanding the Price Effects of the MillerCoors Joint Venture. *forthcoming in Econometrica*.

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

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.

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

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

Sullivan, C. J. (2016). The Ice Cream Split: Empirically Distinguishing Price and Product Space Collusion. *mimeo*.

Sweeting, A. and X. Tao (2017). Dynamic Oligopoly Pricing with Asymmetric Information: Implications for Mergers. *mimeo*.

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