Temporary Price Reductions and Competition: Evidence from the Retail Beer Market

DongIk Kang* and Andrew Usher†

August 19, 2022

Abstract

We find that temporary price reductions (or sales) are an efficient response by colluding firms to privately observed shocks to cost. Using a model of dynamic competition, we show that sales can be permitted in a collusive separating equilibrium in response to private information cost shocks. We estimate the demand and cost primitives of this model, using data from the retail beer market where collusion has previously been documented. We find that profits in the estimated equilibrium with sales is 4% to 13% greater than profits from a counterfactual equilibrium that does not allow for sales. In addition, we find that this counterfactual no-sales equilibrium is less sustainable than the estimated equilibrium in terms of threshold discount factor it supports.

Keywords: Collusion, Sales, Price setting, Dynamic games, Retail beer market

JEL Classification: L10, K11, L13, D4, C73

*Korea Institute of Public Finance. Email: dikang@kipf.re.kr. The views expressed in the paper do not necessarily reflect those of the Korean Institute of Public Finance or the Bank of Canada's Governing Council.

†Bank of Canada. Email: ausher@bankofcanada.ca.

1 Introduction

Collusive arrangements in which firms cooperate to raise prices and increase profits have been observed across many markets and situations. Traditionally, collusive arrangements between firms have often been associated with the rigidity of prices (Carlton 1986, 1989). Prices are less responsive with collusion, the reasoning lies, as to prevent price wars and ensure discipline.[1] Within this framework, price reductions are often interpreted as breakdowns in collusion or price wars, either on or off the equilibrium collusive arrangement.

The presence of temporary price fluctuations, or sales, challenge this interpretation. In sales, we observe frequent price reductions that increase firm market share, precisely what is sought to be avoided by collusion. Furthermore, sales do not resemble price wars or temporary breakdowns in collusion in the mold of Green and Porter (1984) or Rotemberg and Saloner (1986), as they are more often unilateral price reductions followed by a return to regular prices. However, sales are pervasive, even in markets where firms have been known to collude, and strategic utilization of sales would seem to undermine the collusive structure.[2] Yet, the presence and function of sales in collusion are mostly ignored in extant empirical studies of collusion.[3]

In this paper we study the role of sales in maintaining collusion. We hypothesize that sales are an integral part of the tacitly collusive strategy. In particular, we argue that sales are an efficient response of the collusive cartel to privately observed shocks to cost. Privately observed cost shocks may strain the collusive arrangement as they increase the incentive of a firm to deviate and capture the benefits of increased revenue

[1] Athey et al.(2004), Hanazono and Yang (2007) and Athey and Bagwell (2008) formalize this argument.

[2] According to the CPI micro data from the Bureau of Labor Statistics, Kehoe and Midrigan (2015) find that 72% of all price changes are temporary and 50% of all changes revert to their initial regular price in short order. Moreover, using scanner data from Dominick's Finer Foods from 1989 to 1994, they find that 35% of all goods sold were sold on sale.

[3] For example, in papers such as Miller and Weinberg (2017), Michel and Weiergraeber (2018) and Igami and Sugaya (2022), temporary fluctuations in the price disappear with the aggregation of data to higher levels of analysis, without further consideration. Miller and Weinberg note that sales are not important in the increase in collusion.

shares, yet cannot be utilized to determine the collusive price. With a single collusive price, cartels may need to lower markups and profits in order to deter deviation in situations with temporary shocks. Even then, collusion may not be sustainable. However, by allowing for temporary deviations from the collusive price within the collusive structure, i.e. sales, cartels may be better able to sustain collusion and increase collective profits.

To demonstrate how sales may facilitate collusive behavior in the face of private information shocks, we model the firm's pricing decision as one in which firms privately observe cost shocks before deciding on their price. We show that the cartel may be able to accommodate private information cost shocks by pre-setting regular and sales prices and playing a separating equilibrium in which only the low cost firms go on sale. To to insure that price reductions only occur when firms incur shocks to costs, deviations to prices that are neither a predestined regular or sales price are punished by returning to a competitive equilibrium without any coordination. Deviations to sales prices without the presence of cost shocks are prevented by setting prices to satisfy incentive compatibility constraints that deter such deviations.

Ultimately, whether sales improve the efficiency of collusive arrangements is a matter that needs to be settled empirically.[4] Therefore, we estimate our model using the Virtual Stakes approach of Black, Crawford, Lu and White (2004) (henceforth BCLW). The difficulty with trying to measure collusion in a dynamic setting is the large number of possible equilibrium outcomes. The well known folk theorems inherent in dynamic games of firm competition forces the researcher to choose one out of many possible equilibrium outcomes. In practice, this is limiting as it typically imposes a specific model of static competitive interaction such as Nash-Bertrand competition or perfect collusion.

BCLW show that by limiting the set of possible equilibria to those that are Pareto

[4]For example, Athey et al. (2004) show that in certain situations a rigid collusive pricing scheme is optimal, whereas allowing for deviations are sometimes necessary with large private information cost shocks.

efficient, the Virtual Stakes approach they introduce can measure competition in a manner consistent with a large range of competitive processes. They show that the outcome of the dynamic game can be captured as the outcome of a static social planner's problem with competitor profits as constraints. This allows us to estimate the model from a large set of possible outcomes and ensures that our results are robust, as long as Pareto efficiency is an expected quality of the equilibrium.[5] Furthermore, BCLW show that their approach can be interpreted as one in which firms act as if they have "virtual stakes" in the profits of their competitors, providing an intuitive interpretation of their parameters governing the collusive arrangement.[6] To the best of our knowledge, we are one of the first papers to utilize the Virtual Stake approach to estimate dynamic models of collusion, along with Sullivan (2017).

To test whether sales indeed increase cartel profits and help sustain collusion, we estimate our model using scanner data of retail beers. We focus on the American retail beer industry because it is a well defined market with standardized products across markets in which a few large producers such as Anheuser-Busch Inbev (ABI), SABMiller PLC (Miller) and Molson Coors Brewing (Coors) dominate the market and have been suspected of tacit collusion. A recent treatment of this industry by Miller and Weinberg (2017) has illustrated the possibility of collusion intensifying after the merger between two of the three largest firms, Miller and Coors. Furthermore, sales are frequent in this market and privately observed cost shocks are pervasive due to the fact that cost shocks occurring at the local retailer level cannot be realistically

[5] As the set of potential equilibria are limited to those that are Pareto efficient, this method is more suitable for analyzing the characteristics and the dynamics of a collusive arrangement, as opposed to testing for the existence of collusion. This aligns with our purpose, as the objective of our paper is to study the dynamics of collusion and the role of sales in a setting where collusion has been previously documented.

[6] An alternative approach attempts to measure competition by estimating "conduct parameters" that, loosely speaking, governs how the effect of the products own price on the demand for other products influences the determination (through the first-order condition) of the product's own price. When these conduct parameters take on certain values, they correspond exactly to perfect competition, monopoly or Cournot competition models. However, outside of these specific values and models, the conduct parameter approach may fail to measure market power accurately (for a more detailed discussion see, Corts 1999, Fan and Sullivan, 2020).

observed and kept track of by the headquarter, the level at which the general pricing strategy and thus collusion is likely to occur.

Our estimation results are consistent with collusive behavior by ABI, Miller, and Coors. We find that each firm in the cartel places considerable weight on the profit levels of their collusion partners and values ensuring that their collusion partners do not deviate from their agreement. Furthermore, we find evidence that suggest the intensity of collusion increases after the merger between Miller and Coors. The estimated markups of both ABI and the joint MillerCoors venture both increase significantly after the merger, whereas the markups of other players in the industry thought to be outside the cartel do not show meaningful increases. In addition, the shadow marginal value of maintaining the collusive agreement decreases, suggesting that the merger facilitated collusion, as previous research suggests. All in all, our results align very closely with the findings of Miller and Weinberg (2017) as well as the literature on mergers, which we believe supports the credibility of our results.

More importantly, our results show that temporary price reductions play an important role in facilitating collusion. We find that allowing for sales in equilibrium both increases the profits of firms and the sustainability of the collusive arrangement. When sales are allowed, firm profits are greater by 4% to 13% compared to the profits from a counterfactual equilibrium in which no sales were allowed. Furthermore, we find that the minimum threshold discount factor required to sustain the collusive arrangement is lower in the equilibrium with sales than in the counterfactual equilibrium, implying that collusion is easier to sustain. The difference can be attributed to the fact that both the defection profit from unilateral deviation is lower and the collusive profit is higher in the equilibrium with sales. These results suggest that temporary price reductions in the form of sales serve an important role in collusion between firms.

Our findings suggest that collusive arrangements may be more flexible in their nature than conventional wisdom may suggest. While our paper considers only privately observed cost shocks, the broader implications of our finding suggest that other factors that are usually associated with impeding collusive behavior, such as imperfect monitoring or private demand shocks, may also be accommodated by cartels through more flexible collusive arrangements. Such arrangements could facilitate collusion by increasing the efficiency of cartels. Furthermore, these arrangements, as demonstrated by sales, need not be overly complex or difficult to observe, potentially widening the range of competitive environments in which we expect collusion to take place.

A large literature studies collusion between firms modeled as the result of a super game. Papers by Green and Porter (1984) and Abreu, Pearce, and Stacchetti (1986) famously examined collusion in the cases where firm actions were not perfectly observed. Rotemberg and Saloner (1986) study how the changes in firms' incentives due to demand fluctuations affect collusive arrangements. A long list ensuing of papers further study how collusion is effected under various market conditions.

Our paper is especially closely related to the literature that studies how private information affects collusion. Closely related to our model are the series of papers by Athey and Bagwell (2001, 2008) and Athey, Bagwell and Sanchirico (2004). Athey and Bagwell (2001) study optimal collusion in an environment with privately observed i.i.d. cost shocks. Athey, Bagwell and Sanchirico (2004) show that price rigidity may be an outcome of collusion in such environments. They also demonstrate the possibility that "escape clauses" may be needed under certain conditions, where firms are allowed to temporarily cut their prices. This mechanism is reminiscent of sales in our framework. Athey and Bagwell (2008) extends the analysis to persistent costs shocks. In other work, Hazano and Yang (2007) study price ridigity with private signals. Sweeting et al. (2022) study the effect of signaling on mergers in an environment with asymmetric information. All of the aforementioned papers are theoretical in their nature and, to the best of our knowledge, we are the first paper to estimate a fully dynamic model of collusion with private information.

Early empirical research such as Porter (1983) and Bresnahan (1987) explicitly estimate market outcomes as the result of a game between market participants. In this vein, various recent studies have examined collusive arrangements by estimating collusive equilibria in several industries. For example, Ciliberto and Williams (2014) study collusion in the airline industry, Miller and Weinberg (2017) study the effect of mergers on collusion in the market for beer, Igami and Sugaya (2022) study cartels in the vitamin market, and Sullivan (2017) models collusion in the choice of product mix in the premium ice cream market. In each of these papers, as do we, the authors do not directly observe collusion or the primitives of the supply side and therefore must make substantive assumptions on the nature of competition, which they do so using dynamic games.

While the empirical literature on collusion has largely overlooked temporary price reductions, Nava and Schiraldi (2014) present a model in which firms must periodically reduce prices in order to sustain collusion when goods are storable and the market is large. In their model, sales foster collusion by intensifying the intertemporal links in consumer behavior. Roos and Smirnov (2020) examine optimal cartel behavior with imperfectly attentive consumers and show that periodic low prices can increases the sustainability of collusion. In our model, sales facilitate collusion by allowing firms to exploit temporary cost shocks to maximize their current-period profits.

The remainder of the paper is organized as follows. In Section 2, we present our model of collusion and sales. We also describe the Virtual Stakes approach that we utilize to estimate the model and discuss how the degree of competition is captured using this approach. Section 3 details our data set, estimation procedure, and estimation results. In Section 4, we compare our estimation results with a counterfactual equilibrium outcome in which there are no sales. In Section 5, we conclude.

2 Model

In this section we present our model of dynamic price setting among competing firms. Although the model we eventually estimate will be of multi-market, multi-product firms, we abstract from these considerations in this section for the purpose of exposition.[7]

2.1 A Model of Tacit Collusion

We consider a dynamic price setting game in which each producer chooses their price every period. There is a finite number F firms competing in the industry. Each period, firms choose the price of their product. The prices chosen by each firm are observed by all firms. Firms choose their price to maximize their expected profit,

$$E_t \sum_{\tau=0}^{\infty} \beta^{\tau} \pi_{ft+\tau}$$

where $\pi_{ft+\tau}$ is the period $t+\tau$ expected profit of the firm and can be expressed as,

$$E_t \pi_{ft+\tau} = E[(p_{ft+\tau} - c_{ft+\tau}) q_f(p_{ft+\tau}, P_{-ft+\tau}; \Theta_{t+\tau}) | \Theta_t]. \tag{1}$$

The variable p_{ft} denotes the period t price of firm f, P_{-ft} is the vector of the prices of other firms, c_{ft} is the marginal cost, and Θ_t contains all other payoff relevant information and is revealed at the beginning of each period. Θ_t is common knowledge to all firms. The demand function q_f of firm f is a function of its own price p_{ft} and other firms' prices P_{-ft}, conditional on Θ_t. The expectation in (1) is over the actions (prices) of other firms as well as future realizations of costs and $\Theta_{t+\tau}$. Although the firms can freely change their price every period, the price that firms choose in the current period can affect their future profits through its influence on the action of

[7]The extension from the setup in this section to multi-product, multi-market firms is mostly straightforward, although some issues regarding the level of the decision making unit of the firm are discussed in detail in Section 3.3 when we take our model to the data.

competing firms.

We assume the following cost structure for all firms. At the beginning of each period the cost structure of every firm is revealed. After the costs are revealed, each firm is subject to a cost shock with probability α_{ft} where they temporarily face a lower marginal cost. The "regular" and "shocked" costs $\{c_{ft}^R, c_{ft}^S\}_{f\in F}$, as well as the probabilities $\{\alpha_{ft}\}_{f\in F}$ are common knowledge and thus all firms know the potential costs of their competitors. However, whether or not a firms received the cost shock is private information. In other words, the realization of the actual cost of the firm $c_{ft} \in \{c_{ft}^R, c_{ft}^S\}$ is only known to the firm itself. The cost structure implies that firms have incomplete information about the state.

In the competitive equilibrium of this game, the firms would choose prices that maximized their own static profit each period subject to the other firms in the industry doing the same. However, in many cases, firms have the incentive to collude with their competitors to collectively charge higher than competitive prices to guarantee higher returns. In such a case, the literature has demonstrated that collusion can be maintained through tacit arrangements where deviations are deterred by punishment schemes that lower prices and profits when they are detected.

With privately observed cost shocks, the collusive equilibrium can theoretically be either pooling or separating.[8] However, in a pooling equilibrium all firms will charge only the collusive regular price, indicating that the sheer presence of sales rules out this equilibrium. In the separating equilibrium, firms with the regular cost will charge the collusive regular price and firms with the low cost will charge the sales price, which aligns with our observations of frequent sales. Therefore, in what follows we describe the potential competitive equilibrium and the separating equilibrium of this model.

We choose to use cost shocks to drive sales, and simplify the structure of the costs to capture the the fundamental tradeoffs that cartels face with private information

[8]We do not consider hybrid equilibria in this section due to its complex nature and the burden it levys on firms to maintain such equilibria, especially with multi-product and multi-market firms. In a previous version of this paper, we explored hybrid equilibria and its implications.

and simultaneously replicate the sales feature of the data, in a simple and tractable manner. This is critical to our ability to estimate the model. However, we also believe that our model is a good representation of the true competitive environment and is well supported by the previous findings and descriptions regarding sales and firm behavior. For example, Eichenbaum et al. (2011) finds that almost all sales are accompanied by price changes, which supports our choice to model sales as arising from cost shocks in equilibrium. Anderson et al. (2017) report that sales often require manufacturer - retailer coordination and the broad outlines of how sales are to be conducted is partly known in advance, whereas whether to go on sale itself is determined by the retail stores, suggesting that the general cost structure can be known by the cartel but the actual realizations are unobservable in advance.

2.1.1 The Competitive Equilibrium

Without cooperation between firms, the firms will play the Nash-Bertrand equilibrium of the dynamic game in which the competitive equilibrium of the static game is played each period. The Nash-Bertrand equilibrium is often considered as the default equilibrium in a dynamic game among firms.

First, we lay out some notation for clarity of exposition. Let ω be a vector with F elements where $\omega = [\omega_1 \omega_2 ... \omega_F]$ and each $\omega_f \in \{0, 1\}$ for all $f \in F$. The ω_f's indicate the idiosyncratic state of each firm, where if $\omega_f = 1$ then firm f has the regular cost c_f^R and if $\omega_f = 0$ then firm f has the shocked cost c_f^S. Furthermore, denote ω_{-f} as the vector ω without the fth element. For example $\omega_{-1} = [\omega_2 \omega_3 ... \omega_F]$ is a $(F-1) \times 1$ vector. Let Ω_{-f} be the set of all possible outcomes of ω_{-f}.

Then, given cost c_{ft}^i where $i \in \{R, S\}$, the firms would choose their prices in the Nash-Bertrand equilibrium to solve,

$$\pi_{ft}^{iB} = \max_p \sum_{\omega_{-f} \in \Omega_{-f}} prob_t(\omega_{-f})(p - c_{ft}^i) q_f(p, P_{-ft}(\omega_{-f}); \Theta_t) \tag{2}$$

where $prob_t(\omega_{-f})$ is the probability of state ω_{-f} occurring in period t and $P_{-ft}(\omega_{-f})$ is the $F-1$ vector of prices of the other firms in period t in state ω_{-f}. Note that the period t expected payoff of firm f in the Nash-Bertrand equilibrium is then,

$$\pi^B_{ft} = (1-\alpha_{ft})\pi^{RB}_{ft} + \alpha_{ft}\pi^{SB}_{ft}, \tag{3}$$

and the Nash-Bertrand Equilibrium can be defined as follows.

The Nash-Bertrand Equilibrium
Each period, each firm f chooses price p^B_{ft} such that given the vector of prices P_{-ft} and cost c^i_{ft} where $i \in \{R, S\}$, p^B_{ft} solves (2).

2.1.2 The Collusive Separating Equilibrium

Alternatively, it may be of interest to firms to agree to the collusive arrangement and charge a higher price to increase total profits. If firms are sufficiently patient, these types of arrangements can usually be sustained by a threat of punishment when deviation is detected. The punishment will decrease the profits of the deviating firm (and also possibly other firms) so the firms do not have the incentive to deviate from equilibrium play. Therefore, we consider a collusive equilibrium in which firms utilize the Nash-Bertrand equilibrium as a punishment stage when a deviation is detected. While, punishment schemes other than the Nash-Bertrand can also be considered, such schemes entail an additional no-deviation constraint to those we shall consider. Thus, much of the literature often assumes punishment schemes that revert to Nash-Bertrand.[9]

As previously stated, even a naive observation of the data shows that firms sometimes deviate from their normal price and charge a considerably lower sales price.

[9]See, Asker and Nocke (2021) for a survey.

Therefore, a full description of the collusive arrangement between firms should provide an explanation for such behavior within the competitive framework. We model such sales prices as a part of a separating equilibrium where "excused" temporary deviations from the collusive regular price are allowed when incentives to deviate are temporarily high. These excused deviations can serve to either facilitate and sustain the collusive agreement or to increases profits, or both. In this paper, we model increased deviation pressures through temporarily low prices. Thus, in the separating collusive equilibrium we consider, the firms agree to adhere to the collusive regular price but are allowed to temporarily charge a lower sales price if their costs are low.

Now, consider the following collusive separating equilibrium. At the beginning of each period, the firms negotiate (outside the model) a set of collusive prices $(p_t^R, p_t^S) = (\{p_{ft}^R\}_{f\in F}, \{p_{ft}^S\}_{f\in F})$ for each to set as their regular and sales prices. The set of collusive prices are known to all firms. The regular price is the normal price charged by the firm for its product and the sales price is a lower price the firm may instead charge when they have low costs. If a firm deviates, all firms revert to their Nash-Bertrand strategies forever as punishment (grim-trigger). The firms observe their realized costs before their decision to deviate. We will assume that firms determine the collusive price before they observe the realization of the shock. We note that for the purposes of this section, we treat the collusive prices as given and study the firm's decision problem of setting their actual price. In Section 2.2 we describe in detail our treatment of the collusive prices, their interpretation, and how we proceed to estimate them. For now, we assume that the collusive regular price is always weakly greater than the Nash-Bertrand equilibrium price.

For the firms to actually charge the agreed upon collusive prices, the expected gain from deviating to a non-collusive price must be smaller than their expected loss from playing the Nash-Bertrand equilibrium in the subsequent periods. Let $\pi_{ft}^i(p)$ denote

$$\pi_{ft}^i(p) = \sum_{\omega_{-f}\in\Omega_{-f}} prob_t(\omega_{-f})(p - c_{ft}^i)q_f(p, P_{-ft}(\omega_{-f}); \Theta_t). \tag{4}$$

Then the following incentive compatibility constraints must hold for $i \in \{R, S\}$ for the firms to not deviate from the separating equilibrium:

$$(5) \qquad E_t \sum_{\tau=1}^{\infty} \beta^{\tau} \left[\pi_{ft+\tau}(p^R_{ft+\tau}, p^S_{ft+\tau}) - \pi^B_{ft+\tau} \right] \geq \pi^i_{ft}(p) - \pi^i_{ft}(p^i_{ft}) \qquad \text{for any } p$$

where π^B_{ft} is described by (3) and $\pi_{ft}(p^R_{ft}, p^S_{ft})$ equals

$$(6) \qquad \pi_{ft}(p^R_{ft}, p^S_{ft}) = (1 - \alpha_{ft})\pi^R_{ft}(p^R_{ft}) + \alpha_{ft}\pi^S_{ft}(p^S_{ft}).$$

Equation (5) is the no-deviation condition that states that the present value of the loss from reverting back to the Nash-Bertrand equilibrium from period $t + 1$ onward is larger than the profit that can be gained in the current period from deviation for costs c^R_{ft} and c^S_{ft} respectively.[10]

Additionally, note that because firms have incomplete information about the realization of the costs, firms will be unable to differentiate between firms that charge a low price because they have low costs, and those that have regular costs but are deviating from equilibrium play. This poses a threat to the sustainability of the collusive arrangement because firms often have the incentive to charge lower costs to capture market share. Firms will often have the incentive to act as if they have the low cost and charge the agreed upon collusive sales price p^S_{ft}, even when they actually have the regular cost. Thus, the cartel must ensure that the firms with the regular cost do not have the incentive to charge the sales price even without the threat of detection, and vice versa. Therefore, the following incentive compatibility constraints must hold

[10]In equation (5), we have assumed that firms play a grim trigger strategy, where firms revert to Nash-Bertrand indefinitely once they detect deviation. While we assumed that firms play a grim trigger strategy for convenience of exposition, our setup and estimation is consistent with any punishment strategy in which firms revert to the Nash-Bertrand equilibrium for an extended period of time and return to the collusive arrangement. This is due to the fact that the punishment scheme does not depend on current period price at time of potential deviation.

for all firms $f \in F$:

$$(7)\quad (p^R_{ft} - c^R_{ft})\Big\{ \sum_{\omega_{-f}\in\Omega_{-f}} prob_t(\omega_{-f}) q_f(p^R_{ft}, P_{-ft}(\omega_{-f}); \Theta_t)\Big\}$$

$$\geq \quad (p^S_{ft} - c^R_{ft})\{ \sum_{\omega_{-f}\in\Omega_{-f}} prob_t(\omega_{-f}) q_f(p^S_{ft}, P_{-ft}(\omega_{-f}); \Theta_t)\},$$

$$(8)\quad (p^S_{ft} - c^S_{ft})\Big\{ \sum_{\omega_{-f}\in\Omega_{-f}} prob_t(\omega_{-f}) q_f(p^S_{ft}, P_{-ft}(\omega_{-f}); \Theta_t)\Big\}$$

$$\geq \quad (p^R_{ft} - c^S_{ft})\{ \sum_{\omega_{-f}\in\Omega_{-f}} prob_t(\omega_{-f}) q_f(p^R_{ft}, P_{-ft}(\omega_{-f}); \Theta_t)\}.$$

We follow Athey and Bagwell(2004) and label the constraints (7) and (8) as the *on-schedule* no-deviation constraints and the incentive compatibility constraint in (5) as the *off-schedule* deviation constraint. Condition (5) assures that firms will not deviate to any price other than p^R_{ft} or p^S_{ft} for fear of punishment. Condition (7) implies that firms with the regular cost will not deviate to the sales price, and Condition (8) implies that firms with the low cost will not deviate to the regular price.

As first discussed in Athey, Bagwell and Sanchirico (2004), privately observed temporary cost shocks may increase the incentive of firms to deviate from the collusive arrangement, making collusion unsustainable.[11] In addition, even if collusion were still possible, firms may simply wish to accommodate cost shocks into their collusive arrangement by allowing firms that receive the shocks to temporarily charge lower prices, increasing profitability. We accommodate such cost shocks by playing a separating equilibrium in which only firms who receive the low cost shocks charge a temporarily low sales price, and all other firms charge the collusive regular price. We note that the sales prices charged in our setting is similar to the "*escape clauses*" in Athey, Bagwell and Sanchirico that allow temporary deviation for large cost shocks

[11]Ideally, our setting would include all situations in which a firm's incentive to deviate experiences temporary fluctuations. For example, temporary changes to the price elasticity of demand may cause firms to wish (more than usual) to lower their prices to capture larger market shares. However, including all such possibilities is infeasible and thus we focus on cost to be parsimonious and because we believe its the most representative of such incentives.

in order to sustain collusion.

We define the collusive equilibrium as follows.

The Collusive Separating Equilibrium

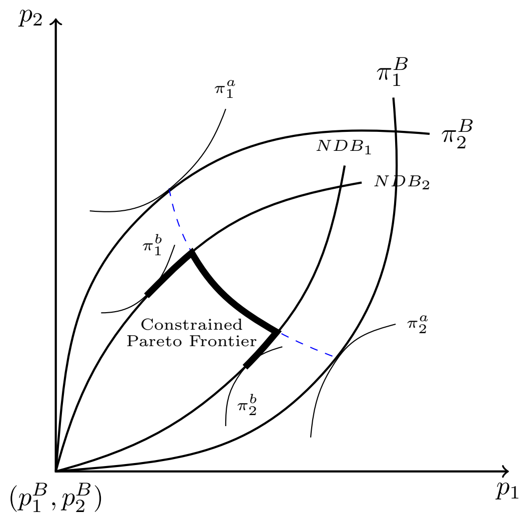
Let $(p_t^R, p_t^S) = (\{p_{ft}^R\}_{f \in F}, \{p_{ft}^S\}_{f \in F})$ *denote the given collusive prices in period* t. *A Collusive Separating Equilibrium exists if* (p_t^R, p_t^S) *satisfies* (5) *for* $i \in \{R, S\}$ *and* (7), (8) *for all* $f \in F$. *Let* p_{fs} *denote the actual price charged by firm* f *in period* s. *In this equilibrium, each firm* f *chooses price* p_{ft}^R *if the realized cost is* c_{ft}^R *and price* p_{ft}^S *if the realized cost is* c_{ft}^S, *if* $p_{gs} \in \{p_{gs}^R, p_{gs}^S\}$ *for all* $g \in F$ *and* $s < t$. *Otherwise, each firm chooses the price* p_{ft}^B.

2.2 Virtual Stakes Approach

In the collusive separating equilibrium in Section 2.1.2, we leave the determination of the collusive prices (p_t^R, p_t^S) outside the model and treat it as given at the beginning of the period. We do so to focus on the equilibrium actions of the firms given the parameters of the collusive agreement.

In this section, we describe how we measure collusion among firms, and how the collusive price is determined in our methodology. As previously noted, we adopt the Virtual Stakes approach of BCLW to measure the degree of collusion among firms. We limit the set of equilibrium to those that are Pareto efficient, which allows us to capture the equilibrium as an outcome of a social planner's problem with Pareto weights and current period constraints. Figure 1 shows Figure 4(A) in BCLW, which illustrates a Pareto efficient set in an environment with two firms. The thick black line represents the Pareto efficient set on which the potential equilibrium may lie. The depicted Pareto efficient set includes points on the unconstrained Pareto Frontier (the dashed blue line) in between the no-deviation boundaries labeled NDB_1 and NDB_2 respectively. The no-deviation boundaries show the potential prices that would exactly satisfy the conditions under which each firm would not have the incentive to

Figure 1: Pareto Optimal Set



Note: This figure shows Figure 4(A) of BCLW, that depicts the constrained Pareto optimal set with binding no-deviation constraints with two competing firms.

deviate. Thus, the equilibrium prices must lie inside the two no-deviation constraints. The rest of the Pareto efficient set lies on these no-deviation constraints between the tangent point to the iso-profit curves of each firm and the unconstrained Pareto Frontier. The figure highlights that a static representation of the dynamic game can successfully capture any equilibrium outcome on the depicted Pareto set as a result of the full dynamic game.

BCLW further demonstrate that the static social planner's problem is equivalent to their Virtual Stakes representation, in which firms act as if they have "virtual stakes" in their competitors, i.e., partial ownership of their competitor's profits. This duality in the representation of a dynamic equilibrium provides a natural and economic interpretation of the estimated parameters governing collusion. In our setting, the Virtual Stakes approach entails expressing the collusive equilibrium outcome as the solution to the following problem for each firm $f \in F$ as follows:

(9)

$$\max_{p^R_{ft}, p^S_{ft}} \quad \pi_{ft}(p^R_{ft}, p^S_{ft}; P_{-ft}) + \sum_{g \in F, g \neq f} \theta_{fgt} \pi_{gt}(p^R_{ft}, p^S_{ft}; P_{-ft})$$

$s.t.$

$$E_t \sum_{\tau=1}^{\infty} \beta^{\tau} (\pi_{gt+\tau}(p^R_{gt+\tau}, p^S_{gt+\tau}; P_{-gt+\tau}) - \pi^B_{gt+\tau}) \geq \pi^R_{gt}(p^{*R}_{gt}; P_{-gt}) - \pi^R_{gt}(p^R_{gt}; P_{-gt}) \qquad \forall g$$

$$E_t \sum_{\tau=1}^{\infty} \beta^{\tau} (\pi_{gt+\tau}(p^R_{gt+\tau}, p^S_{gt+\tau}; P_{-gt+\tau}) - \pi^B_{gt+\tau}) \geq \pi^S_{it}(p^{*S}_{gt}; P_{-gt}) - \pi^S_{gt}(p^S_{gt}; P_{-gt}) \qquad \forall g$$

$$\pi^R_{gt}(p^R_{gt}; P_{-gt}) \geq \pi^R_{gt}(p^S_{gt}; P_{-gt}) \qquad \forall g$$

$$\pi^S_{gt}(p^S_{gt}; P_{-gt}) \geq \pi^S_{gt}(p^R_{gt}; P_{-gt}) \qquad \forall g$$

where P_{-ft} denotes the vector of equilibrium regular and sales prices of all firms except firm f. We can re-state the above problem using Lagrange multipliers as the static objective function for the firm f at time t as:

(10)

$$\max_{p^R_{ft}, p^S_{ft}} \quad \pi_{ft}(p^R_{ft}, p^S_{ft}; P_{-ft}) + \sum_{g \in F, g \neq f} \theta_{fgt} \pi_{gt}(p^R_{ft}, p^S_{ft}; P_{-ft})$$
$$+ \sum_{g \in F} \lambda^R_{fgt} \Big[E_t \sum_{\tau=1}^{\infty} \beta^{\tau} (\pi_{gt+\tau}(p^R_{gt+\tau}, p^S_{gt+\tau}; P_{-gt+\tau}) - \pi^B_{gt+\tau}) - (\pi^R_{gt}(p^{*R}_{gt}; P_{-gt}) - \pi^R_{gt}(p^R_{gt}; P_{-gt})) \Big]$$
$$+ \sum_{g \in F} \lambda^S_{fgt} \Big[E_t \sum_{\tau=1}^{\infty} \beta^{\tau} (\pi_{gt+\tau}(p^R_{gt+\tau}, p^S_{gt+\tau}; P_{-gt+\tau}) - \pi^B_{gt+\tau}) - (\pi^S_{it}(p^{*S}_{gt}; P_{-gt}) - \pi^S_{gt}(p^S_{gt}; P_{-gt})) \Big]$$
$$+ \sum_{g \in F} \mu^R_{fgt} \Big[\pi^R_{gt}(p^R_{gt}; P_{-gt}) - \pi^R_{gt}(p^S_{gt}; P_{-gt}) \Big] + \sum_{g \in F} \mu^S_{fgt} \Big[\pi^S_{gt}(p^S_{gt}; P_{-gt}) - \pi^S_{gt}(p^R_{gt}; P_{-gt}) \Big]$$

for each firm $f \in F$. The prices p^{*R}_{gt} and p^{*S}_{gt} are the prices that maximize static profits $\pi^R_{gt}(\cdot)$ and $\pi^S_{gt}(\cdot)$ respectively. Note that the choice variables p^R_{ft}, p^S_{ft} are reflected in the vector of prices P_{-gt} in the no-deviation constraints. In the objective functions of (9) and (10), we have assumed that the timing of the choice of the collusive prices occurs after the cost structure is revealed but before the actual realization of the shocks. Thus the firms do not know which of regular or shocked cost they will have when making this choice.

The Virtual Stakes representation highlights that with collusion, producer actions can be interpreted as if firms held stakes (θ_{fgt}) in the profits of each firm. In addition, the multipliers $\{\lambda^N_{fgt}, \lambda^S_{fgt}, \mu^R_{fgt}, \mu^S_{fgt}\}$ of the no-deviation constraints can be interpreted as a measure of the internalized value of keeping each rival firm from deviating and thus maintaining the collusive arrangement. Intuitively producers behave as if they care about the impact of their actions on (i) their own profit, (ii) the profits of their competitors and (iii) maintaining the collusive equilibrium. The Virtual Stakes approach captures these motives in a succinct way, and provide a way to measure the value of each of these components. Thus a higher value of θ_{fgt} can be interpreted as firm f putting more weight on the profit level of firm g. Similarly, a higher value of λ^i_{fgt} would indicate that the marginal cost of insuring that firm g does not deviate from the collusive strategy is greater for firm f.

BCLW demonstrate that the first-order conditions and hence the equilibrium prices obtained from (10) are equivalent to those from a Pareto optimal social planner's problem where the weights for each firm θ_{fgt} are equivalent to the Lagrange multipliers for the constraint that each firm's profits are equal to or greater than the profit specified by the Pareto optimal allocation. The no-deviation constraints restrict the set of Pareto optimal equilibria to those that can be sustained by the punishment strategy. The multipliers $\{\lambda^R_{fgt}, \lambda^S_{fgt}, \mu^R_{fgt}, \mu^S_{fgt}\}$ indicate the degree to which each constraint is binding. Together, these multipliers determine where within the set of possible Pareto optimal outcomes the actual equilibrium lies. This interpretation provides a deeper understanding of the economic foundations of the estimated parameters.

Furthermore, the equivalence between the Virtual Stakes equilibrium and Pareto efficient equilibria further restricts the estimated parameters such that the Virtual

Stakes parameters of each firm only differ by a scale factor, allowing for more efficient estimation. This can be seen by considering the Pareto efficient social planner's problem that the firm would face. The planner would face a problem similar to (10) with the same objective and constraints, but instead of choosing only one price the planner would choose the price of all firms. Then, the Lagrange multipliers from the planner's first-order conditions will only be identified up to scale, meaning they can be re-scaled freely and have identical implications for each firm. This places cross-restrictions on the Virtual Stakes parameters such that the parameters of each firm differ by only a scale factor.

Our approach is reminiscent of the conduct parameter models in the spirit of Bresnahan (1981) and our interpretations of the parameters are somewhat similar. However, the profit-weight conduct parameter models are subject to criticism that the approach arbitrary and impose behavior on firms that are not well supported by economic theory. Notably, it is unclear as to why a firm would include competitor profits in their policy functions outside the limiting cases that conform exactly to competitive, Cournot, or monopolistic outcomes. In contrast, the Virtual Stakes approach of BCLW provide a interpretation of the parameters well grounded in economic theory. They show that the static solution concept of Virtual Stakes equilibria can be utilized to represent the dynamic behavior of competition among firms for the type of supergames considered in this paper.

Lastly, we discuss some aspects of our modeling choices. First, we note that the Virtual Stakes approach allows for a wide range of punishment schemes, consistent with the approach. In equation (5) we present a grim trigger strategy for the sake of exposition, but the methodology is consistent with a large set of alternative punishment schemes. Any punishment scheme, in which the future actions and payoffs do not depend on the prices at the time of deviations, conditional on punishment, is consistent with our approach. For example, the length of punishment period maybe adjusted to be finite without any substantive effect on the estimation. Adjusting the length of the punishment phase has no effect on the first-order-conditions used to estimate the model, as can be seen by the fact that future excess profits from collusion $\sum_{\tau=1}^{\infty} \beta^{\tau}(\pi_{it+\tau} - \pi_{it+\tau}^{B})$ in equation (5) do not enter the first-order condition directly, as firms can freely adjust their prices in the future.[12] Furthermore, the cost structure assumed in this paper can be further relaxed such that the sales cost c_{ft}^{S} is a random variable from the perspective of rival firms instead of a known constant. We avoid this approach due to the computational difficulty of estimation, as the probability of each state would then have to also incorporate the distribution of costs c_{ft}^{S} of each firm. However, in theory this could be implemented in a straightforward manner.

3 Estimation

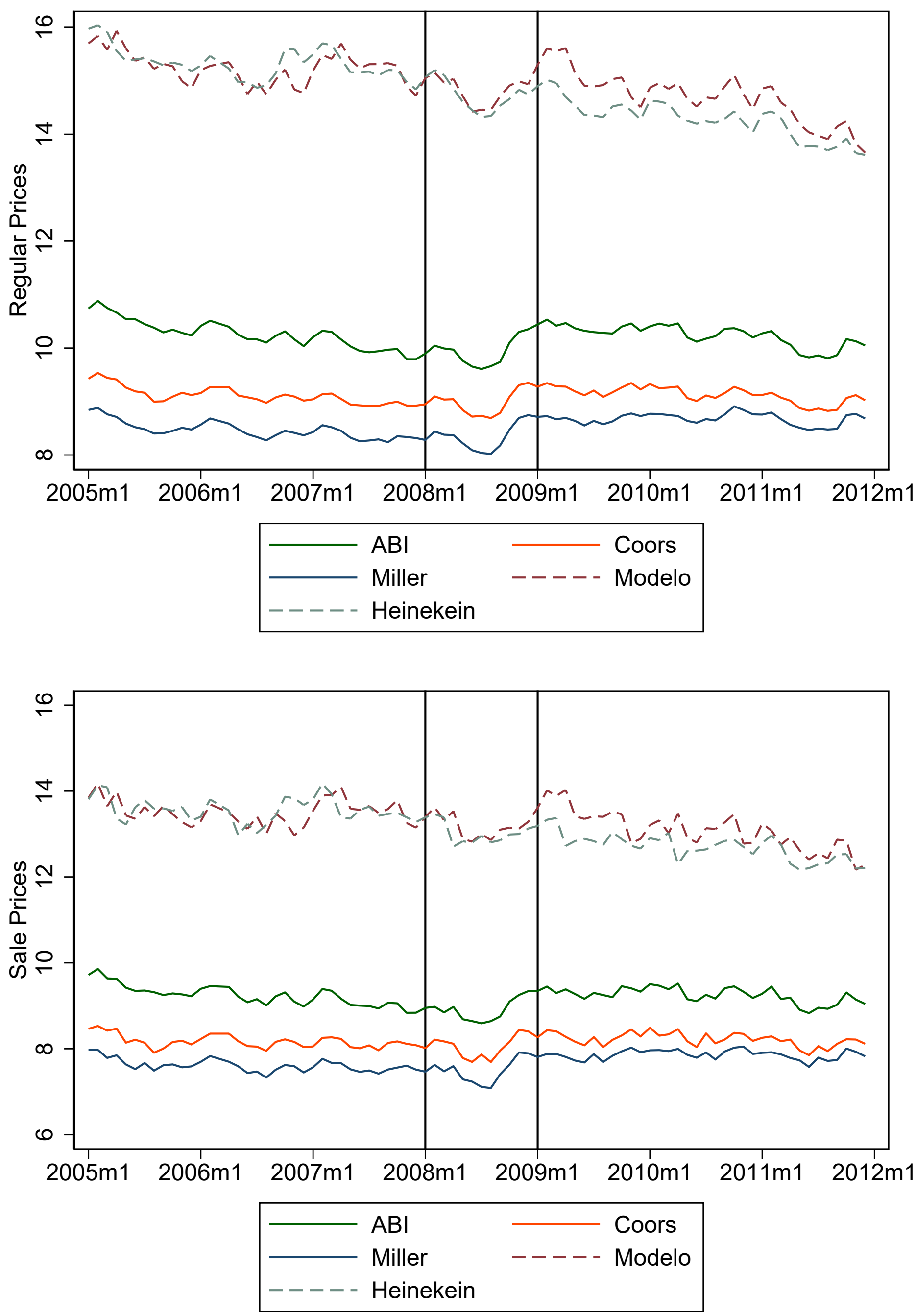
3.1 Industry Background and Data

Competition in the beer industry has been examined in various papers including those of Asker (2016), De Loecker and Scott (2016) and Miller and Weinberg (2017). Miller and Weinberg, in particular, document evidence of price coordination between ABI and MillerCoors following the merger of Miller and Coors. They utilize a differentiated-products pricing model to test and reject the hypothesis that the price increases can be explained by movement from one Nash–Bertrand equilibrium to another and find large increases in the prices and markups increase significantly due to the coordination. Furthermore, Miller and Weinberg provide documentary records from US Department of Justice complaints indicating coordination between ABI and MillerCoors, and these documents also suggests that Groupo Modelo (Modelo) and Heineken were not part of the coalition.

Therefore, throughout our analysis, we assume that the three largest domestic beer companies, ABI, Coors and Miller, and after the merger of Miller and Coors the joint

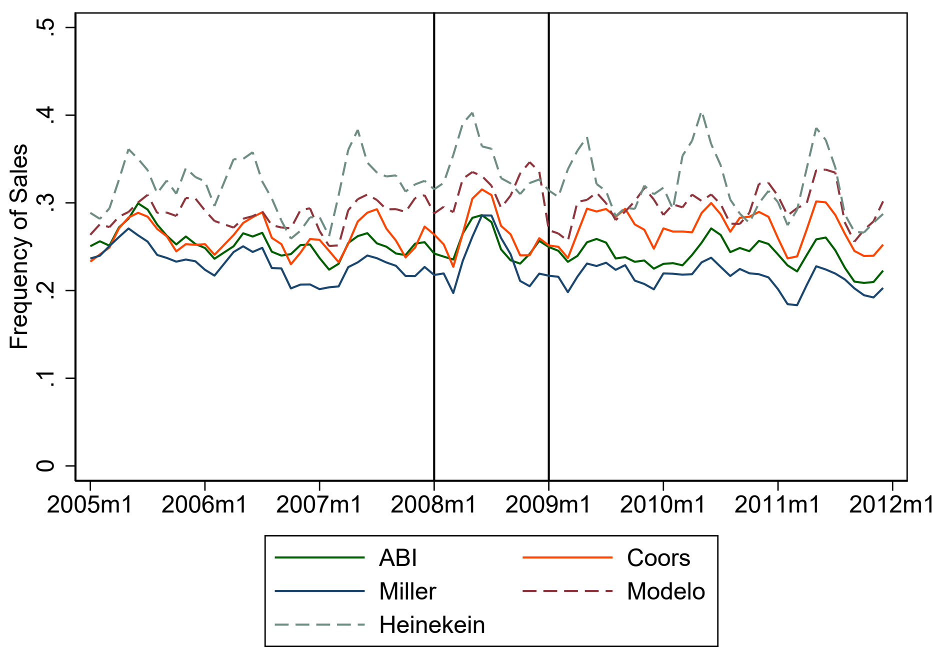
[12]However, the length of punishment will effect the threshold discount factor computation. Shortening the punishment period would increase the threshold discount factor needed to satisfy the no-deviation constraints.

Figure 2: Average Price



Note: This figure shows the regular and sales prices of each firm, averaged across products and markets. The solid horizontal lines show the time of announcement of the merger between Miller and Coors and the actual date of the merger respectively. Prices are converted to 144 ounce equivalents and deflated by the CPI deflator for all urban consumers.

Figure 3: Sale Frequency



Note: This figure shows the frequency of sales of each firm, averaged across products and markets. Sales are identified using the algorithm introduced in Midrigan and Kehoe (2015). The solid horizontal lines show the time of announcement of the merger between Miller and Coors and the actual date of the merger respectively.

MillerCoors company are coordinating their prices through tacit agreements. We also include Modelo and Heineken in our analysis, but assume they act competitively (Nash-Bertrand) and do not partake in the coalition.

The primary data this paper uses is the IRI academic scanner dataset. We use data of prices and quantities of beer products from 39 markets from 2005 to 2012.[13] We aggregate weekly data to monthly to alleviate concerns about measurement error. In addition, this aggregation alleviates concerns that demand may represent dynamic considerations driven by storage by the consumers. We deflate the prices by the CPI deflator for all urban consumers. We make use of the algorithm introduced by Kehoe and Midrigan (2015) to compute regular prices because not all of the sale prices are flagged in the data. The algorithm determines the regular price as the modal price in an 11-week window surrounding each week provided the modal price is used sufficiently often.[14] If the modal price is not used more than two-thirds of the time in a given 11 week window, the current price is determined to be the regular price. We use this as our measure of whether a given good is on sale in a given week in a given store.

Some previous papers have utilized different definitions of regular price than that described above. For example, Hendel and Nevo (2006) use the maximum or mode price within a certain period as the regular price. Nevertheless, we find that the algorithm from Kehoe and Midrigan (2015) is more likely to pick up sales than staggered price increases in our setting. Additionally, the literature varies on the depth of sale to classify a price as a sale or not. Hendel and Nevo (2006) consider various depths of sale (1%, 10%, 25%, 50%). However, we find that sales of such large depth are infrequent in our sample and for this reason we use 1%.

We use the highest selling products from ABI (Budweiser, Bud Light, Michelob and Michelob Light), Coors (Coors and Coors Light), Miller (Miller Genuine Draft,

[13]We provide a list of the markets in our dataset in Appendix A.

[14]A detailed description of the algorithm we use to compute regular prices can be found in Kehoe and Midrigan (2015).

Miller High Life, and Miller Lite) as well as Modelo (Corona and Corona Light) and Heineken for our estimation. We follow the sample selection of Miller and Weinberg (2017) closely to allow for comparison. However, to reduce the dimension of the supply side of our estimation, we aggregate the data to 12 packs of 12 ounce equivalents by dividing by the total number of ounces in the product and multiplying by 144.

To reduce the computational burden of estimation, we assume that stores within a market-firm are identical in their characteristics regarding the determination of costs and prices. They differ only as to whether they receive a shock to marginal cost in the current period or not, which is motivated by the fact that during any give period some stores in a market are on sale while others are not. We assume that the shock is independent across each product and store. Therefore, instead of treating each store within the same market-firm combination individually, we construct an average regular price across all stores for each product-market-firm and also the average depth of sales in all stores for each product-market-product in a given period. We construct a representative regular price by averaging the regular price charged at every store weighted by the volume of sales. From this we generate the sale price for each market-firm by taking the average depth below the regular price. Finally, we compute the frequency of sales for each product as the number of stores within the market that had a sale on that good (weighted by average revenue) over the total number of stores. This gives us a continuous measure of the frequency of sales in each market.

In Figures 2 and 3, we plot the regular and sales prices of each firm and sales frequency in our dataset averaged over products and markets each period. The horizontal lines show the time of announcement of the merger between Miller and Coors and the actual date of the merger respectively. We find that prices rise notably for each of ABI, Coors and Miller but not significantly for Modelo and Heineken following the merger of Miller and Coors. Regarding frequency of sales, we observe considerable cyclicality in the frequency of sales but no discernible difference in their frequency pre or post merger.[15]

In Table 1, we report the summary statistics of the prices and the sales frequency. In Table 2 we show the market share of each of the firms we use in our estimation. The firms we consider cover approximately 80% of all sales in the beer industry.

3.2 Demand Estimation

We estimate the demand system for each product using a nested logit demand system which allows for flexible estimation of the cross-price elasticities similar to the models used by De Loecker and Scott (2016) and Miller and Weinberg (2017) for the beer industry.[16] While we aggregate across sizes for most of our analysis, as previously described, we use the full set of products with differentiated packaging sizes to estimate our demand system because the variation in the availability and prices of the pack sizes helps identify the demand parameters. We then take the parameters to the aggregated data and compute the market-time-product taste residuals, and the implied aggregate elasticities and outside diversion.

The utility of consumer i for product j in market m at time t is determined as:

$$u_{ijmt} = \alpha_{price} * price_{jmt} + \tau_j + \tau_m + \tau_t + \zeta_{jt} + \chi_{jmt} + (1 - \rho) * \epsilon_{ijmt}$$

where ϵ_{ijmt} follows a independent and identically distributed extreme value type 1 distribution. ρ is the nesting coefficient on the inside option on the consumption of any beer in our sample. The nested logit parameter determines the degree of substitutability between products in our sample versus the outside option. Variables τ_j, τ_m and τ_t denote product, market and time fixed effects respectively. As we have credible exogenous variation due to transportation costs, we use the distance to the

[15] In Appendix B, we show the regression results regarding the change in prices and the frequency of sales pre and post merger. We find that the prices increase for ABI, Coors and Miller as seen in Figure 2, and that there is a small decrease in the frequency of sales.

[16] However, we use a simpler demand system than the full mixed logit used in the aforementioned papers to alleviate the computational burden of numeric integration in the supply estimation.

Table 1: SUMMARY STATISTICS

	regular price	sale price	frequency of sale
Average	11.03	9.88	0.26
Median	10.07	9.10	0.25
Std. Dev	2.83	2.49	0.16
Observations	41,378	41,378	41,378

Note: This table shows the regular and sales prices and the frequency of sales, averaged across products, markets, firms and time. The solid horizontal lines show the time of announcement of the merger between Miller and Coors and the actual date of the merger respectively. Prices are converted to 144 ounce equivalents and deflated by the CPI deflator for all urban consumers. Sales are identified using the algorithm introduced in Midrigan and Kehoe (2015).

Table 2: MARKET SHARE BY FIRM

	pre merger	post merger
ABI	0.33	0.31
Coors	0.08	0.27
Miller	0.20	-
Modelo	0.14	0.12
Heineken	0.07	0.07
Other	0.18	0.22
Total	1.00	1.00

Note: This figure shows the market share of each firm for the pre and post merger periods respectively. The pre merger period is from January 2007 to May 2008 and the post merger period and May 2009 to December 2012. The period between June 2008 to April 2009 is excluded.

Table 3: DEMAND ESTIMATION RESULTS

Estimates	coefficient	std. err.
α_{price}	-0.13	0.10
ρ	0.64	0.09
Median value	demand sample	aggregate value
Elasticity	-3.65	-3.18
Outside diversion	0.29	0.29
N	110,826	41,882

Note: Standard errors are clustered by month.

nearest brewery for each firm to each market and the distance times the cost of diesel fuel as instruments to identify the price coefficient. We use the number of products offered in each market to instrument for the nest parameter. Intuitively, this is because a larger variety of products make consumers more likely to purchase the inside good. We also include a dummy variable for the merger between Miller and Coors as in Miller and Weinberg (2017) and also interact each instrument with the dummy variable.

Table 3 reports the estimated coefficients for the demand system, as well as the median elasticity and outside diversion. The coefficient on price α_{price} is equal to -0.13, having a negative value as should be expected. The parameter on the inside nest has a value of 0.64 which implies that most of the substitution away from a given product would be to another product rather than the outside option. The median elasticity is equal to 3.65 which is in line with the literature. The median outside diversion of 0.29 implies that when consumers choose to stop purchasing a product, they will substitute to another product rather than stop purchasing the beer at all with greater than 70% probability. When we aggregate across product sizes to align with the supply model presented in the next section, the estimated demand parameters imply a median elasticity of 3.18 and outside diversion of 0.29. Thus the aggregation does not materially change these parameters.

3.3 Supply Estimation

In this section, we discuss how we estimate the dynamic game of collusion among firms with temporary price reductions. While fundamentally we assume prices and costs are determined for each store, recall that we assume that the relevant cost primitives are identical across stores of a given firm in the same market. Therefore, prices and costs are effectively determined at the product-firm-market level and each firm has a price for each product in every market.

When taking our approach in Section 2 to the data, complications arise from the fact that we must deal with multi-product firms operating in many markets simultaneously. First, with multiple layers of potential decision making units, including at product or market levels, we must determine the level at which collusion takes place. Does the decision to collude occur at the product-market level or would such decisions be made at a higher level of aggregation? Similarly, should each product-market-firm profit get individual consideration in the collusive equilibrium or should the profits be aggregated to the firm level? Related, is the question of whether collusion breaks down in all markets even when a firm deviates in just one market.

We mostly follow the standard approach in the literature, which has been to aggregate across markets and products of a given firm. We treat the decision to collude as being made at the firm level, assuming that collusion itself is a more of an all or nothing proposition where a deviation in a single market can result in the collapse of the collusive equilibrium. Likewise we consider total profits of a firm in aggregation. Our choice is motivated by the findings of Bernheim and Whinston (1990) who argue for such treatment and support this choice by demonstrating that the degree of multi-market contact helps determine collusive outcomes.

However, we do not follow the standard approach of considering only the aggregate no-deviation constraint of firms in the cartel, and consider the no-deviation constraints for each firm-market pair separately. This choice is motivated by the fact that, in the brewing industry, the largest firms may face agency issues with regard to store managers and distributors, who may have some discretion over the pricing of products at individual stores. In other words, while the decision on whether or not to collude is decided at the level of the national firm, the firms must satisfy no-deviation constraints at the store (market) level as the national headquarters may not have the complete control over each store necessary to fully enforce collusive pricing when individual stores have the incentive to deviate.[17] Therefore, the cartel will be aware of this fact and consider the necessary incentive constraints in advance, as they decide on their collusive pricing schemes.[18]

We have assumed that only ABI, Miller and Coors are in the collusive arrangement before the merger, while only ABI and MillerCoors are in the price setting coalition after the merger. Therefore, Miller's and Coors' products are attributed to a single profit maximizing agent and share the same parameters after the merger. We write the Virtual Stakes objective function for each firm f in the coalition F as follows:

$$\max_{\{p^R_{jmft}, p^S_{jmft}\}} \quad \pi_{ft} + \sum_{g \in F} \theta_{fgt} \pi_{gt} + \sum_{g \in F} \sum_{m \in M} \lambda^R_{fgmt} \big(ND^R_{gmt} \big) + \sum_{g \in F} \sum_{m \in M} \lambda^S_{fgmt} \big(ND^S_{gmt} \big) \tag{11}$$

where π_{ft} is the expected profit of the firm, ND^R_{fmt} is the no-deviation constraint for firm f in market m with a regular cost, and ND^S_{fmt} is the no-deviation constraint for firm f in market m with a low cost. The expected profit of the firm f is the sum of

[17] We believe that this setup best captures the incentives latent in the pricing of beer that arise due to the complex structure of the industry. Asker (2016) describes that due to state laws regulating alcohol, large brewers are prohibited from selling beer directly to retail outlets so they typically sell to state-licensed distributors, who, in turn, sell to retailers. Furthermore, retail price maintenance is technically illegal in many states yet distributors are often induced to sell at wholesale prices set by brewers. Anderson et al.'s (2017) description of retailer-manufacturer interactions also support this choice.

[18] We note that while the punishment for deviation can be thought of as being implemented at the national level, the aggregate no-deviation constraint of the firm is redundant with the inclusion of market level non-deviation constraints.

the expected profits of each product-market-firm is

$$\pi_{ft} = \sum_{j\in J}\sum_{m\in M} \pi_{jmft}(p^R_{jfmt}, p^S_{jfmt}; P_{-jmft}), \tag{12}$$

where the expected profit for each product-market-firm given prices in time t equals

$$\begin{aligned}
&\pi_{jmft}(p^R_{jmft}, p^S_{jmft}; P_{-jmft}) \\
&= \sum_{\omega_{-jmft}\in\Omega_{-jmft}} prob_t(\omega_{-jmft}) \quad \Big[(1-\alpha_{jmft})(p^R_{jmft} - c^R_{jmft})q(p^R_{jmft}, P_{-jmft}(\omega_{-jmft}))) \\
&\qquad\qquad + \alpha_{jmft}(p^S_{jmft} - c^S_{jmft})q(p^S_{jmft}, P_{-jmft}(\omega_{-jmft})))\Big].
\end{aligned}$$

The no-deviation constraints ND^R_{fmt} and ND^S_{fmt} represent the off-schedule no-deviation constraints for firm f in market m. Since we can write the off-schedule no-deviation constraint for firms with a regular cost as:

$$\begin{aligned}
&E_t \sum_{\tau=1}^{\infty} \beta^\tau \Big(\sum_{j\in J}\Big(\pi_{jmft+\tau} - \pi^B_{jmft+\tau}\Big)\Big) \\
&\geq \sum_{j\in J}\Big(\pi^R_{jmft}(p_{jmft}; P_{-jmft}) - \pi^R_{jmft}(p^R_{jmft}; P_{-jmft})\Big) \quad \text{for all } \{p_{jmft}\}_{j\in J},
\end{aligned} \tag{13}$$

the no-deviation constraint ND^R_{fmt} is,

$$\begin{aligned}
ND^R_{fmt} =& E_t \sum_{\tau=1}^{\infty} \beta^\tau \Big(\sum_{j\in J}\Big(\pi_{jmft+\tau} - \pi^B_{jmft+\tau}\Big)\Big) \\
&- \sum_{j\in J}\Big(\pi^R_{jmft}(p^{R*}_{jmft}; P_{-jmft}) - \pi^R_{jmft}(p^R_{jmft}; P_{-jmft})\Big)
\end{aligned} \tag{14}$$

where $\{p^{R*}_{jmft}\}_{j\in J}$ jointly maximizes $\sum_{j\in J}\pi^R_{jmft}(p_{jmft}; P_{-jmft})$ for any $\{p_{jmft}\}_{j\in J}$.

Similarly, the no-deviation constraint in the objective function ND^S_{fmt} is,

$$(15) \qquad ND^S_{fmt} = E_t \sum_{\tau=1}^{\infty} \beta^{\tau} \Big(\sum_{j \in J} \Big(\pi_{jmft+\tau} - \pi^B_{jmft+\tau} \Big) \Big) - \sum_{j \in J} \Big(\pi^S_{jmft}(p^{S*}_{jmft}; P_{-jmft}) - \pi^S_{jmft}(p^S_{jmft}; P_{-jmft}) \Big).$$

We note that we have aggregated across products in equations (12) through (15) to address multi-product firms. This reflects the assumption that deviations will occur within the store for the entire menu of firm products rather for each product separately.

Equation (11) is noticeably missing the on-schedule no-deviation constraints; i.e. the estimation equation equivalent to the constraint in the inequalities (7) and (8). The reason that equation (11) does not have on-schedule constraints and their associated multipliers is for estimation reasons. Unlike the off-schedule no-deviation constraint, the on-schedule deviation constraint is unique to each product-firm-market. Therefore, if we try to estimate (11) with the on-schedule constraints and their associate multipliers included in the Lagrangian, we end up with too many unknowns and not enough equations from which to identify them. Thus, we address this issue by restricting the action space to satisfy the on-schedule no-deviation constraints (16) and (17) for each j, m and f:

$$(16) \quad (p^R_{jmft} - c^R_{jmft}) \Big\{ \sum_{\omega_{-jmft} \in \Omega_{-jmft}} prob_t(\omega_{-jmft}) q(p^R_{jmft}, P_{-jfmt}(\omega_{-jmft}); \Theta_t) \Big\} \geq (p^S_{jmft} - c^R_{jmft}) \{ \sum_{\omega_{-jmft} \in \Omega_{-jmft}} prob_t(\omega_{-jmft}) q(p^S_{jmft}, P_{-jmft}(\omega_{-jmft}); \Theta_t) \}.$$

$$(17) \quad (p^S_{jmft} - c^S_{jmft}) \Big\{ \sum_{\omega_{-jmft} \in \Omega_{-jmft}} prob_t(\omega_{-jmft}) q(p^S_{jmft}, P_{-jmft}(\omega_{-jmft}); \Theta_t) \Big\} \geq (p^R_{jmft} - c^S_{jmft}) \{ \sum_{\omega_{-jmft} \in \Omega_{-jmft}} prob_t(\omega_{-jmft}) q(p^R_{jmft}, P_{-jmft}(\omega_{-jmft}); \Theta_t) \}.$$

We implement this procedure by first choosing and evaluating costs from the set of costs that satisfy condition (16) during the estimation procedure. Then, we check whether (17) is satisfied.

The first-order conditions from equations (11) are determined as follows:

(18)

$$[p^R_{\hat{j}mft}]: \quad 0 = \sum_{j\in J} \frac{\partial \pi_{jmft}}{\partial p^R_{\hat{j}mft}} + \sum_{g\in F} \theta_{fgt}\Big(\sum_{j\in J} \frac{\partial \pi_{jmgt}}{\partial p^R_{\hat{j}mft}}\Big)$$
$$+\sum_{g\in F} \lambda^R_{fgmt}\Big(\sum_{j\in J}\left(\frac{\partial \pi^{R*}_{jmgt}}{\partial p^R_{\hat{j}mft}} - \frac{\partial \pi^R_{jmgt}}{\partial p^R_{\hat{j}mft}}\Big)\right) + \sum_{g\in F} \lambda^S_{fgmt}\Big(\sum_{j\in J}\left(\frac{\partial \pi^{S^*}_{jmgt}}{\partial p^R_{\hat{j}mft}} - \frac{\partial \pi^S_{jmgt}}{\partial p^R_{\hat{j}mft}}\Big)\right)$$

(19)

$$[p^S_{\hat{j}mft}]: \quad 0 = \sum_{j\in J} \frac{\partial \pi_{jmft}}{\partial p^S_{\hat{j}mft}} + \sum_{g\in F} \theta_{fgt}\Big(\sum_{j\in J} \frac{\partial \pi_{jmgt}}{\partial p^S_{\hat{j}mft}}\Big)$$
$$+\sum_{g\in F} \lambda^R_{fgmt}\Big(\sum_{j\in J}\left(\frac{\partial \pi^{R*}_{jmgt}}{\partial p^S_{\hat{j}mft}} - \frac{\partial \pi^R_{jmgt}}{\partial p^S_{\hat{j}mft}}\Big)\right) + \sum_{g\in F} \lambda^S_{fgmt}\Big(\sum_{j\in J}\left(\frac{\partial \pi^{S*}_{jmgt}}{\partial p^S_{\hat{j}mft}} - \frac{\partial \pi^S_{jmgt}}{\partial p^S_{\hat{j}mft}}\Big)\right).$$

for all $\hat{j}mft$, where $\frac{\partial \pi_{jmgt}}{\partial p^{i}_{\hat{j}mft}}$ is the partial derivative of expected profit of firm g when the firms follow the equilibrium play, $\frac{\partial \pi^{\hat{i}}_{jmgt}}{\partial p^{i}_{\hat{j}mft}}$ is the partial derivative of expected profit of firm g when firm g has the cost $\hat{i}$, and $\frac{\partial \pi^{\hat{i}}_{jmgt}}{\partial p^{i*}_{\hat{j}mft}}$ is the partial derivative of maximum profit attainable from deviation for firm g when firm g has the cost $\hat{i}$. Note that because there is no relationship between prices across markets, the cross-price derivatives are all zero for prices not in their own market. The derivative of the deviation profit to the firms own prices are also all equal to zero.

3.4 Identification

While the Virtual Stakes approach provides a natural interpretation of the collusion parameters, the actual estimation procedure leans heavily on the duality between the Virtual Stakes approach and Pareto optimality. BCLW show that we can leverage Pareto optimality to impose cross-restrictions on the Virtual Stakes parameters. As previously discussed, the set Virtual Stakes parameters from each firm's problem can differ only by a scale factor. While the parameters $\{\theta_{fgt}, \lambda^R_{fgmt}, \lambda^S_{fgmt}\}$ have subscripts for both the optimizing firm f and their counterpart g, and this representation allows for a economic natural interpretation of these parameters, Pareto optimality implies that these parameters only differ by a scale factor for each optimizing firm. In other words, only $\#F - 1$ (or $(\#F - 1) \times M$) parameters are independent. Thus, for the remainder of this section, we write the Virtual Stakes parameters with only one subscript indicating the firm $\{\theta_{ft}, \lambda^R_{fmt}, \lambda^S_{fmt}\}$.

Now, of the five variables $(p^R_{jmft}, p^S_{jmft}, c^R_{jmft}, c^S_{jmft}, \alpha_{jmft})$ and three parameters $\{\theta_{ft}, \lambda^R_{fmt}, \lambda^S_{fmt}\}$ in our model, three variables $(p^R_{jmft}, p^S_{jmft}, \alpha_{jmft})$ are directly observable from the data. Furthermore, we parameterize costs c^R_{jmft} and c^S_{jmft} as

$$\begin{aligned} c^R_{jmft} &= \rho^R_{ft} + \rho^R_{mt} + \epsilon^R_{jmft} \\ c^S_{jmft} &= \rho^S_{ft} + \rho^S_{mt} + \epsilon^S_{jmft}, \end{aligned} \tag{20}$$

implying that we must estimate $\rho^R_{ft}, \rho^R_{mt}, \rho^S_{ft}, \rho^S_{mt}$. The fixed effects $\rho^R_{ft}, \rho^R_{mt}, \rho^S_{ft}, \rho^S_{mt}$ are partialled out via OLS.[19]

The Virtual Stakes parameters are estimated via GMM. We take the cost residuals $(\epsilon^R_{jmft}, \epsilon^S_{jmft})$ and form moments with a set of instruments to estimate the parameters. In the first-order conditions (18) and (19), we have $2 \times J \times M \times F$ equations for each period. This gives us as many residuals ϵ^R_{jmft} and ϵ^S_{jmft}. We construct moments conditions with those residuals interacted with our instruments to identify the Virtual Stakes parameters. We pool across time to estimate our market level parameters from the moment conditions which allows us to better identify the Virtual Stakes parameters. Because Miller and Weinberg (2017) establish a shift in the conduct of the colluding firms between the pre and post merger periods, we allow the Virtual Stakes parameters to have a different values between the pre and post merger periods.

[19]This method is similar to Nevo's (2000) treatment of demand side fixed effects.

Lastly, we briefly discuss the instruments that we use to estimate our supply model. To estimate the Virtual Stakes parameters, we need instruments that shift market prices that are independent of the producers' local costs. We need to do this to solve the endogeneity problem that the prices in each market are determined in equilibrium. We use the average quality of the products other than the firm's own in each market, the number of products, the average fuel costs across competitors, as well as the average income in each market as instruments. Since we have both sale and regular price shocks, we have eight moment conditions for each period. Our choice of instruments follows the insights of Berry and Haile (2014). We weight each observation by the fraction of that it contributes to the total revenue of the product's producer.

To summarize, we follow the following procedure for estimating the parameters of the supply model. First, we fix demand parameters estimated in Section 3.2. Then, for a given guess of the Virtual Stakes parameters $\{\theta_{ft}, \lambda^N_{fmt}, \lambda^S_{fmt}\}$, we solve each first-order equation to back out the marginal costs. We stack the marginal costs and estimate the fixed effects using linear regressions to estimate equations in (20). We take the residuals from this regression and form moments using the aforementioned instruments. Finally, we estimate the parameters using GMM separately for the pre-merger period and the most-merger period.

3.5 Estimation Results

We now present the results of our estimation of the supply side of the model. First, in Table 4, we show the average values of each of the Virtual Stakes parameter estimates, before and after the merger of Miller and Coors. As we have assumed that Modelo and Heineken do not participate in the collusive arrangement, they are estimated as acting competitively, separate from the Virtual Stakes objective. We have normalized the profit weights θ_{ft} to sum to one in our estimation ($\sum_f \theta_{ft} = 1$). We only report Virtual Stakes parameters for ABI and the joint MillerCoors company (as Coors) in the post merger periods.

In the first two rows of Table 4, we show the average value of the profit weight parameters θ_{ft} for each firm. We have estimated the model separately for the pre and post merger periods, which has the added benefit that we may cross-check our results regarding the effect of the merger with those from the previous studies, most notably Miller and Weinberg (2017). Before the merger of Coors and Miller, the profit weight for ABI is the highest among the firms suggesting that their importance to the collusive arrangement was large compared to the other two firms. However, after the merger we can see that the relative weight of the merged MillerCoors company increased significantly.

For the purposes of interpreting the Virtual Stakes parameters governing the no-deviation conditions, it can be more useful to re-scale the parameters such that the profit weight parameter of a firm θ_{ft} is equal to one. Then, the Virtual Stakes parameters governing the no-deviation conditions $(\lambda^R_{mft}, \lambda^S_{mft})$ show the value of insuring that the incentive constraints hold, in units the firm's flow profit. For example, the average value of enforcing the regular cost no-deviation constraint of Coors to ABI in the pre-merger period is captured by the re-scaled value of 6.90 (= 2.90/0.42). In other words, the shadow value of ensuring that Coors does not deviate from the collusive arrangement is equivalent to 6.90 units of profit in the current period to ABI. A general overview of Table 4 suggests that, broadly speaking, the cost of maintaining the collusive arrangement decreased following the merger of Coors and Miller, as evidenced by the decrease in the average values of $(\lambda^R_{mft}, \lambda^S_{mft})$ post merger. This is consistent with previous studies that find mergers facilitate collusion, and also the findings of Miller and Weinberg (2017) who find evidence of increased collusion after the merger of Coors and Miller.

In Figure 4, we report the regular and sales prices and estimated costs in real dollars averaged over markets and products for each firm partaking in the coalition, weighted by revenue. In all panels of Figure 4, the black lines mark the time of announcement

Table 4: VIRTUAL STAKES PARAMETER ESTIMATES (AVERAGE VALUE)

	Period	ABI	Coors	Miller
θ_{ft}	pre merger	0.42	0.37	0.21
	post merger	0.44	0.56	-
λ^R_{mft}	pre merger	2.55	2.90	2.63
	post merger	1.48	1.49	-
λ^S_{mft}	pre merger	2.98	3.15	3.02
	post merger	1.54	1.93	-

This table shows the estimated values of the Virtual Stakes parameters, averaged over market and time. The pre merger period is from January 2007 to May 2008 and the post merger period and May 2009 to December 2012. The period between June 2008 to April 2009 is excluded. The estimated parameter values for the MillerCoors joint venture in the post merger period is shown as Coors.

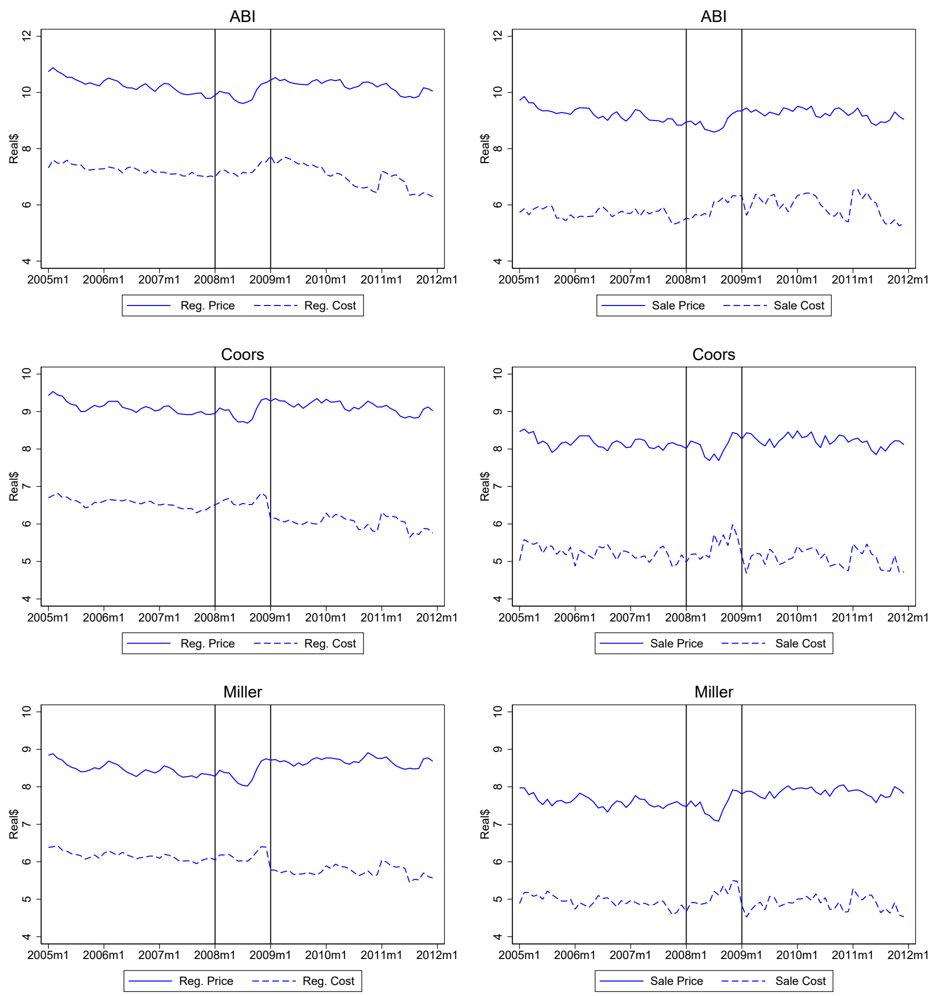
of the merger between and Miller and Coors and the timing of the actual merger respectively. The period between January 2007 to May 2008 constitutes the pre-merger period and May 2009 to December 2012 constitutes the post merger period. The pictures in the first column shows the regular prices and costs and the second column shows the sales prices and costs. For all three firms, we find that the price of beer rises significantly leading up to the merger of Miller and Coors. The costs of ABI do not exhibit a large systematic change with the merger as in the case of the price. For Miller and Coors respectively, the marginal costs seem to decrease following their merger, suggesting potential synergies and benefits from the merger. These patterns appear for both regular and sales prices and costs.

These results suggest that the markups seem to increase after the merger of Miller and Coors, which is consistent with the previous findings by Miller and Weinberg (2017) of increased collusion due to the merger. Furthermore, these results indicate that the increase in the markups post merger seem to be driven in part by the increase in prices, especially in the case of ABI. For Miller and Coors, both price increases and cost decreases contribute to the increase in markups.

The results in Figure 4 stand in contrast to the prices and costs from firms not partaking in the coalition, Modelo and Heineken. Figure 4 shows the regular and sales prices and estimated costs of Modelo and Heineken respectively. For Modelo and Heineken, there does not seem to be a notable increase in the price following the merger of Miller and Coors, and there are no noticeable structural changes to the estimated costs either. Therefore, the spike in the markups of firms seem to be concentrated in firms that are members of the pricing coalition.

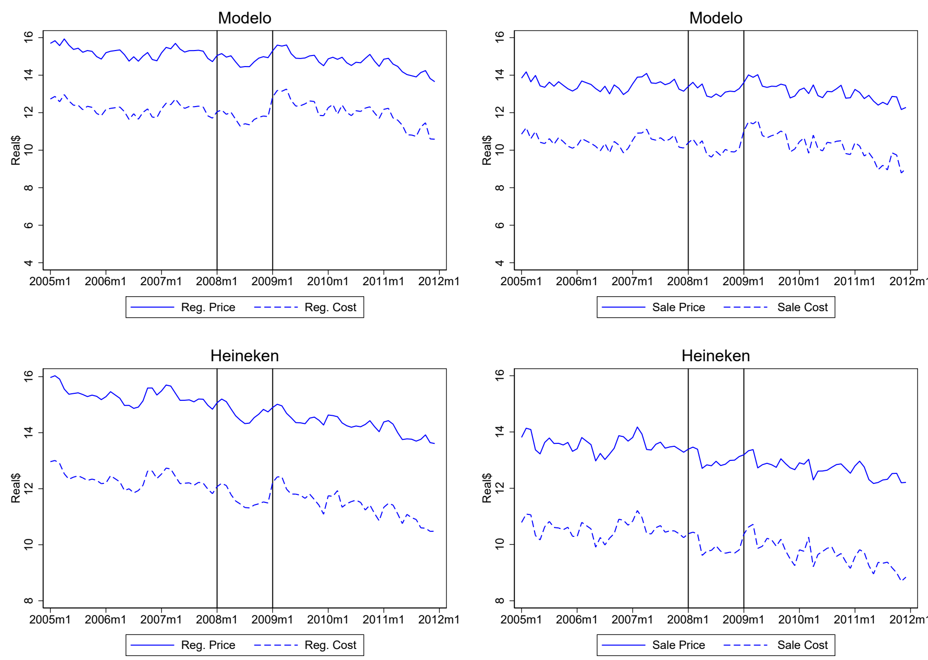
In Tables 5 and 6 we report the markups, prices and costs for each firm, averaged over the pre and post merger periods. The estimated markups are between 37% and 62% of the marginal cost for members of the coalition, which is generally in line with the upper bound of markups generally reported for consumer non-durables. As was seen in Figure 4, the average markup increased in the periods following the merger of Miller and Coors. As with markups, it is apparent from this table that the regular prices rose significantly for each firm, while the marginal costs stayed mostly constant for ABI while decreasing for Miller and Coors. The same cannot be said for Modelo and Heineken, whose markups are between 23% to 31% of the marginal cost and does not change significantly between the pre and post periods. The central reason for this difference seems to be in the pricing decision of these firms, as the behavior of costs do not show systematic differences between Modelo/Heineken versus ABI.

Figure 4: Average Prices and Costs by firm (ABI, Miller, Coors)



Note: This figure shows the prices and estimated costs for each firm, averaged over products and markets. The solid horizontal lines show the time of announcement of the merger between Miller and Coors and the actual date of the merger respectively. The costs are estimated using the Virtual Stakes approach outlined in Section 3.3.

Figure 5: Average Prices and Costs by firm (Modelo, Heineken)



Note: This figure shows the prices and estimated costs for each firm, averaged over products and markets. The solid horizontal lines show the time of announcement of the merger between Miller and Coors and the actual date of the merger respectively. The costs are estimated separately from the Virtual Stakes estimation, from a model in which Modelo and Heineken act competitively.

Table 5: Estimation Results Averages: Regular

	Period	Markup (dollars)	Markup (% of cost)	Price	Cost
ABI	pre merger	2.94	0.41	10.18	7.23
	post merger	3.24	0.46	10.24	7.01
Coors	pre merger	2.51	0.38	9.08	6.57
	post merger	3.11	0.52	9.13	6.03
Miller	pre merger	2.27	0.37	8.43	6.16
	post merger	2.94	0.51	8.68	5.74
Modelo	pre merger	3.03	0.25	15.09	12.06
	post merger	2.71	0.23	14.71	12.00
Heineken	pre merger	3.02	0.25	15.18	12.16
	post merger	2.87	0.25	14.28	11.41

Note: This table shows the markups, prices and costs for each firm, averaged over products, markets, and time. The pre merger period is from January 2007 to May 2008 and the post merger period and May 2009 to December 2012. The period between June 2008 to April 2009 is excluded.

Table 6: Estimation Results Averages: Sales

	Period	Markup (dollars)	Markup (% of cost)	Price	Cost
ABI	pre merger	3.41	0.59	9.15	5.75
	post merger	3.48	0.60	9.25	5.77
Coors	pre merger	2.86	0.54	8.14	5.27
	post merger	3.15	0.62	8.22	5.08
Miller	pre merger	2.60	0.52	7.57	4.97
	post merger	2.98	0.61	7.86	4.88
Modelo	pre merger	3.06	0.30	13.40	10.34
	post merger	2.87	0.28	13.09	10.22
Heineken	pre merger	3.03	0.29	13.42	10.39
	post merger	3.04	0.31	12.70	9.66

Note: This table shows the markups, prices and costs for each firm, averaged over products, markets, and time. The pre merger period is from January 2007 to May 2008 and the post merger period and May 2009 to December 2012. The period between June 2008 to April 2009 is excluded.

4 Counterfactual Analysis

In Section 2 we have constructed an equilibrium in which firms allow temporary deviations from the collusive regular price in the form of sales when stores receive transitory cost shocks. In this section, we explore why the cartel may be motivated to allow for such considerations. In other words, we wish to demonstrate that there is some benefit to allowing for sales within the collusive arrangement that incentivizes the cartels to do so. There are two potential avenues that we consider in which allowing for sales is advantageous to the cartel. Allowing for sales may increase the profitability of each firm or the ease at which the tacitly collusive arrangement can be sustained may increase. We show that for our sample and estimated primitives, the equilibrium allowing for sales is preferable in both criteria to a counterfactual equilibrium without sales.

We consider for comparison the equilibrium where each firm charges the regular price we observe in the data every period, and does not ever go on sale. We believe this is a natural equilibrium to consider. In this counterfactual equilibrium, firms cooperate by charging the collusive regular price in each period and punishment for defection is via the same grim-trigger strategy as before. We consider this equilibrium with the same primitives as in the collusive separating equilibrium estimated in the previous section. Thus, the Nash-Bertrand equilibrium is identical as well.

First, we compare the average profit of each firm in the collusive separating equilibrium with sales to the counterfactual no-sales equilibrium. We do so by computing the ratio of profits between the observed equilibrium with sales and the counterfactual no-sales equilibrium. In Table 7 we report the average values of this ratio pre and post merger for each firm. During the period before the merger, allowing for sales improves the profitability of ABI by 6%, Coors by 13%, and Miller by 5%. Following the merger, allowing for sales increases profits by 5% for ABI and 4% for the merged MillerCoors company. These results show that firms have the incentive to allow for sale in their collusive arrangement to improve their profits.

Next, we turn our attention now to the sustainability of collusion. In order to evaluate the sustainability of the equilibria we propose, we need a measure of the sustainability that we can construct. We focus on the threshold discount factor following the literature studying of dynamic games of collusion beginning with Abreu (1986, 1988). The threshold discount factor is the discount factor that would make the producers indifferent between collusion and defection given the equilibrium strategies and the firms' beliefs about the future. Intuitively, the higher this discount factor the more difficult collusion is to sustain. Recent papers such as Eizenberg et al. (2021) also use the threshold discount factor to characterize the stability of collusive arrangements.

As discussed in Sullivan (2017), the researcher must take a stance on the producers beliefs about the future flow of profits under both collusion and defection to compute the minimum discount factor. We assume that the payoff relevant primitives Θ_t and the corresponding firm profits follow a random walk. Paired with the fact the the firms are risk-neutral, this equates to the commonly utilized approach of assuming that Θ_t is held constant at the time of the decision to defect and that firms believe that collusive profits today would continue identically if collusion were sustained. We compute the threshold discount factors for the strategy in which firms play a grim-trigger strategy of defecting to the Nash-Bertrand equilibrium in perpetuity when they detect defection.[20]

Under these assumptions we can compute the threshold discount factor from (14) and (15) for each market-firm as follows:

$$\hat{\beta}_{mft} = \frac{\pi^*_{mft} - \pi_{mft}}{\pi^*_{mft} - \pi^B_{mft}} \tag{21}$$

[20]In contrast to our estimation results, the assumption of grim-trigger strategies are substantive to the computation of the threshold discount factors. A less permanent punishment strategy will have the effect of increasing the threshold discount factors, making that collusion strategy more difficult to sustain.

where

$$\pi^*_{mft} \equiv (1-\alpha_{mft})\pi^R_{mft}(\{p^{R*}_{jmft}\}_{j\in J}) + \alpha_{mft}\pi^S_{mft}(\{p^{S*}_{jmft}\}_{j\in J}),$$
$$\pi^i_{mft}(\{p^{i*}_{jmft}\}_{j\in J}) \equiv \sum_{j\in J} \pi^i_{jmft}(p^{i*}_{jmft}; P_{-jmft}),$$

and $\{p^{i*}_{jmft}\}_{j\in J}$ jointly maximizes $\sum_{j\in J} \pi^i_{jmft}(p^{R*}_{jmft}; P_{-jmft})$. The threshold discount factor holds each firm f indifferent between defecting and continuing to collude. Profits reflect the total profits aggregated across products in each market for each firm at that time. We report the 90th percentile values of $\hat{\beta}_{mft}$ for each firm in Table 8.

In Table 8 we report the 90th percentile threshold minimum discount factors that supports each of the estimated and counterfactual equilibria. We find that the 90th percentile threshold discount factors are higher in the counterfactual equilibria than the observed equilibrium with sales for all three firms, both in the pre merger and post merger periods. In addition, we also observe that for ABI the threshold discount factor slightly increases after the merger of Coors and Miller. However, the discount factor still hovers below 0.6 in the observed equilibrium, which is significantly lower than the usual discount factors of firms. On the other hand, the threshold discount factor for Coors and Miller are high in the pre-merger period and decreases significantly following their merger. Considering the fact that the collusive arrangement depends on all of the participants maintaining the arrangement, this suggests that the sustainability of collusive arrangement overall benefited from the merger of Coors and Miller.

In Figure 6, we plot three-month moving average of the time series of the threshold minimum discount factors for each of ABI, Coors and Miller. The figures illustrate that the observed equilibrium in which sales are allowed is also more easily sustained in most periods individually throughout our sample, and even more so in the post merger period.

We can further decompose the sustainability results by considering the following

Table 7: Ratio of profits between equilibria

	Period	Average	10th percentile	90th percentile
ABI	pre merger	1.06	0.90	1.23
	post merger	1.05	0.90	1.22
Coors	pre merger	1.13	0.87	2.03
	post merger	1.04	0.94	1.25
Miller	pre merger	1.05	0.94	1.07
	post merger	-	-	-

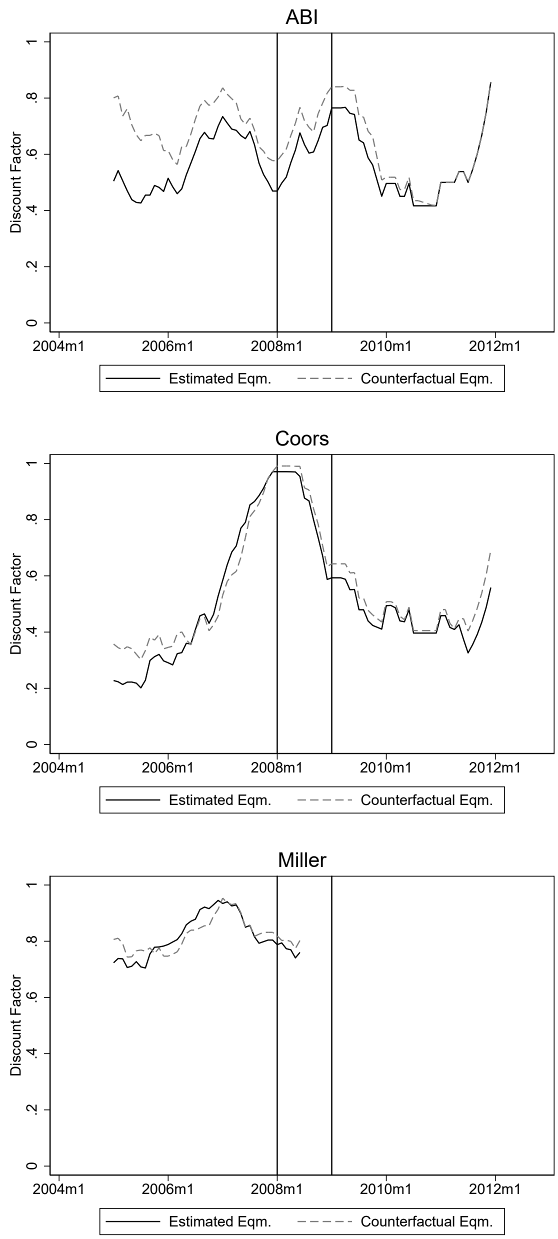
Note: This table shows the ratio of the profits between the estimated equilibrium and the counterfactual equilibrium. The pre merger period is from January 2007 to May 2008 and the post merger period and May 2009 to December 2012. The period between June 2008 to April 2009 is excluded. The values for the MillerCoors joint venture in the post merger period is shown as Coors.

Table 8: Discount factor by period

		90th percentile $\hat{\beta}_{mft}$	
	Period	estimated equilibrium	counterfactual
ABI	pre merger	0.57	0.72
	post merger	0.59	0.66
Coors	pre merger	0.78	0.80
	post merger	0.53	0.59
Miller	pre merger	0.76	0.76
	post merger	-	-

Note: This table shows the 90th percentile values of the threshold discount factor needed to sustain the estimated equilibrium and the counterfactual equilibrium, respectively. The pre merger period is from January 2007 to May 2008 and the post merger period and May 2009 to December 2012. The period between June 2008 to April 2009 is excluded. The values for the MillerCoors joint venture in the post merger period is shown as Coors.

Figure 6: Threshold Discount Factor: 90th percentile



Note: This figure shows the 90th percentile values of the threshold discount factor needed to sustain the estimated equilibrium and the counterfactual equilibrium, respectively. The values for the MillerCoors joint venture in the post merger period is shown as Coors.

equation,

$$\frac{\hat{\beta}_{mft}}{1-\hat{\beta}_{mft}} = \frac{\frac{\pi^*_{mft}-\pi_{mft}}{\pi^B_{mft}}}{\frac{\pi_{mft}-\pi^B_{mft}}{\pi^B_{mft}}}. \tag{22}$$

Equation (22) shows that the threshold discount factor can be decomposed as the difference between the deviation profits and the equilibrium profits over the difference between the equilibrium profits and the Nash-Bertrand profits from the punishment stage, as a ratio of the Nash-Bertrand profit. Intuitively, the larger the deviation profits the larger the discount factor will have to be sustain collusion with punishment in the future. In contrast, if the difference between the equilibrium profit and the Nash Bertrand profit is large, collusion will be easier to sustain.

In Table 9 we show an example of the values of these ratios averaged for markets falling within a two percent band around the 90th percentile threshold discount factors. We find that overall, the profit from defection is larger and the benefit from collusion smaller in the counterfactual equilibrium than the estimated equilibrium with sales. Thus both factors, the fact that deviations are less profitable and that collusion is more beneficial, contribute to the increased sustainability of collusion when sales are allowed in response to temporary shocks to marginal cost.

We conclude that not only is the collusive separating equilibrium more profitable than a natural counterfactual equilibrium, but it is also more sustainable. Not only do sales allow the firms to take advantage of opportunities that rise due to cost reductions to increase profit, but price reductions also relieve the pressure to defect making collusion easier to sustain.

5 Conclusion

In this paper we show that explicitly considering temporary price reductions adds to our understanding of competition and collusion. Including sales in our analysis not

Table 9: Components of the Discount Factor

		Defection Profit		Collusive Profit	
	Period	estimated	counterfactual	estimated	counterfactual
ABI	pre merger	1.51	1.55	1.20	1.02
	post merger	1.82	1.98	1.42	1.16
Coors	pre merger	4.67	4.75	1.44	1.37
	post merger	1.33	1.36	1.07	1.03
Miller	pre merger	2.23	2.20	1.19	1.12
	post merger	-	-	-	-

Note: This table shows decomposition of the threshold discount factor into defection profits and collusion profits for the estimated equilibrium and the counterfactual equilibrium, respectively. The values shown are the averaged value of markets within a two percent band around the 90th percentile discount factor. The pre merger period is from January 2007 to May 2008 and the post merger period and May 2009 to December 2012. The period between June 2008 to April 2009 is excluded. The values for the MillerCoors joint venture in the post merger period is shown as Coors.

only adds a real world feature, but also alters our inference into collusive arrangements in a substantive manner. We find that sales allow some flexibility in the tacit collusive arrangement, increasing the profitability and sustainability of pricing cartels.

The broader implications of our findings suggest that collusive arrangements may be more flexible in their nature than previously thought. They may be adaptable to various circumstances such as environments with shocks to cost/demand and private/imperfect information that have previously been suspected of making collusion more difficult. While our paper explores the case of privately observed cost shocks and sales, other competitive mechanisms may serve similar functions in alternative environments.

References

Abreu, Dilip. "Extremal equilibria of oligopolistic supergames." Journal of Economic Theory 39, no. 1 (1986): 191-225.

Abreu, Dilip. "On the theory of infinitely repeated games with discounting." Econometrica: Journal of the Econometric Society (1988): 383-396.

Abreu, Dilip, David Pearce, Ennio Stcchetti. "Optimal cartel equilibria with imperfect monitoring." Journal of Economic Theory 39, no.1 (1986): 251-269

Anderson, Eric, Benjamin A. Malin, Emi Nakamura, Duncan Simester, and Jon Steinsson. "Informational rigidities and the stickiness of temporary sales." Journal of Monetary Economics 90 (2017): 64-83.

Asker, John. "Diagnosing Foreclosure Due to Exclusive Dealing." Journal of Industrial Economics 64, no.3 (2016): 375-410

Asker, John and Volker Nocke. "Collusion, Mergers, and Related Antitrust Issues." no. w29175. National Bureau of Economic Research, 2021.

Athey, Susan, and Kyle Bagwell. "Optimal Collusion with Private Information," RAND Journal of Economics 32, no. 3 (2001): 428-465.

Athey, Susan, Kyle Bagwell, and Chris Sanchirico. "Collusion and Price Rigidity," Review of Economic Studies 71, (2004): 317-349

Athey, Susan, and Kyle Bagwell. "Collusion with Persistent Cost Shocks," Econometrica 76, no. 3 (2008): 493-540

Bernheim, B. Douglas, and Michael D. Whinston. "Multimarket contact and collusive behavior." RAND Journal of Economics 21, no.1 (1990): 1-26.

Berry, Steven T., and Philip A. Haile. "Identification in differentiated products markets using market level data." Econometrica 82, no. 5 (2014): 1749-1797.

Black, R. Michael, Gregory S. Crawford, Shihua Lu, and Halbert White. "A Virtual Stakes Approach to Measuring Competition In Product Markets." Working Paper (2004).

Bresnahan, Timothy F. "Duopoly Models with Consistent Conjectures." American

Economic Review 71, no. 5 (1981): 934-945

Bresnahan, Timothy F. “Competition and collusion in the American automobile industry: The 1955 price war.” Journal of Industrial Economics (1987): 457-482.

Carlton, Dennis W. “The Rigidity of Prices.” American Economic Review 76, no. 4 (1986): 637-658

Carlton, Dennis W. “The Theory and the Facts of How Markets Clear: Is Industrial Organization Valuable for Understanding Macroeconomics?” in R. Schmalensee and R. D. Willig (eds.) Handbook of Industrial Organization, Volume 1, Handbooks in Economics, No. 10 (Oxford: North-Holland) (1989): 909-946.

Ciliberto, Federico, and Jonathan W. Williams. “Does multimarket contact facilitate tacit collusion? Inference on conduct parameters in the airline industry.” RAND Journal of Economics 45, no.4 (2014): 764-791

Corts, Kenneth S. “Conduct parameters and the measurement of market power.” Journal of Econometrics 88, (1988): 227-250.

De Loecker, Jan, and Paul T. Scott. “Estimating market power Evidence from the US Brewing Industry.” no. w22957. National Bureau of Economic Research, (2016).

Eichenbaum, Martin, Nir Jaimovich, and Sergio Rebelo. “Reference Prices, Costs, and Nominal Rigidities.” American Economic Review 101 (2011): 234-262

Eizenberg, Alon, Dalia Shilian, Daniel D. Blanga. “Estimating the Potential Effect of Multi-Market Contact on the Intensity of Competition.” Working Paper (2021)

Fan, Ying, and Chris Sullivan. “Measuring Competition with a Flexible Supply Model” Working Paper (2020).

Green, Edward J., and Robert H. Porter. “Noncooperative Collusion Under Imperfect Price Information.” Econometrica 52, no.1 (1984): 87-100.

Hanazono, Nakoto, and Huanxing Yang. “Collusion, Fluctuating Demand, and Price Rigidity.” International Economic Review 48, no. 2 (2007): 483-515

Hendel, Igal, and Aviv Nevo. “Measuring the implications of sales and consumer inventory behavior.” Econometrica 74, no. 6 (2006): 1637-1673.

Igami, Mitsuru, and Takuo Sugaya. "Measuring the Incentive to Collude: The Vitamin Cartels, 1990-1999." Review of Economic Studies 89, no. 3 (2022): 1460-1494.

Kehoe, Patrick, and Virgiliu Midrigan. "Prices are sticky after all." Journal of Monetary Economics 75 (2015): 35-53.

Michel, Christian, and Stefan Weiergraeber. "Estimating Industry Conduct in Differentiated Products Markets: The Evolution of Pricing Behavior in the RTE Cereal Industry." Working Paper (2018)

Miller, Nathan H., and Matthew C. Weinberg. "Understanding the price effects of the MillerCoors joint venture." Econometrica 85, no. 6 (2017): 1763-1791.

Nava, Francesco, and Pasquale Schiraldi. "Sales and Collusion in a Market with Storage." Journal of the European Economic Association 12, no. 3 (2014): 791-832

Nevo, Aviv. "A Practitioner's Guide to Estimation of Random-coefficients Logit Models of Demand." Journal of Economics and Management Strategy 9, no. 4 (2000): 513-548

Porter, Robert H. "A study of cartel stability: the Joint Executive Committee, 1880-1886." The Bell Journal of Economics (1983): 301-314.

Roos, Nicolas de, and Vladimir Smirnov. "Collusion with intertemporal price dispersion." RAND Journal of Economics 51, no. 1 (2020): 158-188.

Rotemberg, Julio J., and Garth Saloner "A Supergame-Theoretic Model of Price Wars during Booms." American Economic Review 76, no. 3 (1986): 390-407.

Sullivan, Christopher. "Three Essays on Product Collusion." PhD Thesis (2017).

Sweeting, Andrew, Xuezhen Tao, and Xinlu Yao. "Dynamic Oligopoly Pricing with Asymmetric Information: Implications for Horizontal Mergers." no. w28589. National Bureau of Economic Research, (2022).

Appendix A Markets

Following Miller and Weinberg (2017), we choose 39 markets on which we conduct our analysis. The designated market areas are defined by IRI and correspond approximately (although not exactly) to a metropolitan statistical area (MSA) consisting of several counties but sometimes extending to a larger region such as New England. The full list is provided in Table 10.

Table 10: Markets

Name
ATLANTA
BIRMINGHAM MONTGOMERY
BOSTON
BUFFALO ROCHESTER
CHARLOTTE
CHICAGO
CLEVELAND
DALLAS
DES MOINES
DETROIT
GRAND RAPIDS
GREEN BAY
HARTFORD
HOUSTON
INDIANAPOLIS
KNOXVILLE
LOS ANGELES
MILWAUKEE
MISSISSIPPI
NEW ORLEANS
NEW YORK
OMAHA
PEORIA SPRINGFLD
PHOENIX AZ
PORTLAND OR
RALEIGH DURHAM
RICHMOND NORFOLK
ROANOKE
SACRAMENTO
SAN DIEGO
SAN FRANCISCO
SEATTLE TACOMA
SOUTH CAROLINA
SPOKANE
ST. LOUIS
SYRACUSE
TOLEDO
WASHINGTON DC
WEST TEX NEW MEXICO

Appendix B Evidence on the Effect of the Merger

In this section we document the behavior of prices and sales in our sample, and the impact of the merger of Miller and Coors on their outcomes. As many papers including Miller and Weinberg (2017) demonstrate, a merger can significantly impact the competitive environment. Therefore, we investigate whether these previous findings are borne out in our data, and whether we should account for this factor in our analysis of collusion and estimation, as previous studies suggest. Recall from Figure 2, that the prices of each of ABI, Miller and Coors seemed to increase post merger relative to Modelo and Heineken, while there was no immediately discernible pattern in the frequency of sales. In this section, we explore whether such patterns can be identified from a descriptive regression analysis of pre and post merger outcomes. We run the following regressions at the market-firm-month level:

$$
\begin{aligned}
y_{mft} =& \alpha + \beta_{MC}\mathrm{MC}_{mf} + \beta_{ABI}\mathrm{ABI}_{mf} + \beta_{post}\mathrm{post}_t \\
&+ \beta_{MC,post}(\mathrm{MC}_{mf} \times \mathrm{post}_t) + \beta_{ABI,post}(\mathrm{ABI}_{mf} \times \mathrm{post}_t) + \epsilon_{mft}
\end{aligned}
$$

where y_{mft} is the strategic variable of interest. These include the average regular price, the average sales price, and the frequency of sales by market-firm-month. MC_{mf} is a dummy variable indicating whether a product is sold by either Miller or Coors prior to the merger and MillerCoors post merger, and ABI_{mf} is a dummy variable indicating whether a product is sold by ABI. post_t is a dummy variable indicating the time periods after the merger was consummated. We include Modelo and Heineken as a control group of firms that neither merge nor are thought to be part of the pricing coalition. We exclude the period from June 2008 to April 2009 to account for immediate complications leading up to and after the merger.

First, Table 11 shows the results for the average regular price. The first column are the results from the specification that includes only firm fixed effects, the second column includes firm and time fixed effects, and the last column controls for firm specific trends. All standard errors are clustered on month. The results show that the regular prices of ABI, Miller, and Coors all rose significantly after the merger compared to the prices of Modelo and Heineken. In the specifications with fixed effects, the price of ABI increased by approximately 1.02 to 1.03 dollars after the merger depending on the specification, while the price of Miller and Coors increased by 0.96 dollars post merger. The results disappear with the inclusion of firm-specific trends. The results for the sales price, shown in Table 12 show similar results. Table 13 shows the regression results for the frequency of sales. In the specifications that do not include firm-specific trends, the frequency of sales for ABI seem to decreases by around 1.5 to 1.7 percentage points after the merger, while the effect seems to be smaller Miller and Coors. Firm-specific trends do seem to effect the results.

This section has shown descriptive evidence of the effect of the merger between Miller and Coors on important strategic variables. While this merger is not the major focus of this paper, it allows some inference into a major shift in the competitive environment, which we incorporate into our analysis in the estimation of our model parameters in Section 3.

Table 11: Average regular price

	(1)	(2)	(3)
β_{post}	-0.223*		0.692***
	(0.112)		(0.206)
$\beta_{ABI,post}$	1.018***	1.028***	-0.205
	(0.159)	(0.158)	(0.278)
$\beta_{MC,post}$	0.956***	0.965***	-0.543**
	(0.108)	(0.107)	(0.266)
Observations	41,879	41,879	41,879
R-squared	0.107	0.113	0.975
Firm FE	Y	Y	Y
Time FE	N	Y	N
Firm Trend	N	N	Y

Note: Standard errors clustered on month.
*** p<0.01, ** p<0.05, * p<0.1

Table 12: Average sales price

	(1)	(2)	(3)
β_{post}	0.0943		-5.188***
	(0.140)		(0.743)
$\beta_{ABI,post}$	0.544***	0.915***	3.904***
	(0.177)	(0.135)	(0.718)
$\beta_{MC,post}$	0.235	0.881***	-1.550***
	(0.214)	(0.0998)	(0.472)
Observations	41,879	41,879	41,879
R-squared	0.101	0.118	0.914
Firm FE	Y	Y	Y
Time FE	N	Y	N
Firm Trend	N	N	Y

Note: Standard errors clustered on month.
*** p<0.01, ** p<0.05, * p<0.1

Table 13: Frequency of Sales

	(1)	(2)	(3)
β_{post}	0.00155		0.00860
	(0.0100)		(0.0139)
$\beta_{ABI,post}$	-0.0170*	-0.0154***	-0.0185**
	(0.00845)	(0.00314)	(0.00801)
$\beta_{MC,post}$	-0.00821	-0.00816***	-0.00902
	(0.00865)	(0.00297)	(0.00842)
Observations	41,879	41,879	41,879
R-squared	0.044	0.393	0.749
Firm FE	Y	Y	Y
Time FE	N	Y	N
Firm Trend	N	N	Y

Note: Standard errors clustered on month.

*** p<0.01, ** p<0.05, * p<0.1