NBER WORKING PAPER SERIES

# DYNAMIC OLIGOPOLY PRICING WITH ASYMMETRIC INFORMATION: IMPLICATIONS FOR HORIZONTAL MERGERS

Andrew Sweeting
Xuezhen Tao
Xinlu Yao

Working Paper 28589
http://www.nber.org/papers/w28589

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2021

This research was supported by US National Science Foundation Grant SES-1658670 (Sweeting) and National Natural Science Foundation of China Grant 72003120 (Tao). We thank Nate Miller, Matt Weinberg, Gloria Sheu, Dan Vincent, Bob Majure, Joe Harrington, Kyle Bagwell and a number of seminar participants for valuable comments, which have been especially useful in framing the paper. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Dynamic Oligopoly Pricing with Asymmetric Information: Implications for Horizontal Mergers
Andrew Sweeting, Xuezhen Tao, and Xinlu Yao
NBER Working Paper No. 28589
March 2021
JEL No. D43,D82,L13,L41,L90

## ABSTRACT

We model differentiated product pricing by firms that possess private information about serially-correlated state variables, such as their marginal costs, and can use prices to signal information to rivals. In a dynamic game, signaling can raise prices significantly above static complete information Nash levels even when the privately observed state variables are restricted to lie in narrow ranges. We calibrate our model using data from the beer industry, and we show that our model can explain changes in price levels and price dynamics after the 2008 MillerCoors joint venture.

Andrew Sweeting
Department of Economics
University of Maryland
Tydings Hall 3135
College Park, MD 20742
and NBER
sweeting@econ.umd.edu

Xinlu Yao
University of Maryland
Tydings Hall
College Park, MD 20742
United States
yao@econ.umd.edu

Xuezhen Tao
Shanghai University of Finance and Economics
318 Wuchuan Rd
Wujiaochang, Yangpu District
Shanghai 200083
China
tao.xuezhen@mail.shufe.edu.cn

A data appendix is available at http://www.nber.org/data-appendix/w28589

# 1 Introduction

Theoretical and empirical analyses of differentiated product markets usually assume that firms have complete information (CI) and set prices to maximize their current profits. If an alternative is considered, it is typically tacit collusion with repeated CI stage games. These assumptions provide tractability, but there is surprisingly little evidence that they accurately predict how prices change after events such as mergers. The CI assumption is also inconsistent with how firms closely guard information about the margins of individual product lines and how sensitively this information is treated during merger investigations.

This paper considers what happens when we relax the static and CI assumptions. Specifically, we will assume that each firm has a payoff-relevant state variable, such as its marginal cost, which is imperfectly serially-correlated and unobserved by rivals. In this environment, each firm may want to choose its price strategically to affect its rivals' inferences. We will consider fully separating equilibria where, in equilibrium, a firm's chosen price perfectly reveals its current cost, and beliefs have a simple form. In these equilibria, all firms that do not have the lowest possible marginal cost set prices above static best response levels to credibly signal this information to their rivals. This can, in turn, cause static best response prices to increase, and signaling prices to rise further, a positive feedback that can cause equilibrium prices to be significantly above static CI Nash levels, although, as we discuss, separating equilibria may not exist if prices rise too much. While a small theoretical literature has shown that oligopoly signaling can affect equilibrium prices in two- or three-period models, we provide the first analysis of how large these effects may be, and the first empirical application.

We apply our model to horizontal merger analysis. Signaling is a strategic investment to raise rivals' future prices, and like many strategic investments, the equilibrium incentive to invest can rise when the number of competitors is reduced. We use an example to illustrate how a standard static CI merger simulation can significantly underpredict post-merger price increases if the firms are playing a dynamic signaling game. We then apply the model to data from the U.S. beer market around the time of the 2008 Miller-Coors (MC) joint venture (JV). Miller and Weinberg (2017) (MW) show that, after the JV, domestic brewers' prices increased in a way that is inconsistent with static CI Nash pricing. We calibrate our dynamic signaling model using only data on pre-JV price dynamics and show that it predicts the observed change in the level of prices accurately and that it also predicts directional changes in measures of observed price dynamics. We also extend MW's conduct parameter framework (Bresnahan (1982), Lau (1982), Nevo (1998), Berry and Haile (2014)) to show that the CI tacit collusion explanations for the post-JV price increase advanced by MW and Miller, Sheu, and Weinberg (2020) (MSW) do not fully describe the pricing of domestic brewers, suggesting the need to explore new explanations, such as ours.

Before discussing the related literature, we should be clear about several limitations of our analysis. First, we have to assume that each firm has exactly one privately-known state and can send exactly one signal per period. This imposes restrictions on how firms are modeled after mergers. Second, we only consider fully separating equilibria, even though these may not exist for some parameters and we can only prove existence and uniqueness in special cases. Third, while we can reject some specific tacit collusion models, folk theorems imply that collusive models may exist that could fit the data perfectly.

The rest of this introduction reviews the related literature. Section 2 lays out the model and the equilibrium concept. Section 3 presents some examples and illustrates the implications for merger analysis. Section 4 provides our empirical application. Section 5 concludes. The online Appendices detail the computational algorithms; additional examples; a proof of existence and uniqueness for the case of linear demand; and, further details of the data and empirical analysis.

**Related Literature.** Shapiro (1986) and Vives (2011) examine how equilibrium prices and welfare change when marginal costs are private information in one-shot oligopoly models. Most of our focus will be on models where marginal costs lie in quite narrow intervals and the static effects that these papers identify are very small. A large theoretical literature has considered one-shot signaling models where only one player has private information. The classic Industrial Organization example is the Milgrom and Roberts (1982) limit pricing model, where an incumbent monopolist may lower its first period price to deter entry in a two-period game. Sweeting, Roberts, and Gedge (2020) develop finite and infinite-horizon versions of this model where an incumbent monopolist's type changes over time, as we will assume in this paper.[1] They

[1] Kaya (2009) and Toxvaerd (2017) analyze one-sided, dynamic signaling games where the informed firm's type is fixed, and, in equilibrium, the informed firm signals until its reputation is established.

estimate the model and show that it can explain why incumbent airlines dropped prices by as much as 15% when Southwest threatened entry on monopoly routes. The oligopoly setting considered here is potentially applicable to many more markets.

The literature on games where multiple players signal simultaneously is much more limited.[2] Mailath (1988) identifies conditions under which a separating equilibrium will exist in an abstract two-period game with continuous types, and shows that the conditions on payoffs required for the uniqueness of each player's separating best response function are similar to those shown by Mailath (1987) for models where only one player is signaling (Mailath and von Thadden (2013) generalize these conditions). Mailath (1989) applies these results to a two-period pricing game where differentiated firms have static linear demands and marginal costs that are private information but fixed. Firms raise their prices in the first period in order to try to raise their rivals' prices in the second period.[3] Mester (1992) extends this approach to a three-period quantity-setting model where marginal costs change over time, and she shows that signaling, which leads to increased output in this case, happens in the first two periods.

We rely on Mailath's results to characterize best response signaling pricing functions, and we will focus on the magnitude, empirical relevance and implications of the equilibrium effects in multi-period settings with more standard forms of differentiated product demand. Fershtman and Pakes (2012) and Asker, Fershtman, Jeon, and Pakes (2020) develop an alternative approach to discrete state and discrete action dynamic games with asymmetric information. They reduce the computational burden using the concept of Experience-Based Equilibrium (EBE) where firms have beliefs about their payoffs from different actions rather than rivals' types.[4] Our equilibrium concept is more standard, and the computational burden is reduced by focusing on fully separating equilibria in continuous action games.

We discuss the relationship between our paper and discussions of coordinated effects in horizontal merger analysis (Ordover (2007), Baker and Farrell (forthcoming), Farrell and Baker

[2] Bonatti, Cisternas, and Toikka (2017) analyze linear signaling strategies in a continuous-time Cournot game where each firm's marginal cost is private information and fixed, but firms cannot perfectly observe the quantities that their rivals choose. We will assume that prices are perfectly observable.

[3] Caminal (1990) considers a two-period linear demand duopoly model where firms have private information about the demand for their own product, and also raise prices to signal that they will set higher prices in the final period.

[4] The rest of the literature on dynamic games, following Ericson and Pakes (1995) and Pakes and McGuire (1994), has assumed that players observe all state variables up to iid payoff shocks so that there is no role for signaling.

(2021)) in the conclusion. Our paper is partly motivated by the empirical merger retrospectives literature. Ashenfelter, Hosken, and Weinberg (2014) find that 36 of 49 studies across several industries identify significant post-merger price increases.[5] Peters (2009) and Garmon (2017) show that merger simulations and other methods, such as pricing pressure indices, that are derived from static CI first-order conditions often perform poorly at predicting price changes after airline and hospital mergers. This leads naturally to the question of which alternative models can do better.

# 2 Model

In this section, we present our general model. More specific assumptions will be made in our examples and application.

## 2.1 Outline.

There are discrete time periods, $t = 1, ..., T$, where $T \leq \infty$, with discount factor $0 < \beta < 1$. $\beta = 0.99$ in the rest of the paper. There are a fixed set of $N$ risk-neutral firms. Each firm either sells a single-product or sells multiple products, which are symmetric in demand and are produced at the same marginal cost, at a single price. There may be observed and fixed differences in demand and costs across firms, but exactly one dimension of a firm's type is private information. In the text, we will assume that the type is continuous on a known compact interval $[\underline{\theta_i}, \overline{\theta_i}]$, but Appendix B.1 uses examples where firms can have two discrete types, $\underline{\theta_i}$ and $\overline{\theta_i}$. Types are assumed to evolve exogenously, and independently, from period-to-period according to a first-order Markov process, $\psi_i : \theta_{i,t-1} \rightarrow \theta_{i,t}$.[6]

[5] Ashenfelter, Hosken, and Weinberg (2014) note that retrospectives have not typically found price increases in banking. Interestingly, the Mester (1992) analysis of a Cournot oligopoly model with asymmetric information was explicitly motivated by a desire to explain why, contrary to the predictions of Nash and tacit collusion models, concentration appeared to lead to more competitive behavior in banking.

[6] This assumption seems unrealistic, but it is consistent with how the empirical literature on horizontal mergers and the production function literature that has followed Olley and Pakes (1996) (see Doraszelski and Jaumandreu (2013) for an exception) has treated marginal cost or productivity changes.

## 2.2 Within-Period Timing.

In each period $t$ of the game, timing is as follows. Firms enter period $t$ with their $t-1$ types, which then evolve according to $\psi_i$. Firms observe their own new types, but neither the previous nor the new type of other firms.[7] Each firm then simultaneously chooses a price, $p_{i,t}$, with no menu costs. Once a firm sets its period $t$ price, it is unable to change it. A firm's profits are given by $\pi_i(p_{i,t}, p_{-i,t}, \theta_{i,t})$ and we assume that $\frac{\partial \pi_i}{\partial p_{-i,t}} > 0$ for all $-i$. Note that $\pi_i(p_{i,t}, p_{-i,t}, \theta_{i,t})$ only depends on current prices and the firm's type, consistent with static and time-invariant demand. Current and past prices are assumed to be perfectly observed by each firm.

## 2.3 Assumptions.

For continuous types, we make the following assumption.

**Assumption 1** ***Type Transitions for the Continuous Type Model.*** *The conditional pdf* $\psi_i(\theta_{i,t}|\theta_{i,t-1})$

1. *has full support, so that the type can transition from any value on the support to any other value in a single period.*

2. *is continuous and differentiable (with appropriate one-sided derivatives at the boundaries).*

3. *for any* $\theta_{i,t-1}$ *there is some* $\theta'$ *such that* $\frac{\partial \psi_i(\theta_{i,t}|\theta_{i,t-1})}{\partial \theta_{i,t-1}}|_{\theta_{i,t}=\theta'} = 0$ *and* $\frac{\partial \psi_i(\theta_{i,t}|\theta_{i,t-1})}{\partial \theta_{i,t-1}} < 0$ *for all* $\theta_{i,t} < \theta'$ *and* $\frac{\partial \psi_i(\theta_{i,t}|\theta_{i,t-1})}{\partial \theta_{i,t-1}} > 0$ *for all* $\theta_{i,t} > \theta'$. *Obviously it will also be the case that* $\int_{\underline{\theta_i}}^{\overline{\theta_i}} \frac{\partial \psi_i(\theta_{i,t}|\theta_{i,t-1})}{\partial \theta_{i,t-1}} d\theta_{i,t} = 0$.

This assumption implies types are positively, but not perfectly, serially correlated so that a higher type in one period implies that a higher type in the next period is more likely.

Beliefs about rivals' types play an important role in our game. In a fully separating equilibrium, each firm will (correctly) believe that each rival has a particular type in the previous period. For convenience, we assume that beliefs about types in $t = 1$ have the same structure.

**Assumption 2** ***Initial Period Beliefs.*** *Firms know what their rivals' types were in a fictitious prior period,* $t = 0$.

[7] Our fully separating equilibria would be unchanged if $t-2$ types were revealed.

## 2.4 Fully Separating Equilibrium in a Finite Horizon and Continuous Type Game.

We now describe the equilibrium for a game with two ex-ante symmetric single-product duopolists, which we will use in our first example.

### 2.4.1 Final Period ($T$).

In the final period, each firm maximizes its expected payoff given its own type, its beliefs about the types of the other firms and their pricing strategies. Play is therefore consistent with a Bayesian Nash Equilibrium. If firm $j$ believes that firm $i$'s period $T-1$ type was $\widehat{\theta_{i,T-1}^{j}}$ and $j$'s period $T$ pricing function is $P_{j,T}(\theta_{j,T}, \theta_{j,T-1}, \widehat{\theta_{i,T-1}^{j}})$[8], then a type $\theta_{i,T}$ $i$ will set a price

$$p_{i,T}^{*}(\theta_{i,T}, \theta_{j,T-1}, \widehat{\theta_{i,T-1}^{j}}) = \arg\max_{p_{i,T}} \int_{\underline{\theta_j}}^{\overline{\theta_j}} \pi(p_{i,T}, P_{j,T}(\theta_{j,T}, \theta_{j,T-1}, \widehat{\theta_{i,T-1}^{j}}), \theta_{i,T}) \psi(\theta_{j,T}|\theta_{j,T-1}) d\theta_{j,T}.$$

### 2.4.2 Earlier Periods ($1, .., T-1$).

In earlier periods, $i$ may choose not to set a static best response price in order to affect $j$'s belief about its type. The equilibrium concept that we use is symmetric Markov Perfect Bayesian Equilibrium (MPBE) (Roddie (2012), Toxvaerd (2008)). An MPBE specifies period-specific pricing strategies for each firm $i$ as a function of its current type, and its belief about $j$'s previous type, and $j$'s belief about $i$'s previous type; and, each firm's belief about its rival's type given observed histories of prices. Equilibrium beliefs should be consistent with Bayes Rule given equilibrium pricing strategies. If there are multiple rivals, they should all have the same beliefs given an observed history. While only current types and prices are directly payoff-relevant, history can matter in this Markovian equilibrium because it affects beliefs. We will only consider fully separating MPBEs where, in every period, a firm's equilibrium pricing strategy perfectly reveals its current type, and $j$'s belief about $i$'s current type will come from inverting $i$'s pricing function.

[8]This notation reflects the fact that we are assuming that player $j$ used an equilibrium strategy in $T-1$ that revealed its type ($\theta_{j,T-1}$), but we are allowing for the possibility that firm $i$ may have deviated so that $j$'s beliefs about $i$'s previous type are incorrect.

### 2.4.3 Characterization of Separating Pricing Functions in Period $t < T$.

We follow Mailath (1989), which shows that one can apply the results in Mailath (1987) to this problem, in characterizing fully separating pricing functions using a definition of firm $i$'s period-specific "signaling payoff function", $\Pi^{i,t}(\theta_{i,t}, \widehat{\theta^j_{i,t}}, p_{i,t})$. This is the present discounted value of firm $i$'s expected current and future payoffs when its current type is $\theta_{i,t}$, it sets price $p_{i,t}$ and $j$ believes, *at the end of period* $t$, that $i$ has type $\widehat{\theta^j_{i,t}}$. $\Pi^{i,t}$ is assumed to be continuous and at least twice differentiable in its arguments. It is implicitly conditional on (i) $j$'s period $t$ pricing strategy, which will depend on beliefs about types at $t-1$, and (ii) both players' strategies in future periods. As $j$'s end-of-period $t$ belief about $i$'s type enters as a separate argument, $p_{i,t}$ only affects $\Pi^{i,t}$ through period $t$ profits. Given conditions on $\Pi^{i,t}$ that will be listed in a moment, the fully separating best response function of firm $i$, which is also implicitly conditioned on $j$'s current pricing strategy and beliefs about previous types, can be uniquely characterized as follows (see Appendix C for a restatement of the Mailath (1987) theorems): $i$'s pricing function will be the solution to a differential equation where

$$\frac{\partial p^*_{i,t}(\theta_{i,t})}{\partial \theta_{i,t}} = -\frac{\Pi^{i,t}_2\left(\theta_{i,t}, \widehat{\theta^j_{i,t}}, p_{i,t}\right)}{\Pi^{i,t}_3\left(\theta_{i,t}, \widehat{\theta^j_{i,t}}, p_{i,t}\right)} > 0, \tag{1}$$

and a boundary condition. The subscript $n$ in $\Pi^{i,t}_n$ denotes the partial derivative of $\Pi^{i,t}$ with respect to the $n^{\text{th}}$ argument. Assuming that lower types want to set lower prices (e.g., a type corresponds to the firm's marginal cost), the boundary condition will be that $p^*_{i,t}(\underline{\theta_i})$ is the solution to

$$\Pi^{i,t}_3\left(\underline{\theta_i}, \widehat{\theta^j_{i,t}}, p_{i,t}\right) = 0, \tag{2}$$

i.e., the lowest type's price maximizes its static expected profits given $j$'s pricing policy. The numerator in (1) is $i$'s marginal future benefit from raising $j$'s belief about $\theta_{i,t}$, and the denominator is the marginal effect of a price increase on $i$'s current profit. For prices above a static best response price, the denominator will be negative, and the pricing function will slope upwards in the firm's type.

This characterization of a separating best response will be valid under four conditions on $\Pi^{i,t}$, in addition to continuity and differentiability,

**Condition 1** ***Shape of*** $\Pi^{i,t}$ ***with respect to*** $p_{i,t}$***.*** *For any* $(\theta_{i,t}, \widehat{\theta^j_{i,t}})$*,* $\Pi^{i,t}\left(\theta_{i,t}, \widehat{\theta^j_{i,t}}, p_{i,t}\right)$ *has a unique optimum in* $p_{i,t}$*, and, for all* $\theta_{i,t}$*, for any* $p_{i,t}$ *where* $\Pi^{i,t}_{33}\left(\theta_{i,t}, \widehat{\theta^j_{i,t}}, p_{i,t}\right) > 0$*, there is some* $k > 0$ *such that* $\left|\Pi^{i,t}_3\left(\theta_{i,t}, \widehat{\theta^j_{i,t}}, p_{i,t}\right)\right| > k$*.*

**Condition 2** ***Type Monotonicity.*** $\Pi^{i,t}_{13}\left(\theta_{i,t}, \widehat{\theta^j_{i,t}}, p_{i,t}\right) \neq 0$ *for all* $(\theta_{i,t}, \widehat{\theta^j_{i,t}}, p_{i,t})$*.*

**Condition 3** ***Belief Monotonicity.*** $\Pi^{i,t}_2\left(\theta_{i,t}, \widehat{\theta^j_{i,t}}, p_{i,t}\right)$ *is either* $> 0$ *for all* $(\theta_{i,t}, \widehat{\theta^j_{i,t}})$ *or* $< 0$ *for all* $(\theta_{i,t}, \widehat{\theta^j_{i,t}})$*.*

**Condition 4** ***Single-Crossing.*** $\frac{\Pi^{i,t}_3\left(\theta_{i,t}, \widehat{\theta^j_{i,t}}, p_{i,t}\right)}{\Pi^{i,t}_2\left(\theta_{i,t}, \widehat{\theta^j_{i,t}}, p_{i,t}\right)}$ *is a monotone function of* $\theta_{i,t}$ *for all* $\widehat{\theta^j_{i,t}}$ *and for* $(\theta_{i,t}, p_{i,t})$ *in the graph of* $p^*_{i,t}(\theta_{i,t}, \theta_{j,t-1})$*.*

To interpret these conditions, assume that types correspond to marginal costs. The first condition will be satisfied if, for any marginal cost and distribution of prices that the rival may set, a firm's expected current period profit is quasi-concave in its own price. This will hold for common forms of differentiated product demand such as the multinomial and nested logit models. Type monotonicity requires that, when a firm increases its price, the profit that it loses will be lower if it has higher marginal costs. This will hold for constant marginal costs. Belief monotonicity requires that a firm's expected future profits should increase when rivals believe that it has a higher cost, holding its actual cost fixed. This condition may fail: Appendix B.1 discusses in detail a two-type example where $j$ will respond to $i$ having a higher cost by setting a lower price in the next period. The single-crossing condition requires that a firm with a higher marginal cost should always be more willing to raise its price, reducing its current profits, in order to raise its rival's belief about its marginal cost. This condition can also fail.

For completeness, we also need to define beliefs that a firm will have if the rival sets a price that is outside the range of the pricing function (i.e., a price that is not on the equilibrium path). When types correspond to marginal costs, we will assume that when a firm sets a price below (above) the lowest (highest) price in the range of the pricing function, it will be inferred to have the lowest (highest) possible cost type.

### 2.4.4 Existence and Uniqueness of a Fully Separating Equilibrium.

The conditions defined above guarantee the existence and uniqueness of fully separating best responses in any period, but this does not prove the existence or uniqueness of a fully separating equilibrium in the whole game. Mailath (1989) proves existence and uniqueness in a two-period duopoly game with linear demand and there is private information about marginal costs. Appendix C shows the existence and uniqueness in a finite horizon, linear demand duopoly game where marginal costs are private information. The proof requires that the marginal cost interval $(\overline{\theta} - \underline{\theta})$ is small enough so that a single-crossing condition holds when prices rise.

In our application, we will assume nonlinear demand and, to reduce the computational burden, an infinite horizon. We will therefore proceed without proofs of existence or uniqueness. Appendix A details how we compute equilibrium strategies, and verify belief monotonicity and single-crossing as part of the algorithm. We will discuss examples where we cannot find a separating equilibrium below. We have only ever found a single equilibrium in finite horizon games and infinite horizon games with continuous types, but we have found examples of multiplicity in infinite horizon games with two types even when, as we describe below, we impose a refinement that is needed to guarantee unique best response functions in that case.[9]

# 3 Examples

This section uses examples to illustrate the equilibrium of our game and the effects of a merger. Additional examples described in Appendix B are also discussed.

## 3.1 Continuous-Type Duopoly Example.

### 3.1.1 Specification.

There are two ex-ante symmetric single-product firms. Demand is determined by a nested logit model, with both products in one nest, and the outside good in its own nest. Consumer $c$'s indirect utility from buying from product $i$ is $u_{i,c} = 5 - 0.1p_i + \sigma\nu_c + (1-\sigma)\varepsilon_{i,c}$ where $p_i$ is

[9] In examples where we have found multiplicity, the algorithm that we use elsewhere in the paper appears to consistently pick out an equilibrium that is the limit of the equilibrium in the early periods of a finite horizon game as the number of periods grows.

the dollar price, $\varepsilon_{i,c}$ is a draw from a Type I extreme value distribution, $\sigma = 0.25$, and $\nu_c$ is an appropriately distributed draw for $c$'s nest preferences. For the outside good, $u_{0,c} = \varepsilon_{0,c}$. We will set market size equal to 1, so that our welfare numbers have a "per-consumer" interpretation. We first examine what happens to strategies in a finite horizon game with $T = 25$ periods. The game is solved backwards, starting at the last period.

We assume that marginal cost is private information, and that, for each firm, it lies in the interval $[\underline{c}, \overline{c}] = [\$8, \$8.05]$. Costs evolve independently according to an exogenous truncated AR(1) process where

$$c_{i,t} = \rho c_{i,t-1} + (1-\rho)\frac{\overline{c} + \underline{c}}{2} + \eta_{i,t} \tag{3}$$

where $\rho = 0.8$ and $\eta_{i,t} \sim TRN(0, \sigma_c^2, \underline{c} - \rho c_{i,t-1} - (1-\rho)\frac{\overline{c}+\underline{c}}{2}, \overline{c} - \rho c_{i,t-1} - (1-\rho)\frac{\overline{c}+\underline{c}}{2})$, where $TRN$ denotes a truncated normal distribution, and the first two arguments are the mean and variance of the untruncated distribution, and the third and fourth arguments are the lower and upper truncation points. $\sigma_c = \$0.025$.

Two features of this parameterization are worth highlighting. First, marginal costs are restricted to a narrow range (diverging by less than 0.32% from mean value) and the probability that a firm will switch from a relatively high cost to a relatively low cost across periods is quite high.[10] Therefore, no signal should affect a rival's posterior belief about a firm's next period marginal cost very much. Despite this, we find large signaling effects. Second, the demand parameters imply high margins and limited substitution to the outside good in both static and dynamic equilibria. As we will discuss, these features contribute to the existence of a fully separating equilibrium with large price effects.

#### 3.1.2 Equilibrium Outcomes and Strategies.

Table 1 shows expected price levels, the standard deviation of prices and various welfare measures when we simulate data using equilibrium strategies in different periods of the finite horizon game. For comparison, expected joint-profit maximizing prices and static Nash equilibrium prices under CI (given average costs) are \$45.20 and \$22.62, with small standard deviations (\$0.007 and \$0.011). Signaling MPBE prices are higher and significantly more volatile than

[10] For example, the probability that a firm with the highest marginal cost has a cost in the lower half of the support in the next period is 0.32.

Table 1: Equilibrium Prices and Welfare in the Duopoly Game

| | | | | Expected Welfare Measures Per Market Size Unit | | |
|---|---|---|---|---|---|---|
| Period | Nature of Equilibrium | Mean Price | Std. Dev. Price | Cons. Surplus | Producer Surplus | Total Welfare |
| T-24 | MPBE | \$24.76 | \$0.47 | \$30.91 | \$15.96 | \$46.87 |
| T-13 | MPBE | \$24.76 | \$0.47 | \$30.91 | \$15.96 | \$46.87 |
| T-10 | MPBE | \$24.75 | \$0.47 | \$30.92 | \$15.95 | \$46.87 |
| T-7 | MPBE | \$24.68 | \$0.45 | \$30.98 | \$15.89 | \$46.88 |
| T-4 | MPBE | \$24.25 | \$0.36 | \$31.40 | \$15.51 | \$46.91 |
| T-2 | MPBE | \$23.38 | \$0.17 | \$32.23 | \$14.74 | \$46.97 |
| T-1 | MPBE | \$22.88 | \$0.06 | \$32.71 | \$14.29 | \$47.00 |
| T | BNE | \$22.62 | \$0.01 | \$32.96 | \$14.05 | \$47.01 |
| Infinite Horizon | Stationary MPBE | \$24.76 | \$0.47 | \$30.91 | \$15.96 | \$46.87 |

Notes: except for the last row, all prices are based on equilibrium strategies in a finite horizon model with parameters described in the text. The last line reports results for the stationary strategies in an infinite horizon model with the same parameters.

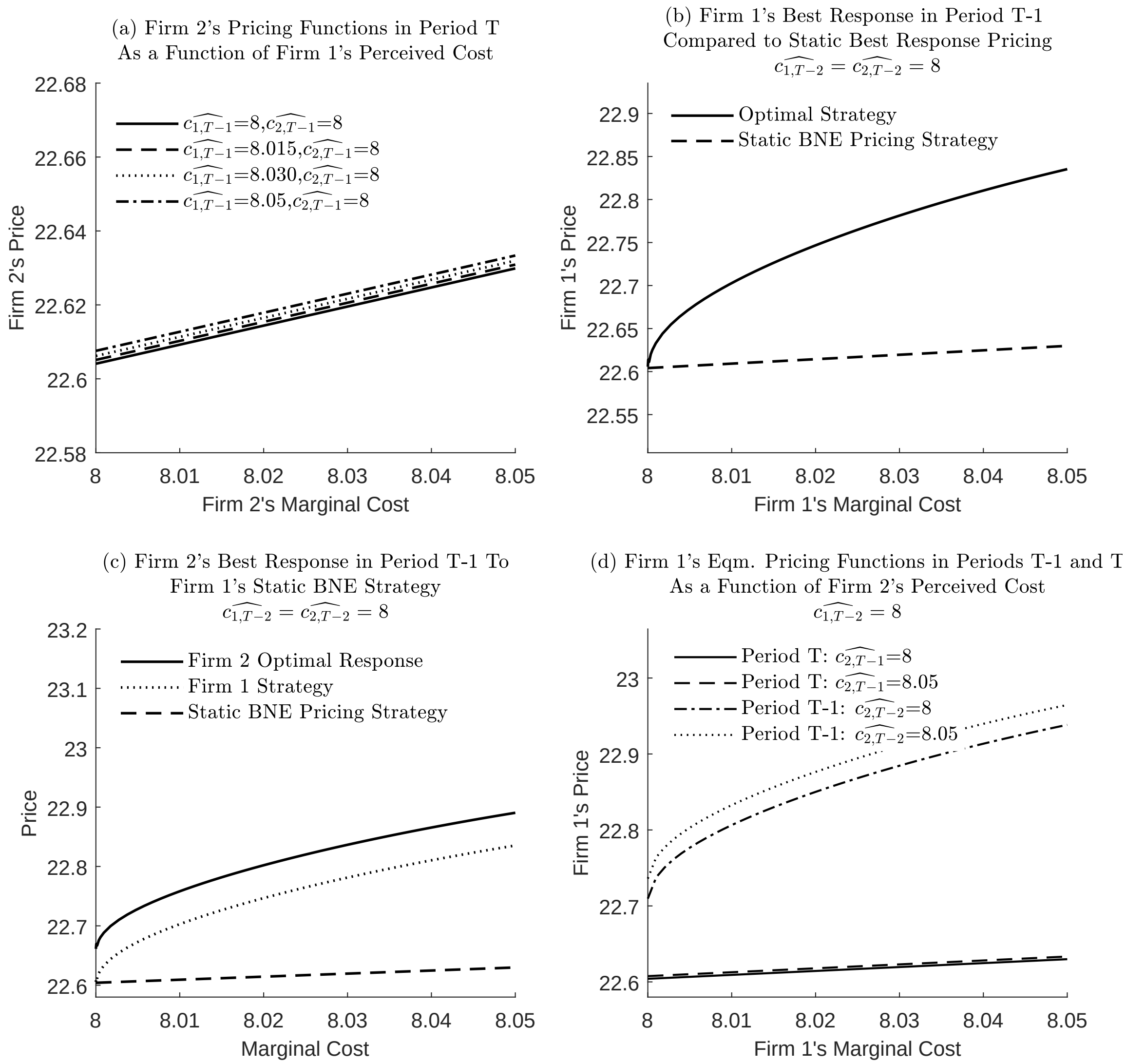
Nash prices when the game is more than a couple of periods from the end, but they are always much lower than joint profit-maximizing prices. We now describe the strategies that result in these outcomes.

Figure 1(a) shows four static BNE period $T$ pricing functions for firm 2, for different values of firm 1's period $T-1$ marginal cost ($c_{1,T-1}$), assuming that both firms know/believe that $c_{2,T-1} = \$8$. Firm 2's price increases with $c_{1,T-1}$ as firm 1's expected period $T$ price rises with $c_{1,T-1}$. However, the variation in firm 1's prior cost affects firm 2's price by less than one cent, and, averaging across all possible cost realizations, average prices and welfare are almost identical to outcomes with CI.[11] Therefore the existence of asymmetric information alone (i.e., when not combined with some form of dynamics) does not generate interesting effects given our parameters.

There is an incentive to signal in period $T-1$ because a firm's price can affect its rival's price in period $T$. Assuming both firms' period $T-2$ costs were \$8, Figure 1(b) shows firm 1's signaling pricing function (found by solving the differential equation in (1) given the boundary condition (2)) if it expected that firm 2 was using its period $T$ strategy. We reproduce the

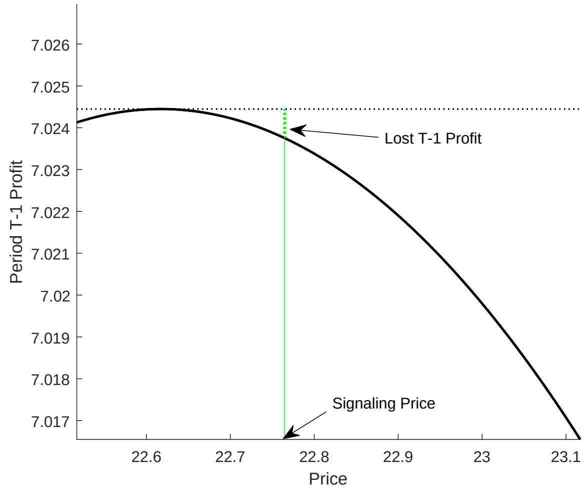
[11]Expected producer and consumer surplus differ by less than \$0.0001 across these models.

Figure 1: Period $T$ and $T-1$ Pricing Strategies in the Finite Horizon, Continuous Type Signaling Game




period $T$ pricing strategy for comparison. The pricing functions intersect for $c_{1,T-1} = \$8$, but signaling may lead firm 1 to raise its price by as much as 20 cents for higher costs. At first blush, this large increase may seem surprising given that we know the effect on firm 2's price can only be small. However, the assumed demand implies that firm 1's profit function, shown in Figure 2, is sufficiently flat that, if $c_{1,T-1} = \$8.025$, its expected lost period $T-1$ profit from using a signaling price of \$22.76, rather than the statically optimal period $T-1$ price of \$22.61, is only \$0.00070 per consumer, which is less than the (discounted) expected period $T$ profit gain of \$0.00079 from being viewed as a firm with a $c_{1,T-1} = \$8.025$ rather than $c_{1,T-1} = \$8.0001$ (which is how firm 2 would interpret a price of \$22.61).

Figure 2: Expected $T-1$ Period Profit Function: $c_{1,T-1} = \$8.025$ and $c_{1,T-2} = c_{2,T-2} = \$8$



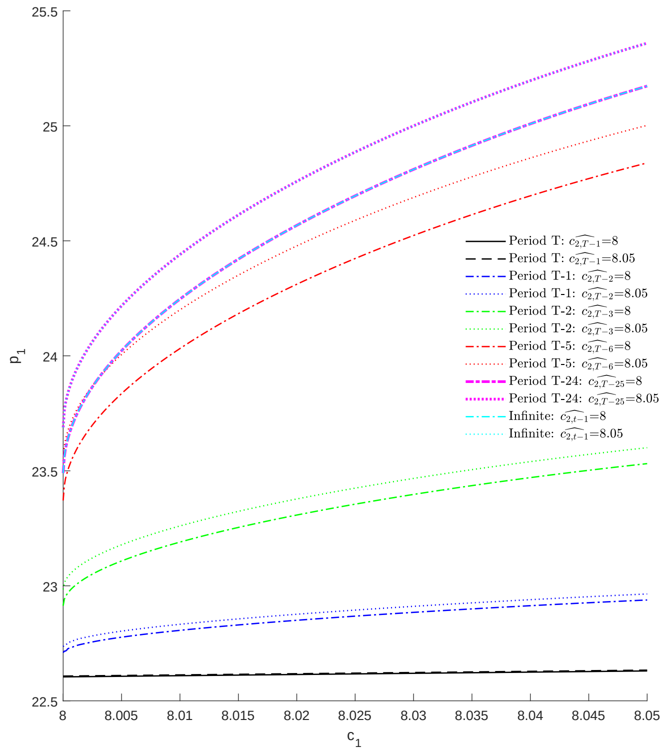

Notes: the profit function is drawn "per potential consumer" for a firm assumed to have a marginal cost of \$8.025, and with a rival using the static BNE pricing strategy when both firms' previous period marginal costs were \$8.

Figure 1(b) assumed that firm 2 was using its period $T$ strategy with no signaling. Figure 1(c) shows firm 2's best signaling response when firm 1 uses the strategy in Figure 1(b) (repeated in the new figure as a comparison). As firm 1's expected price has increased, firm 2's static best response pricing function shifts upwards. Of course, this positive feedback will cause firm 1's pricing function to rise as well. Figure 1(d) shows the equilibrium period $T-1$ pricing functions. The increase in the slope and the dispersion of the pricing functions means that period $T-1$ prices will be higher and more volatile than period $T$ prices.

The increased vertical spread also means that period $T-1$ prices are more sensitive to perceived period $T-2$ costs which increases period $T-2$ signaling incentives. Figure 3 shows a selection of equilibrium pricing functions for period $T-2$ and earlier periods. The pricing functions become more spread out and the level of prices increases, although by successively smaller amounts, in earlier periods. Further back than period $T-15$ equilibrium pricing functions and average prices barely change. The figure also plots the stationary pricing strategies that we compute for an infinite horizon game with the same parameters. They are indistinguishable from the strategies in the early periods of the finite horizon game.[12]

[12] We have consistently found this convergence except in cases when the conditions required for separation are violated or are very close to being violated (in which case the infinite horizon strategies may not converge).

Figure 3: Equilibrium Pricing Functions for Firm 1 in the Infinite Horizon Game and Various Periods of the Finite Horizon Game.




Notes: all functions drawn assuming that firm 1's perceived marginal cost in the previous period was \$8.

## 3.2 Merger Analysis.

There are many possible applications of our model, but we will focus on its predictions for horizontal mergers. We present a simple motivating example using the infinite horizon, continuous cost model with the same demand and marginal cost parameters that we have just assumed, although we will allow for more firms. We will assume that a merger occurs as an unanticipated one-off shock, i.e., firms signal assuming the prevailing market structure will last forever. As discussed in a two-type example with up to seven firms in Appendix B.1.3, signaling tends to have more effects on pricing when there are fewer firms, because each firm's price will tend to have a larger effect on its rivals' next period prices.

Table 2 shows the effects of 4-to-3 and 3-to-2 mergers. Before either merger, there are symmetric single-product firms. In the upper panel, we assume that a merger eliminates a product, so that after the merger there are only single-product firms with symmetric demand.

Table 2: The Effects of Signaling on Mergers and Merger Analysis When Firms Use Infinite Horizon Signaling Strategies

(a) Merger Leads to the Elimination of a Product By the Merged Firm

| | 4-to-3 Merger | 3-to-2 Merger |
|---|---|---|
| Signaling MPBE | | |
| Pre-Merger Average Price | \$18.25 | \$19.79 |
| Post-Merger Average Price of Merged Firm if No Marginal Cost Synergy | \$19.81 (+8.5%) | \$24.75 (+25.1%) |
| Post-Merger Average Price of Non-Merging Firm if No Marginal Cost Synergy | \$19.81 (+8.5%) | \$24.75 (+25.1%) |
| Merged Firm Marginal Cost Required to Prevent Merged Firm Average Price from Rising | \$5.73 | -\$2.20 |
| If Merger Analyzed under Complete Information ... | | |
| Implied Pre-Merger Average Marginal Cost | \$8.29 | \$8.62 |
| Merged Firm Marginal Cost Required to Prevent Prices from Rising | \$7.11 | \$5.13 |
| Average Merged Firm Price in Signaling Model if Analyst Required Marginal Cost is Realized | \$19.17 (+5.0%) | \$23.25 (+17.4%) |

(b) Merging Firm Owns Two Products Post-Merger

| | 4-to-3 Merger | 3-to-2 Merger |
|---|---|---|
| Signaling MPBE | | |
| Pre-Merger Average Price | \$18.25 | \$19.79 |
| Post-Merger Average Price of Merged Firm if No Marginal Cost Synergy | \$21.53 (+18.0%) | \$27.18 (+37.3%) |
| Post-Merger Average Price of Non-Merging Firm if No Marginal Cost Synergy | \$19.12 (+4.8%) | \$23.59 (+19.2%) |
| Merged Firm Marginal Cost Required to Prevent Merged Firm Average Price from Rising | \$2.26 | -\$11.92 |
| If Merger Analyzed under Complete Information ... | | |
| Implied Marginal Cost | \$8.29 | \$8.62 |
| Merged Firm Marginal Cost Required to Prevent Prices from Rising | \$3.43 | -\$2.05 |
| Average Price in Signaling Model if Analyst Required Marginal Cost is Realized | \$18.85 (+3.2%) | \$23.00 (+16.2%) |

If we assume that the firms always use equilibrium signaling strategies, then a 4-to-3 merger with no synergy (implying the firms remain symmetric post-merger) will raise average prices by 8.5%. To prevent the merged firm's price from rising, the merger would need to reduce the average marginal cost of the merging firm from \$8.025 to \$5.73, a 29% reduction.[13]

We can compare these effects to the predictions of an analyst who knows demand and uses a standard CI merger simulation model.[14] Using average prices and CI first-order conditions, the analyst would infer that average pre-merger marginal costs are equal to \$8.29 (i.e., higher than they really are), and that a 14% synergy (reducing marginal costs to \$7.11) would prevent price increases. If the 14% synergy was achieved but firms play a signaling equilibrium after the merger, then the merged firm's average price would increase to \$19.17 (a 5% post-merger increase). In the case of a 3-to-2 merger, all of the effects seen in the 4-to-3 case become larger, and, in fact, the merged firm's marginal cost would need to be negative to prevent a price increase.[15] The realization of the synergy identified by a CI simulation would not prevent prices from rising by 17%.

The lower panel assumes that, after the merger, the merged firm has two products, which have the same marginal cost and which are sold at the same price. This restrictive assumption preserves the structure that each firm has one piece of private information and can send exactly one signal. Ownership of two products increases incentives to raise prices, and hence the size of required synergies. As in the upper panel, a CI analysis will underpredict price increases and required synergies.

### 3.3 Signaling Incentives and the Existence of Separating Equilibria.

In our example, signaling incentives are relatively weak because marginal costs are only weakly correlated from period to period. Increasing the AR(1) parameter or $\overline{c}-\underline{c}$, or reducing $\sigma_c$ tend to increase signaling incentives and raise equilibrium prices. However, when price increases are too large, the conditions for characterizing best responses can fail and we may not be able to

[13]We assume that the range of marginal costs, \$0.05, and the process by which marginal costs evolve remain the same after the merger and after any synergy is realized.

[14]This characterization follows how merger simulation is used in the academic literature. Agency economists typically calibrate the price and nesting parameters in the demand system to match average margins given CI Nash pricing. An incorrect static CI Nash assumption would then lead to the wrong demand parameters.

[15]We assume that the firm cannot freely dispose of products so that it cannot choose to produce an infinite amount if it has negative marginal costs.

Table 3: Equilibrium Pricing in a Finite Horizon Game with Alternative Cost Specifications

| | Baseline | Expand Range | | | Reduce Std. Deviation | | Expand Range & Increase Std. Dev. |
|---|---|---|---|---|---|---|---|
| $\bar{c} - \underline{c}$ (\$) | [8,8.05] | [8,8.075] | [8,8.15] | [8,8.3] | [8,8.05] | [8,8.05] | [8,8.50] |
| $\sigma_c$ (\$) | 0.025 | 0.025 | 0.025 | 0.025 | 0.02 | 0.01 | 0.25 |
| T-24 | \$24.76 | \$26.51 | - | - | \$25.71 | - | \$24.90 |
| T-10 | \$24.75 | \$26.59 | - | - | \$25.70 | - | \$24.89 |
| T-9 | \$24.74 | \$26.59 | fails | - | \$25.69 | fails | \$24.89 |
| T-8 | \$24.72 | \$26.57 | \$28.48 | - | \$25.66 | \$28.58 | \$24.89 |
| T-7 | \$24.68 | \$26.50 | \$29.17 | fails | \$25.60 | \$28.76 | \$24.87 |
| T-6 | \$24.61 | \$26.37 | \$29.35 | \$30.40 | \$25.49 | \$28.65 | \$24.85 |
| T-1 | \$22.88 | \$23.05 | \$23.42 | \$23.93 | \$22.93 | \$23.05 | \$23.55 |
| T | \$22.62 | \$22.63 | \$22.67 | \$22.74 | \$22.62 | \$22.62 | \$22.84 |
| $\infty$-Horizon | \$24.76 | \$26.50 | fails | fails | \$25.71 | fails | \$24.90 |

Notes: values in all but the last line are based on the duopoly, continuous type, finite horizon model with demand parameters described in the text (cost parameters indicated in the table). The last line reports results for the stationary strategies in the infinite horizon model with the same parameters. "Fails" indicates that the belief monotonicity or single-crossing conditions fail so that we cannot calculate signaling best response pricing functions.

find a separating equilibrium.

The first six columns of Table 3 show, for different periods, the baseline average prices and average prices when signaling incentives are strengthened. Small parameter changes result in higher equilibrium prices, but larger changes result in the failure of our algorithm as we cannot define best response pricing functions. Pooling or partial pooling equilibria may exist, but we do not know how to characterize them. Appendix B.1.2 uses a two-type example to examine the failure of the conditions, including belief monotonicity, in more detail.[16]

However, as illustrated in the final column, we can sustain separating equilibria if we increase $\bar{c} - \underline{c}$ and increase $\sigma_c$ simultaneously.[17] This pattern will be relevant for our application.

[16]The two-type model has a much lower computational burden but requires imposing a refinement just to identify unique separating best responses. Specifically, we always find the best response that achieves separation at the lowest cost to the signaling firm, consistent with the type of "intuitive criterion" (Cho and Kreps (1987)) refinement that has been widely used in one-sided signaling models with two types. However, even with this refinement, we have found examples of multiple separating equilibria in the infinite horizon version of the two-type model. The algorithm that we use to produce the reported results appears to consistently select the equilibrium that corresponds to the limit of a (seemingly unique) equilibrium in a finite horizon game as the number of periods grows large.

[17]The probability that a cost goes from one extreme of the support to the opposite half of the support is 0.32, which is the same as in the baseline case.

### 3.4 Additional Examples.

Appendix B.1 uses two-type duopoly examples to examine how price effects vary with the number of firms and to examine the relationship between the existence of separating equilibria, the magnitude of price effects, the serial correlation of costs and the extent to which, when a firm's price rises, demand is diverted to the outside good. When there is limited diversion to the outside good we find large increases in prices above static CI Nash levels (an increase of 45% in one case) under duopoly even when there is moderate serial correlation in costs (e.g., $\Pr(c_{i,t} = c_{i,t-1})$=0.75). On the other hand, price increases are small with more than three firms, and only small increases can be sustained in separating equilibria when there is more diversion to the outside good even under duopoly. The examples suggest that thinking about the effects of signaling is most relevant when two or three firms dominate a market or a very distinct segment of a market.

Appendix B.2 present three simple duopoly examples where marginal costs are fixed and known, but firms have private information about some other element of their payoff function (a feature of demand, the weight managers place on revenues rather than profits, or the weight they place on the profits of rivals). Signaling can raise prices significantly above CI Nash levels in each case.

# 4 Empirical Application: The MillerCoors Joint Venture

In this section, we apply our model to data from the U.S. beer industry around the time of the 2008 MC JV. MW show that, relative to the price of imports, the real prices of brands owned by MC and Anheuser-Busch (AB) increased after the JV.[18] We describe the setting and the data, before explaining the calibration of our model using pre-JV pricing data and reporting how well it predicts observed changes in pricing after the JV. Finally, we examine how well the CI models that have previously been used to explain why price increased fit the data.

[18] Anheuser-Busch was purchased by InBev in 2008. Throughout the paper we will use AB to refer to Anheuser-Busch before 2008 and Anheuser-Busch InBev afterwards.

## 4.1 The JV and Its Effects.

The MC JV, announced in October 2007, effectively merged the U.S. brewing, marketing and sales operations of SABMiller (Miller) and MolsonCoors (Coors), the second and third largest U.S. brewers. The Department of Justice (DOJ) decided not to challenge the transaction in June 2008 because it expected "large reductions in variable costs of the type that are likely to have a beneficial effect on prices".[19] For example, the JV was expected to lower transportation costs by producing Coors products at Miller breweries around the country. Ashenfelter, Hosken, and Weinberg (2015) provide evidence that transportation efficiencies were realized.

MW show that, at a national level, the real prices (i.e., deflated by the CPI-U price index) of the most popular domestic brands, such as Bud Light (BL), Miller Lite (ML) and Coors Light (CL), increased after the JV, relative to the prices of imported brands, such as Corona Extra and Heineken, which MW use as controls for industry-wide cost shocks. Regressions in Appendix D quantify these price increases to lie between 40 cents and a dollar per 12-pack, or 3%-6%, depending on the specification. We will proceed assuming that MW's interpretation that the relative price increase was a causal anticompetitive effect of the JV is correct.[20]

An important feature of the relative price change is that AB's prices increased as much as those of Miller and Coors. If AB's marginal costs were unaffected by the JV, this pattern is inconsistent with static CI Nash pricing, as a static best response function would predict that AB should have responded to any JV price increase by raising its prices by a smaller amount.

## 4.2 Data.

We use the same data as MW, which comes from the IRI Academic Dataset (Bronnenberg, Kruger, and Mela (2008)) which provides weekly UPC-store-level scanner data for the beer category from an unbalanced panel of grocery stores from 2001 to 2011. Appendix D provides details, but we will note in the text where our treatment differs from MW. We will follow the typical convention of assuming that retail prices are set directly by brewers, and that any retail

[19] Department of Justice press release, 5 June 2008.

[20] This interpretation is complicated by how the Great Recession may have affected demand and the fall in the deflator, from 220.0 in July 2008 to 210.2 in December 2008, at exactly the same time that the merger was being consummated.

Table 4: Highest-Selling Beer Brands in 2007 with Ownership, Share and Average Nominal Prices per 12-Pack.

| | | | 2007 | | | 2011 | |
|---|---|---|---|---|---|---|---|
| Brand | Company | Packs | % 18+ | Mkt. Share | Price | Mkt. Share | Price |
| Bud Light*,† | AB | 10 | 72.5% | 15.7% | $8.29 | 15.7% | $8.92 |
| Miller Lite*,† | M | 10 | 75.1% | 10.0% | $8.11 | 8.4% | $8.73 |
| Coors Light*,† | C | 10 | 74.8% | 8.3% | $8.36 | 9.4% | $8.98 |
| Budweiser† | AB | 10 | 70.8% | 7.7% | $8.30 | 6.5% | $9.00 |
| Corona Extra†,◊ | GM | 5 | 15.6% | 4.1% | $13.88 | 3.9% | $13.46 |
| Natural Light* | AB | 7 | 68.6% | 3.9% | $6.01 | 3.2% | $7.15 |
| Busch Light* | AB | 9 | 78.4% | 2.8% | $6.07 | 2.5% | $6.96 |
| Miller High Life† | M | 9 | 54.1% | 2.4% | $6.33 | 2.2% | $7.21 |
| Heineken†,◊ | H | 7 | 12.8% | 2.3% | $14.06 | 2.3% | $13.86 |
| Miller Genuine Draft† | M | 10 | 67.0% | 2.3% | $8.26 | 1.3% | $8.94 |
| Michelob Ultra*,† | AB | 9 | 27.4% | 2.1% | $10.05 | 2.4% | $10.51 |
| Busch | AB | 9 | 70.0% | 1.9% | $6.08 | 1.6% | $7.05 |
| Keystone Light* | C | 6 | 81.4% | 1.4% | $5.83 | 1.5% | $7.03 |
| Budweiser Select | AB | 9 | 62.0% | 1.3% | $8.37 | 0.7% | $8.76 |
| Milwaukee's Best Light* | M | 6 | 66.8% | 1.3% | $5.37 | 0.8% | $6.19 |
| Corona Light*,†,◊ | GM | 3 | 2.3% | 1.2% | $14.23 | 1.3% | $13.79 |
| Tecate◊ | H | 7 | 66.3% | 1.2% | $8.65 | 1.2% | $9.04 |
| Natural Ice | AB | 7 | 51.3% | 1.1% | $5.96 | 0.9% | $7.19 |
| Pabst Blue Ribbon | SP | 9 | 49.3% | 1.0% | $6.26 | 1.4% | $7.53 |
| Milwaukee's Best | M | 5 | 61.8% | 0.8% | $5.46 | 0.4% | $6.46 |
| Coors† | C | 10 | 73.3% | 0.8% | $8.44 | 1.0% | $8.84 |
| Michelob Light*,† | AB | 7 | 29.3% | 0.7% | $9.76 | 0.3% | $10.72 |
| Heineken Prem. Light*,†,◊ | H | 5 | 1.9% | 0.6% | $14.28 | 0.5% | $14.18 |

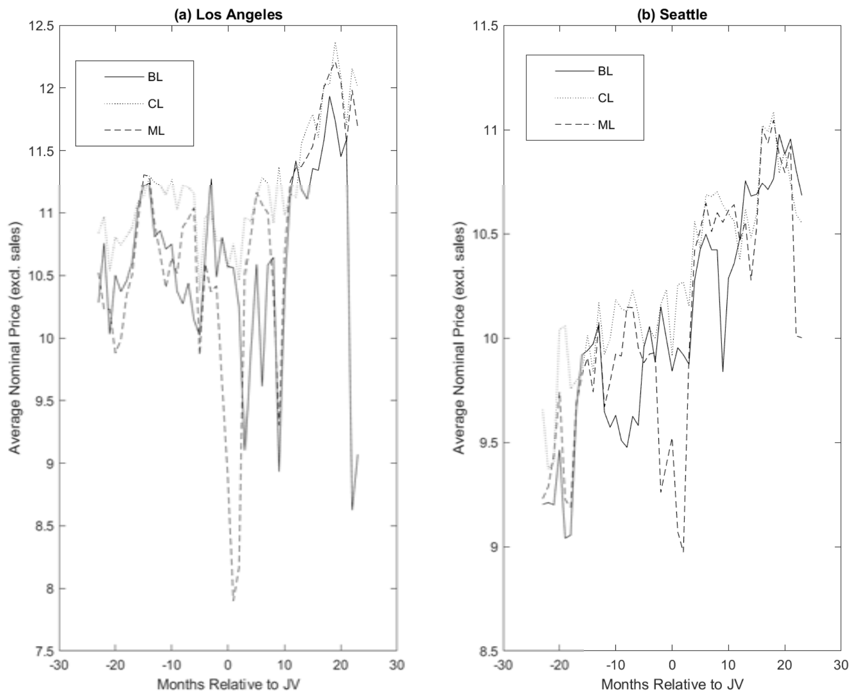
Notes: the table lists the 20 highest-selling brands plus additional brands in MW's sample. Market shares and prices are based on all units sold in packs equivalent to 6, 12, 18, 24, 30 and 36 12oz servings. "Packs" is the number of 2007 bottle/can-pack size combinations for 6, 12, 18, 24 and 30 packs, as 36 packs are rare. "% 18+ " is the percentage of 2007 volume sold in the packs of more than 18 cans or bottles. 2007 companies are: AB=Anheuser-Busch, M=SABMiller, C=MolsonCoors, GM=Grupo-Modelo, H=Heineken, SP=S&P. Prices are nominal prices per 12-pack equivalent (i.e., total dollars sold in all pack sizes divided by total volume in 144oz. units). *=light beers, †=included in MW's sample, ◊=imports.

margin is a fixed component of brewers' marginal costs.[21]

Table 4 lists the 20 brands with the largest sales by volume in 2007, together with additional brands that MW include in their analysis. The table lists market shares and average nominal prices (per 144 oz, the volume in a standard 12-pack) in 2007 and 2011. Most domestic brands are differentiated from imports by being sold primarily in larger packs and at lower prices. The relative prices of domestic brands increased after 2007, but, although CL gained share at ML's expense, the domestic brewers' market shares remained stable: for example, AB's volume share

[21]MW estimate a model that allows for a monopolist retail margin, and cannot reject a model with fixed retail pass-through.

Figure 4: Average Nominal Prices (excluding sales) of 12-Packs of the Domestic Flagship Brands in Two Regional Markets Around the JV.




Notes: averages are calculated as the total dollar sales of 12-packs at prices not identified as temporary price reductions, divided by the number of 12-packs sold. See Appendix D for the same figure with real prices.

was 41.3% in 2007, 41.5% in 2009 and 39.6% in 2011, with light beer shares of 50.0%, 50.8% and 50.6% respectively.[22]

We calibrate the model to match observed pre-JV dynamics of BL, CL and ML prices. As an example of the dynamics in the data, Figure 4 shows monthly average nominal prices of 12-packs for the flagship domestic brands in two large markets for 49 months around the consummation of the JV.[23] Average prices are calculated excluding all sales at prices that IRI indicates are temporary price reductions, as changes in regular prices are more likely to reflect changes in wholesale prices. Within-year volatility is a clear feature of this data, even if we ignore the drop in ML prices during the DOJ's investigation.

[22]Appendix D presents a figure showing the evolution of market shares over this period. The post-JV decline in the shares of several non-flagship domestic brands reflected a continuation of pre-existing trends.

[23]We use nominal prices so that they are not distorted by fluctuations in the CPI-U deflator, including the drop referenced in footnote 20. See Appendix D for the same figure plotted using real prices.

## 4.3 Calibration of the Dynamic Asymmetric Information Model.

We calibrate an infinite horizon, continuous marginal cost three-firm/product version of our model using pre-JV data, and then compare its predictions with post-JV data. We say "calibration", even though we estimate five cost parameters, because of the strong assumptions we make to limit the computational burden. The most important simplification is that our calibration will treat data from different markets as data from independent repetitions of the same game, rather than reflecting markets with different demand and cost primitives.

### 4.3.1 Products.

We model the pricing of three brands. We label these brands as BL, ML and CL, and will estimate the cost parameters to match the observed price of dynamics of these flagship products. However, Appendix D shows that the prices of brands in the same portfolio (e.g., Budweiser and BL) are highly correlated, and one can also view the brands as representing the portfolios of AB, Miller and Coors. Products of other brewers, including imports and craft beers, are included in the outside good.[24] We will assume that ML and CL are symmetric before the JV, as we will have to assume that MC sets the same price for both products after the JV. Appendix D also shows the correlation of ML and CL prices increased after the JV.

### 4.3.2 Demand.

We assume static, time-invariant nested logit demand, with the three brands in the same nest. The parameters are the nesting and price parameters, and the mean utilities (excluding the effect of price) of BL and ML/CL. Our baseline parameters are chosen so that, at average real prices in the pre-JV data, the average own price elasticity is -3, the market shares of the three products are 28% for BL and 14% each for ML/CL and, on average, if the price of one brand increased, 85% of the demand that it loses would go to the other brands (with the remainder to the outside good).[25] When we use weekly data on 6/12/18/24/30-packs and exclude temporary price reductions, the pre-JV cross-market average prices are \$10.09 for BL

[24] In an earlier version, we estimate the model allowing for imports to be a non-signaling fringe that used Bayesian Nash pricing. The model predicted that, after the JV, they would raise their prices by a couple of cents.

[25] These assumed shares overstate the share of BL relative to ML and CL, but understate the share of AB, relative to Miller and Coors, in the beer market and the light beer segment.

and \$9.95 for ML/CL, and the implied nesting and price parameters are 0.772 and −0.098, and the BL and ML/CL mean utilities are 1.044 and 0.863 respectively.

As motivation for the assumed elasticity and diversion, Table 5 reports five sets of demand estimates (the first three will be used in Section 4.4). For these specifications, we follow MW as closely as possible in the choice of data, instruments and controls, except that we use optimal GMM for the nested logit models as doing so affects the estimates.[26] The first three columns contain one nested logit specification, using monthly data, and two random coefficients nested logit (RCNL) specifications, where the 13 MW brands are all included in a single inside nest, and preferences vary with income. The remaining columns estimate nested logit models using monthly and weekly data (we will use weekly price changes when estimating the cost parameters) where flagship products are grouped into a flagship nest, and the remaining products are placed in an "other beer" nest with a different nesting coefficient. The flagship nesting coefficients are larger, consistent with these brands being close substitutes.

The table reports several implied statistics for each specification, including the average ML brand elasticity (i.e., the effect on demand when all ML prices increase), the proportion of lost demand that switches to other flagship products when a flagship price is increased, and the average, across pre-JV observations, predicted change in flagship sales when the prices of all domestic products increase by 75 cents, which is within the range of the observed post-JV price change. The statistics vary across the specifications. Given that the limited decline in flagship brand and domestic brewer market shares after the JV, we assume values for elasticity and diversion that are consistent with the estimates in columns (4) and (5).

#### 4.3.3 Marginal Costs.

We assume that the marginal costs of product $i$, $c_{i,t}$, lie on the interval $[\underline{c_i}, \underline{c_i} + c']$, where we estimate $\underline{c_{BL}}$, $\underline{c_{ML/CL}}$ and $c'$. $c_{i,t}$ evolves according to an AR(1) process with truncated innovations

$$c_{i,t} = \rho c_{i,t-1} + (1-\rho)\frac{\underline{c_i} + \underline{c_i} + c'}{2} + \eta_{i,t} \tag{4}$$

[26]None of the specifications yield exactly the same estimates as MW although the monthly RCNL estimates are almost identical.

Table 5: Estimates of Demand

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | Nested Logit | RCNL | RCNL | Nested Logit | Nested Logit |
| Nests | All Beer | All Beer | All Beer | Flagship/Other | Flagship/Other |
| Data Freq. | Monthly | Monthly | Quarterly | Monthly | Weekly |
| Real Price Coefficient | -0.056 | -0.083 | -0.099 | -0.073 | -0.047 |
| (2010 dollars) | (0.017) | (0.014) | (0.014) | (0.018) | (0.011) |
| Nesting Coefficients | | | | | |
| Single All Brand Nest | 0.741 | 0.838 | 0.831 | - | - |
| | (0.051) | (0.039) | (0.039) | | |
| Two Nests | | | | | |
| Domestic Flagship | - | - | - | 0.838 | 0.898 |
| | | | | (0.049) | (0.040) |
| Other Brands | - | - | - | 0.634 | 0.815 |
| | | | | (0.047) | (0.037) |
| Income Coefficients (RCNL Models) | | | | | |
| *constant | - | 0.014 | 0.014 | - | - |
| | | (0.005) | (0.005) | | |
| *price | - | 0.001 | 0.001 | - | - |
| | | (0.000) | (0.000) | | |
| *calories | - | 0.004 | 0.004 | - | - |
| | | (0.002) | (0.002) | | |
| Median Product Elasticity | -2.31 | -4.71 | -5.41 | -2.51 | -3.12 |
| Mean ML Brand Price Elasticity | -1.66 | -3.68 | -4.22 | -3.06 | -3.09 |
| Mean Flagship Diversion | 0.41 | 0.48 | 0.47 | 0.83 | 0.90 |
| % Change in Flagship Sales Given 75¢ Domestic Price Rise | -5.20% | -8.24% | -9.65% | -4.30% | -2.20% |
| Observations | 94,656 | 94,656 | 31,777 | 94,656 | 405,004 |

Notes: all specifications include time period and product (brand*size) fixed effects, and use data from Jan 2005 to Dec 2011, excluding June 2008 to May 2009. All estimates use two-step optimal GMM. Instruments are the same as in MW for the relevant specification, apart from the two nest models where we define instruments for the number and distance measures for other products based on products in the same nest, and interact instruments with a flagship brand dummy. Market size is defined as 50% more than the highest sales observed in the geographic market for monthly and quarterly specifications. For the weekly specifications it is estimated as 50% more than the sum of the highest sales from stores observed in the scanner data that week. ML Brand Elasticity reflects the change in ML sales when the prices of all ML products are increased. Mean Flagship Diversion is the average proportion of lost sales that go to other flagship products (i.e., BL, ML and CL products) when the price of a flagship product is increased. The change in flagship sales after a 75 cent price rise is the average across pre-JV observations change in total flagship sales when the prices of all domestic products are increased by 75 cents. Standard errors, clustered on the geographic market, in parentheses.

where $\eta_{i,t} \sim TRN(0, \sigma_c^2, \underline{c_i} - \rho c_{i,t-1} - (1-\rho)\frac{\underline{c_i}+\underline{c_i}+c'}{2}, \underline{c_i} + c' - \rho c_{i,t-1} - (1-\rho)\frac{\underline{c_i}+\underline{c_i}+c'}{2})$ and $\sigma_c$ is the standard deviation of the untruncated innovation distribution. The fit of the model improves only slightly if we allow $\rho$, $\sigma_c$ and $c'$ to vary across firms.

#### 4.3.4 Objective Function, Matched Statistics and Identification.

The cost parameters are estimated using indirect inference (Smith (2008)). For a given value of the cost parameters, we solve the model (see Appendix A.2 for the method) and simulate a time-series of data to calculate six statistics/regression coefficients that we match to ones from the data that we describe below. The estimation problem is

$$\widehat{\theta} = \arg\min_{\theta} g(\theta)' W g(\theta)$$

where $g(\theta)$ is a vector where each element $k$ has the form $g_k = \frac{1}{M}\sum_m \tau_{k,m}^{data} - \widehat{\tau_k(\theta)}$ where $\tau_{k,m}^{data}$ is a statistic estimated using the actual data and $\widehat{\tau_k(\theta)}$ is the equivalent coefficient estimated using simulated data from the model solved using parameters $\theta$. $W$ is a weighting matrix. The reported results use an identity weighting matrix, although the choice of $W$ has little effect on the parameters as we match all of the moments almost exactly. The objective function is minimized using `fminsearch` in MATLAB (version 2018a). Standard errors are calculated treating different markets before the JV as independent observations on the same game. Estimation takes between 12 and 24 hours.[27]

For each geographic market, we calculate six statistics using data from January 2001 to the announcement of the JV in October 2007.[28] Our preferred specification uses weekly data and the five most common pack sizes (6, 12, 18, 24 and 30-packs).[29] Market-week-brand-size average real prices per 12-pack equivalent are calculated excluding temporary store price reductions, and using only market-weeks where we observe more than five stores.[30] The first two statistics that we match are the (unweighted) average prices for BL and ML across pack sizes and weeks.

[27]Computationally light two-step approaches, which are often used to estimate dynamic games, cannot be used because they require that all serially-correlated state variables, which in our setting would include beliefs, are observed by the researcher.

[28]MW use data from 2005 when estimating demand because they need to match sales data to data on market demographics. We use a longer sample as we do not include demographics in our estimation of the supply-side.

[29]Our model does not have different pack sizes, market heterogeneity, varying sets of stores or time trends, so the regressions using simulated data do not control for these factors.

[30]See Appendix D for a discussion of the sample selection.

The third statistic is the interquartile range (IQR) of prices for BL. This is calculated as the IQR of the residuals for each market from a regression where, pooling markets, we regress the week-market-size prices of BL products on dummies for the specific set of stores observed in the market-week (interacted with pack size) and week-size fixed effects in order to control for fixed retail price differences across stores and any national promotions. The remaining statistics are coefficients from market-brand-specific regressions of market-week-brand-size prices on the lagged prices of all three brands. Specifically we use the averages of $\rho^{ML,ML}$ and $\rho^{CL,CL}$, $\rho^{BL,CL}$ and $\rho^{BL,ML}$, and $\rho^{ML,CL}$ and $\rho^{CL,ML}$, where $\rho^{i,j}$ is the coefficient on the lagged price of brand $j$ when the dependent variable is the price of brand $i$. These AR(1) regressions include dummies for the exact set of stores observed, interacted with pack size, and a linear time trend.

Assuming that the equilibrium is unique, the intuition for identification is straightforward.[31] Given the assumed demand parameters and the observed price levels, the mark-ups implied by the model will identify the lower bounds on brand marginal costs. The AR(1) coefficients and the dispersion of prices will identify the range of costs and the parameters of the cost innovation process.[32] We will compare additional statistics that we do not match during estimation to understand the fit of the model.

To provide a sense of the AR(1) coefficients, Table 6 shows the coefficients from similar regressions that pool data from all markets for four alternative samples. Panel (a) reports the results for our preferred specification. The serial correlation parameters for a product's own price are between 0.41 and 0.46, while the cross-product correlations are positive but smaller. If price reductions are included (panel (c)), serial correlations fall, which is consistent with sales lasting one week and being proceeded and followed by higher regular prices. Serial correlation is higher if we use only 12-packs (panel (b)). Panel (d) repeats (a) using monthly prices and

[31] The possibility that our game has multiple equilibria may create two issues for estimation. First, the objective function may be hard to minimize if our solution algorithm jumps between different sections of the equilibrium correspondence. In practice, we can match our moments almost exactly across many alternative parameterizations. Second, another equilibrium supported by different parameters might give similar predictions to the equilibrium that our algorithm finds. This is essentially a potential identification problem. Here we have to rely on the fact that we have never found multiple equilibria in continuous-type games, although we suspect that they may exist for some parameters.

[32] Larger cross-brand $\rho$ coefficients imply stronger signaling effects, so that a smaller range of costs may be required to generate the dispersion of prices in the data. Our experience is that we need to match average prices, some measures of own-brand and cross-brand serial correlation, and some measure of either the dispersion of prices or the variance of innovations in prices to identify the parameters. The exact combination of moments used has little effect on the results.

Table 6: AR(1) Price Regressions Using Flagship Market-Pack Size-Week or -Month Data

(a) Week, Price Reductions Excluded, All Pack Sizes, Fixed Effects for Set of Stores

| | (1) $p_{BL,t}$ | (2) $p_{ML,t}$ | (3) $p_{CL,t}$ |
|---|---|---|---|
| $p_{BL,t-1}$ | 0.451 | 0.056 | 0.043 |
| | (0.033) | (0.017) | (0.010) |
| $p_{ML,t-1}$ | 0.030 | 0.409 | 0.016 |
| | (0.011) | (0.036) | (0.014) |
| $p_{CL,t-1}$ | 0.027 | 0.021 | 0.461 |
| | (0.012) | (0.015) | (0.040) |
| Observations | 36,659 | 36,670 | 36,700 |
| R-squared | 0.979 | 0.972 | 0.978 |
| Mean Price ($) | 10.08 | 9.95 | 9.94 |
| SD residuals ($) | 0.184 | 0.221 | 0.197 |

(b) Week, Price Reductions Excluded, 12 Packs Only, Fixed Effects for Set of Stores

| | (1) $p_{BL,t}$ | (2) $p_{ML,t}$ | (3) $p_{CL,t}$ |
|---|---|---|---|
| $p_{BL,t-1}$ | 0.489 | 0.071 | 0.028 |
| | (0.032) | (0.026) | (0.018) |
| $p_{ML,t-1}$ | 0.062 | 0.505 | 0.028 |
| | (0.013) | (0.038) | (0.012) |
| $p_{CL,t-1}$ | 0.004 | 0.016 | 0.549 |
| | (0.012) | (0.015) | (0.043) |
| Observations | 10,829 | 10,817 | 10,828 |
| R-squared | 0.964 | 0.945 | 0.957 |
| Mean Price ($) | 10.30 | 10.22 | 10.19 |
| SD residuals ($) | 0.144 | 0.183 | 0.163 |

(c) Week, Price Reductions Included, All Pack Sizes, Fixed Effects for Set of Stores

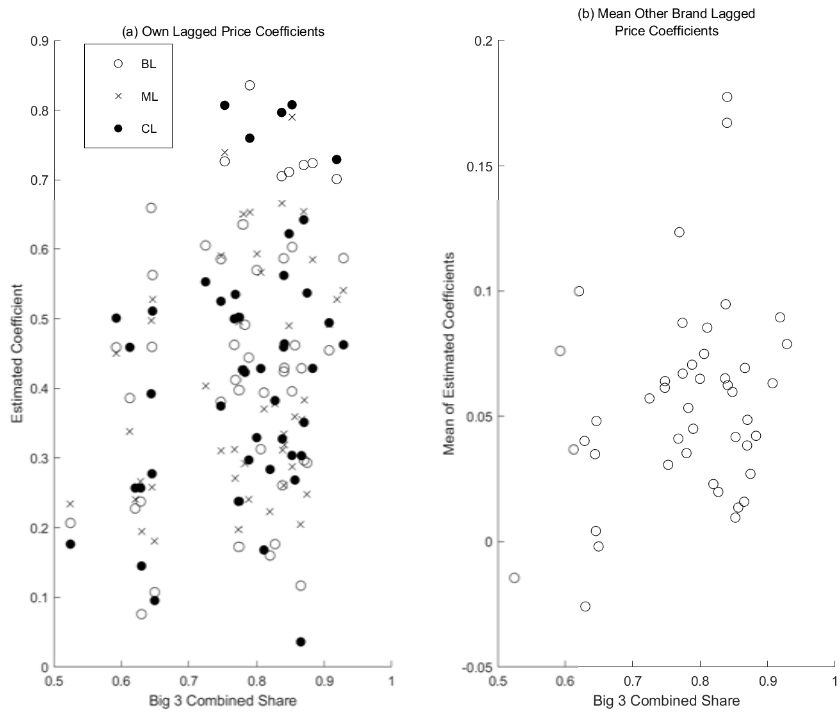
| | (1) $p_{BL,t}$ | (2) $p_{ML,t}$ | (3) $p_{CL,t}$ |
|---|---|---|---|
| $p_{BL,t-1}$ | 0.287 | 0.036 | 0.020 |
| | (0.027) | (0.013) | (0.013) |
| $p_{ML,t-1}$ | 0.045 | 0.322 | 0.010 |
| | (0.009) | (0.027) | (0.012) |
| $p_{CL,t-1}$ | -0.023 | -0.049 | 0.267 |
| | (0.013) | (0.020) | (0.039) |
| Observations | 37,449 | 37,431 | 37,442 |
| R-squared | 0.939 | 0.941 | 0.942 |
| Mean Price | 9.79 | 9.67 | 9.68 |
| SD residuals | 0.337 | 0.342 | 0.336 |

(d) Month, Price Reductions Excluded, All Pack Sizes, Fixed Effects for Markets

| | (1) $p_{BL,t}$ | (2) $p_{ML,t}$ | (3) $p_{CL,t}$ |
|---|---|---|---|
| $p_{BL,t-1}$ | 0.646 | 0.097 | 0.091 |
| | (0.025) | (0.015) | (0.012) |
| $p_{ML,t-1}$ | 0.074 | 0.601 | 0.066 |
| | (0.015) | (0.027) | (0.014) |
| $p_{CL,t-1}$ | 0.100 | 0.097 | 0.682 |
| | (0.010) | (0.016) | (0.025) |
| Observations | 13,972 | 13,973 | 13,975 |
| R-squared | 0.974 | 0.971 | 0.974 |
| Mean Price | 10.08 | 9.95 | 9.94 |
| SD residuals | 0.210 | 0.229 | 0.216 |

Notes: regressions also include time period*pack size interactions and use pack sizes containing volumes equivalent to 6, 12, 18, 24 and 30 12 oz. containers. Market or store fixed effects described in the label to each panel. Standard errors, clustered on the market, are in parentheses. The SD residuals statistic is the standard deviation of the residuals from the regression.

Figure 5: Estimated Pre-JV Price Dynamics and the Combined Market Shares of AB, Miller and Coors.




Notes: The estimated univariate regression coefficients, with standard errors in parentheses, for panel (a) are BL: 0.011 (0.226) + 0.558$C_3$ (0.288), $R^2$ = 0.080; ML : 0.044 (0.192) + 0.465$C_3$ (0.245), $R^2$=0.077; CL : -0.025 (0.215) + 0.568$C_3$ (0.278), $R^2$=0.091; and for panel (b): -0.039 (0.046) + 0.120$C_3$ (0.058), $R^2$=0.088.

market, rather than group-of-store, fixed effects (equivalent regressions will be used in our monthly data specification). In this case, the serial correlation parameters increase, but further investigation reveals that this happens primarily due to the change in the fixed effects.[33]

While our calibration does not seek to match cross-market heterogeneity, the serial correlation coefficients show some interesting patterns across markets. Using data simulated from our model, we typically estimate higher serial correlation parameters when we change the parameters to induce larger signaling effects on prices, by, for example, reducing diversion to the outside good. Given any type of logit or nested logit preferences, diversion to other brands will tend to be lower when the market share accounted for by the signaling brands is higher. Figure 5(a) shows scatter plots of the estimated market-level serial correlation parameters for BL, ML

[33]We have estimated monthly regressions including set of store fixed effects and dropping market-months where the set of stores changes within months. This causes the number of observations to drop dramatically: for example, the number of observations in the BL regression falls to 2,806, and the estimated coefficient on $p_{t-1}^{BL}$ falls to 0.318. For some individual markets, there is not enough data to estimate serial correlation coefficients.

Table 7: Parameter Estimates for Six Specifications

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Data Frequency | Week | Week | Week | Week | Week | Month |
| Sizes | All | 12 only | All | All | All | All |
| Price Reductions | Excl. | Excl. | Incl. | Excl. | Excl. | Excl. |
| Mean Brand Price Elasticity | -3 | -3 | -3 | -2.5 | -3.5 | -3 |
| Mean Flagship Diversion | 85% | 85% | 85% | 90% | 80% | 85% |
| Lower Bound Cost for BL | \$5.259 | \$5.278 | \$4.845 | \$4.248 | \$5.973 | \$4.616 |
| ($\underline{c_{BL}}$) | (0.201) | (0.048) | (0.046) | (0.043) | (0.026) | (0.127) |
| L.B. Cost for ML/CL | \$6.425 | \$6.528 | \$5.984 | \$5.786 | \$6.874 | \$5.711 |
| ($\underline{c_{ML/CL}}$) | (0.020) | (0.014) | (0.022) | (0.024) | (0.017) | (0.020) |
| Width Cost Interval | \$0.625 | \$0.752 | \$1.246 | \$0.556 | \$0.672 | \$1.793 |
| ($\overline{c_i}-\underline{c_i}$) | (0.029) | (0.021) | (0.018) | (0.102) | (0.026) | (0.037) |
| Cost AR(1) Parameter | 1.156 | 0.939 | 0.850 | 1.222 | 0.959 | 0.742 |
| ($\rho$) | (0.020) | (0.011) | (0.026) | (0.013) | (0.012) | (0.025) |
| SD Cost Innovations | \$0.282 | \$0.278 | \$0.566 | \$0.260 | \$0.270 | \$0.400 |
| ($\sigma_c$) | (0.024) | (0.001) | (0.050) | (0.104) | (0.026) | (0.052) |

Notes: BL = Bud Light, ML = Miller Lite and CL=Coors Light. Standard errors in parentheses. The data specifications using weekly data include group-of-store fixed effects when calculating the data statistics. For the monthly specification, the regression using the data only include market fixed effects

and CL against the share of all beer sales accounted for AB, Miller and Coors in 2007 (i.e., the $C_3$). Figure 5(b) shows a similar plot for the average of the six cross-brand coefficients. In both cases there is a positive, and, using a regression analysis, a statistically significant, relationship, consistent with our simulations.[34]

#### 4.3.5 Parameter Estimates and Model Fit.

Table 7 reports estimates from six specifications, using different data or alternative demand parameters. Estimated marginal costs increase when demand is more elastic, and the range of costs and the standard deviation of the innovations increase when we try to match data that contains temporary price reductions. The estimated marginal cost ranges are much larger than in our examples, but the estimated $\sigma_c$s imply that the probability that a marginal cost can go from high to low across periods is quite high.[35] As we will note below, the volatility of observed prices means that the marginal costs implied by CI Nash or conduct parameter models are also

[34]We also find positive, statistically significant relationships when we look at individual cross-brand coefficients.

[35]For example, for the specification in column (1) the probability that a firm with marginal cost $\overline{c_i}$ will have a marginal cost in the lower half of the range in the next period is 0.24, similar to 0.32 in our baseline example.

Table 8: Model Fit for Three Specifications Using Weekly Data, Average Brand Price Elasticity of -3 and Flagship Diversion of 85%

| | (1) | | (2) | | (3) | |
|---|---|---|---|---|---|---|
| Frequency | Week | | Week | | Week | |
| Sizes | All | | 12 | | All | |
| Price Reductions | Excl. | | Excl. | | Incl. | |
| | Data | Model | Data | Model | Data | Model |
| | | | Matched Moments | | | |
| Mean $p_{BL}$ | \$10.09 | \$10.09 | \$10.30 | \$10.30 | \$9.81 | \$9.81 |
| Mean $p_{ML}$ | \$9.96 | \$9.96 | \$10.22 | \$10.22 | \$9.68 | \$9.68 |
| Mean $\rho^{ML,ML}$, $\rho^{CL,CL}$ | 0.402,0.413 | 0.408 | 0.468,0.450 | 0.444 | 0.330,0.290 | 0.313 |
| Mean $\rho^{BL,ML}$, $\rho^{BL,CL}$ | 0.082,0.066 | 0.074 | 0.102,0.056 | 0.076 | 0.070,0.060 | 0.059 |
| Mean $\rho^{ML,CL}$, $\rho^{CL,ML}$ | 0.051,0.036 | 0.033 | 0.065,0.026 | 0.035 | 0.049,-0.004 | 0.028 |
| IQR $p_{BL}$ | \$0.189 | \$0.189 | \$0.185 | \$0.212 | \$0.314 | \$0.313 |
| | | | Unmatched Moments | | | |
| Mean $p_{CL}$ | \$9.95 | \$9.96 | \$10.20 | \$10.23 | \$9.68 | \$9.68 |
| $\rho^{BL,BL}$ | 0.444 | 0.385 | 0.442 | 0.418 | 0.311 | 0.296 |
| Mean $\rho^{ML,BL}$, $\rho^{CL,BL}$ | 0.059,0.0.42 | 0.038 | 0.065,0.040 | 0.038 | 0.076,0.004 | 0.029 |
| SD of BL Res. | \$0.177 | \$0.109 | \$0.136 | \$0.122 | \$0.317 | \$0.188 |
| SD of ML/CL Res. | \$0.204,\$0.189 | \$0.159 | \$0.161,\$0.149 | \$0.179 | \$0.322,\$0.311 | \$0.271 |
| IQR $p_{ML}, p_{CL}$ | \$0.222,\$0.210 | \$0.281 | \$0.228,\$0.206 | \$0.316 | \$0.335,\$0.316 | \$0.462 |
| Skewness of BL Res. | -0.361 | -0.337 | -0.307 | -0.314 | -0.806 | -0.098 |
| ML/CL Res. | -0.100,-0.329 | -0.331 | -0.296,-0.201 | -0.297 | -0.717,-0.696 | -0.080 |

Notes: BL = Bud Light, ML = Miller Lite and CL=Coors Light. SD = standard deviation. Res. = residuals from the AR(1) regressions. For the data we report separate values for the statistics for ML and CL, but, because the model assumes that ML and CL are symmetric, and so predicts identical statistics (ignoring simulation error), we match the average of these values during estimation and report a single prediction.

quite volatile.

The upper panel of Table 8 reports the fit of the moments that we match during estimation for the column (1), (2) and (3) specifications. The lower part of the table reports moments that are not matched, including the skewness of the innovations from the AR(1) regression. The model systematically underpredicts the standard deviation of price residuals for BL. The other moments are matched quite accurately, except that we cannot match the skewness of the residuals when price promotions are included in the data, consistent with our model having no mechanism to match these types of changes.

#### 4.3.6 Predicted Effects of the JV.

Table 9 reports predicted prices when we resolve the six models assuming that ML and CL have the same marginal cost and are sold by a single firm at the same price. We assume that

Table 9: Predicted Average Prices Before and After the MC JV

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Frequency | Week | Week | Week | Week | Week | Month |
| Sizes | All | 12 only | All | All | All | All |
| Price Reductions | Excl. | Excl. | Incl. | Excl. | Excl. | Excl. |
| Brand Elasticity | -3 | -3 | -3 | -2.5 | -3.5 | -3 |
| Flagship Diversion | 85% | 85% | 85% | 90% | 80% | 85% |
| Pre-JV Mean Prices | | | | | | |
| BL | \$10.09 | \$10.30 | \$9.81 | \$10.09 | \$10.09 | \$10.09 |
| ML/CL | \$9.96 | \$10.22 | \$9.68 | \$9.96 | \$9.96 | \$9.95 |
| Assumed ML/CL Synergy | -\$1.18 | -\$1.20 | -\$1.14 | -\$1.50 | -\$0.94 | -\$1.17 |
| Post-JV Mean Prices | | | | | | |
| BL | \$10.62 (+5.3%) | \$10.90 (+5.7%) | \$10.17 (+3.7%) | \$10.98 (+8.7%) | \$10.42 (+3.3%) | fails |
| ML/CL | \$10.48 (+5.2%) | \$10.79 (+5.8%) | \$10.02 (+3.5%) | \$10.82 (+8.5%) | \$10.27 (+3.1%) | fails |

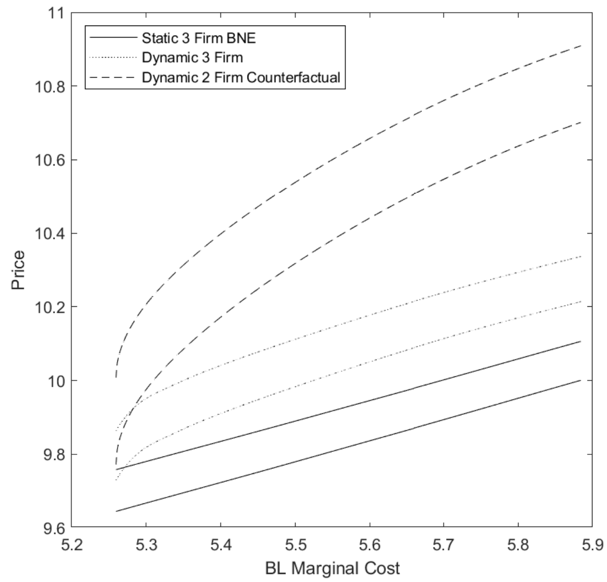
Notes: BL = Bud Light, ML = Miller Lite and CL=Coors Light. For the data we report separate values for the statistics for ML and CL, but, because the model assumes that ML and CL are symmetric, and so predicts identical statistics (ignoring simulation error), we report a single prediction.

MC benefits from a synergy that would have prevented average prices from rising if firms set static CI Nash prices, as this seems consistent with the DOJ's expectation, but the width of the cost interval and the remaining parameters remain the same. The predicted price changes in columns (1)-(5) are all within the estimated 40¢-\$1 or 3-6% ranges.[36] We cannot find an equilibrium for the monthly data specification. In this case, the estimated parameters imply marginal costs are more persistent (the probability that a firm with the cost $\overline{c_i}$ will have a cost less than $\frac{\underline{c_i}+\overline{c_i}+c'}{2}$ is only 0.067) because, in this case, we are matching coefficients from a regression that does not control for cross-store heterogeneity in retail prices, and signaling incentives raise prices so high that the conditions for separation fail.

Figure 6 compares, using the column (1) parameters, BL's equilibrium pricing strategies for the static Bayesian Nash 3-firm model, the estimated signaling 3-firm model and the counterfactual post-JV model. Signaling increases the level and the range of BL prices, which span from the lowest point on the two BL pricing functions to the highest point, especially in the counterfactual.

[36]One might be concerned that our assumed discount factor of $\beta = 0.99$ is too low for weekly data. We have recomputed the column (1) estimates assuming $\beta = 0.998$, implying an annual discount factor of around 0.9. While a higher discount factor increases signaling incentives, the estimated parameters change to rationalize pre-JV dynamics in such a way that the predicted post-JV prices are within 1 cent of those reported in Table 9.

Figure 6: Bud Light Equilibrium Pricing Strategies (for estimates in column (1) of Table 7).




Notes: the strategies shown assume that $c_{t-1}^{BL} = \underline{c^{BL}}$ and $c_{t-1}^{ML} = c_{t-1}^{CL} = \underline{c^{ML/CL}}$ (lower line) and $c_{t-1}^{BL} = \overline{c^{BL}}$ and $c_{t-1}^{ML} = c_{t-1}^{CL} = \overline{c^{ML/CL}}$ (upper line). Therefore, for each type of equilibrium, the maximum range of BL's prices spans from the lowest point on the bottom line to the highest point on the upper line.

Table 10 compares the cross-market averages of the price dynamic statistics before and after the JV in the data, and the values predicted by the column (1) model. The model correctly predicts the directional change in each statistic except the skewness measures, even if it does not predict which statistics change the most. We view our ability to match qualitative changes in dynamics, as well as the increase in average price levels, even though our model is calibrated using only pre-JV data, as an encouraging result.

## 4.4 Testing Alternative Explanations for the Post-JV Price Increases.

Some people have suggested that, even if our model can explain why prices rose after the JV, MW and MSW's CI theories of tacit collusion provide pre-existing and satisfactory explanations. While folk theorems imply that a CI tacit collusion model that fits the data almost perfectly is likely to exist, we can test how well MW and MSW's assumptions fit the data. MSW's supermarkup model of collusion is clearly rejected and, in some specifications, MW's baseline interpretation that there was CI Nash pricing before the JV is also rejected. The estimates also imply that marginal costs are serially correlated and quite volatile, a feature that plays an

Table 10: Observed and Predicted Changes in Price Dynamics

| | Data | | | Fitted Model | | |
|---|---|---|---|---|---|---|
| | Pre-JV | Post-JV | Change | Pre-JV | Post-JV | Change |
| IQR of Prices | | | | | | |
| BL | \$0.189 | \$0.241 | +0.052 | \$0.189 | \$0.353 | +0.164 |
| ML | \$0.222 | \$0.256 | +0.034 | \$0.281 | \$0.350 | +0.069 |
| CL | \$0.210 | \$0.244 | +0.034 | \$0.281 | \$0.350 | +0.069 |
| AR(1) Regression Coefficients | | | | | | |
| $\rho^{BL,BL}$ | 0.444 | 0.524 | +0.080 | 0.385 | 0.415 | +0.030 |
| $\rho^{ML,ML}$ | 0.402 | 0.483 | +0.081 | 0.408 | 0.409 | +0.001 |
| $\rho^{CL,CL}$ | 0.413 | 0.453 | +0.040 | 0.408 | 0.409 | +0.001 |
| $\rho^{BL,ML}$ | 0.082 | 0.092 | +0.010 | 0.074 | 0.122 | +0.048 |
| $\rho^{BL,CL}$ | 0.066 | 0.095 | +0.029 | 0.074 | 0.122 | +0.048 |
| $\rho^{ML,BL}$ | 0.059 | 0.087 | +0.028 | 0.038 | 0.141 | +0.103 |
| $\rho^{CL,BL}$ | 0.042 | 0.080 | +0.038 | 0.038 | 0.141 | +0.103 |
| Std. Dev. of AR(1) regression residuals | | | | | | |
| BL regression | \$0.177 | \$0.188 | +0.011 | \$0.109 | \$0.203 | +0.094 |
| ML regression | \$0.204 | \$0.204 | +0.000 | \$0.159 | \$0.204 | +0.045 |
| CL regression | \$0.189 | \$0.193 | +0.004 | \$0.159 | \$0.204 | +0.045 |
| Skewness of AR(1) regression residuals | | | | | | |
| BL regression | -0.361 | -0.181 | +0.180 | -0.337 | -0.504 | -0.167 |
| ML regression | -0.100 | 0.001 | +0.101 | -0.331 | -0.470 | -0.139 |
| CL regression | -0.329 | -0.104 | +0.225 | -0.331 | -0.470 | -0.139 |

Notes: BL = Bud Light, ML = Miller Lite and CL=Coors Light. The calculation of the statistics is explained in Section 4.3.4. Pre-JV averages are calculated for 45 markets, and post-JV averages are calculated for 44 markets, as one market does not have at least 5 stores observed in consecutive weeks after the JV.

important role in our model.

Our tests extend MW's conduct parameter framework. The framework assumes that pricing is characterized by stacked static, CI first-order conditions

$$\left(\Omega_{mt} \circ \left[\frac{\partial q_{mt}(p_{mt},\theta^D)}{\partial p_{mt}}\right]\right)(p_{mt}-c_{mt})+q_{mt}(p_{mt},\theta^D)=0,$$

where $p_{mt}$, $q_{mt}$ and $c_{mt}$ are vectors of prices, quantities and (constant) marginal costs and $\frac{\partial q_{mt}(p_{mt},\theta^D)}{\partial p_{mt}}$ is a matrix of demand derivatives.

$\Omega_{mt}$ is the "conduct" matrix, with (row $i$, column $j$) element $\Omega_{i,j}$. $\Omega_{i,j}=1$ if products $i$ and $j$ are owned by the same firm. Under static Nash pricing, all other elements of $\Omega_{mt}$ are zero. MW's baseline specification assumes static Nash pricing before the JV, but allows $\Omega_{i,j}=\kappa$ after the JV if $i$ and $j$ are owned by different domestic brewers. $\kappa=1$ is consistent with joint profit-maximization, while $0<\kappa<1$ could be interpreted as reflecting partial internalization of pricing externalities.

Given demand estimates, MW estimate the post-JV $\kappa$ using equations

$$p_{mt} = W_{mt}\gamma - \left(\Omega_{mt}(\kappa) \circ \left[\frac{\partial s_{mt}(p_{mt}, \theta^D)}{\partial p_{mt}}\right]\right)^{-1} s_{mt}(p_{mt}) + \nu_{mt}. \tag{5}$$

where $c_{imt} = W_{imt}\gamma + \nu_{imt}$ and $W$ includes time, product (brand-size) and geographic market fixed effects; a "distance measure" that multiplies distance to the brewery or port with real diesel prices; and, a dummy for MC products after the JV to allow for an additional efficiency. The JV is assumed not to affect AB's marginal costs. The instruments are the variables in $W$ and a dummy for domestic products after the JV. The post-JV $\kappa$ is identified by how much more AB's prices increase than the increase that can be rationalized as a static best response.

MW's single exclusion restriction implies that they cannot estimate separate pre- and post-JV $\kappa$s or test whether a change in conduct is the source of the price increase.[37] We provide this type of test by adding additional instruments and controls.[38] Note, however, that we will only use the model to test MW and MSW's assumptions and we will not interpret positive $\kappa$s as evidence of collusion. As shown by Corts (1999), some forms of tacit collusion may be consistent with estimates of $\kappa$ that are less than or equal to zero, and, as we discuss below, our signaling model tends to imply positive estimates of $\kappa$ even though there is no collusion.

Our specifications include separate pre- and post-JV product and market fixed effects in $W$. To understand our choice of instruments, consider the first-order condition for product $i$ owned by AB

$$p_{imt} = W_{imt}\gamma + \frac{q_{imt}}{\frac{\partial q_{imt}}{\partial p_{imt}}} + \sum_{\substack{j \in AB \\ j \neq i}} \frac{\frac{\partial q_{jmt}}{\partial p_{imt}}}{\frac{\partial q_{imt}}{\partial p_{imt}}}(p_{jmt} - c_{jmt}) + \kappa \sum_{k \in M,C} \frac{\frac{\partial q_{kmt}}{\partial p_{imt}}}{\frac{\partial q_{imt}}{\partial p_{imt}}}(p_{kmt} - c_{kmt}) + \nu_{imt}.$$

Valid instruments will be correlated with $\sum_{k \in M,C} \frac{\frac{\partial q_{kmt}}{\partial p_{imt}}}{\frac{\partial q_{imt}}{\partial p_{imt}}}(p_{kmt} - c_{kmt})$ (i.e., the incremental effect of a change in $i$'s price on a rival's profits), and uncorrelated with the cost unobservable $\nu_{imt}$.

The first six columns in Table 11 report conduct coefficients for the columns (1)-(3) demand

[37] MW re-estimate the post-JV $\kappa$ assuming, but not estimating, different pre-JV $\kappa \leq 0.5$. These estimates imply that $\kappa$ rose after the JV, although by smaller amounts as the assumed pre-JV $\kappa$ rises, as a pre-JV $\kappa$ also implies that AB would increase its prices when MC benefits from an efficiency.

[38] We continue to assume that imported brands use Nash pricing and that $\Omega_{i,j} = 1$ when $i$ and $j$ have the same owner.

Table 11: Testing Alternative Models Using a Generalized Conduct Parameter Framework

| | Tests of MW Conduct Model | | | | | | Tests of MW & MSW Models | | | Test of MSW |
|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Demand Model | NL | RCNL | RCNL | NL | RCNL | RCNL | NL | RCNL | RCNL | RCNL |
| (all one nest) | Monthly | Monthly | Quarterly | Monthly | Monthly | Quarterly | Monthly | Monthly | Quarterly | Quarterly |
| Domestic Firms | | | | | | | | | | FY06, FY07 |
| Pre-JV Conduct | 0.274 | 0.322 | 0.198 | 0.340 | 0.263 | 0.238 | 0.958 | 0.909 | 0.913 | 0.977, 0.924 |
| | (0.098) | (0.193) | (0.221) | (0.057) | (0.166) | (0.147) | (0.005) | (0.016) | (0.024) | (0.007),(0.013) |
| | | | | | | | | | | FY10, FY11 |
| Post-JV Conduct | 0.723 | 0.651 | 0.573 | 0.688 | 0.767 | 0.717 | 0.951 | 0.914 | 0.921 | 0.976, 0.933 |
| | (0.124) | (0.146) | (0.127) | (0.144) | (0.056) | (0.062) | (0.009) | (0.014) | (0.013) | (0.011),(0.015) |
| p-value diff. | 0.004 | 0.013 | 0.017 | 0.000 | 0.000 | 0.001 | 0.483 | 0.638 | 0.558 | - |
| Supermarkup | - | | | | | | Domestic*Market*Fiscal Year Fixed Effects | | | |
| Controls | | | | | | | linear | | | non-linear |
| Excluded IVs | Domestic Rival Distance | | | Dom. Rival Distance, Dom. Rival $\xi$s and Interactions | | | Dom. Rival Distance, Dom. Rival $\xi$s and Interactions | | | |
| ML 12 Packs | | | | | | | | | | |
| Pre-JV: Mean $\widehat{c_{imt}}$ | \$2.37 | \$5.93 | \$6.75 | \$1.99 | \$6.13 | \$6.65 | -\$4.61 | \$1.73 | \$2.79 | -\$1.41 |
| Residual $\rho$ | 0.414 | 0.427 | 0.451 | 0.410 | 0.430 | 0.449 | 0.239 | 0.252 | 0.114 | 0.030 |
| SD AR(1) res. | \$0.31 | \$0.27 | \$0.20 | \$0.31 | \$0.27 | \$0.20 | \$0.63 | \$0.44 | \$0.30 | \$0.35 |
| Post-JV: Mean $\widehat{c_{imt}}$ | -\$1.34 | \$4.20 | \$5.40 | -\$1.00 | \$3.34 | \$4.62 | -\$4.39 | \$1.81 | \$2.93 | -\$5.74 |
| Residual $\rho$ | 0.431 | 0.488 | 0.403 | 0.433 | 0.485 | 0.417 | 0.224 | 0.436 | 0.050 | 0.006 |
| SD AR(1) res. | \$0.43 | \$0.33 | \$0.26 | \$0.41 | \$0.37 | \$0.28 | \$0.56 | \$0.42 | \$0.29 | \$0.43 |
| Observations | 94,656 | 94,656 | 31,777 | 94,656 | 94,656 | 31,777 | 94,656 | 94,656 | 31,777 | 31,777 total |

Notes: Specifications estimated using 2-step GMM. The specifications in columns (1)-(9) contain time period fixed effects, and separate product and market fixed effects for before and after the JV, as well as the distance measure interacted with combinations of dummies for domestic products and periods after the JV. The specification in column (10) is estimated separately for each fiscal year (e.g., the FY06 year runs October 2005-September 2006), and the specification includes product, city and quarter fixed effects, the distance measure (interacted with a dummy for domestic products) as well as non-linear market fixed effects for the domestic products. Conduct parameters are reported for four fiscal years. The "residual $\rho$" statistics are the coefficients on lagged marginal costs ($c_{imt-1}$) from a regression of ML 12-pack marginal costs on their lagged values, market and time fixed effects. These regressions are estimated separately before and after the JV. The "SD AR(1) res." statistics are the standard deviation of the residuals from these regressions. Standard errors in parentheses clustered at the market level.

specifications in Table 5.[39] Columns (1)-(3) use the distance measures of rivals as instruments, as they affect rivals' margins, and, as MW already assume that a product's own distance measure is uncorrelated with $\nu_{imt}$, the additional assumptions required are minimal.[40] Columns (4)-(9) use additional instruments in the form of the average value of the demand unobservables ($\xi$s) for rival brewers over either the pre- or post-JV period, and the interactions of these instruments with the distance instruments.[41] These additional instruments are valid if $\nu_{imt}$ is uncorrelated with the demand unobservables of rivals' products. This is a stronger assumption, although economists sometimes assume that a product's own demand and marginal costs unobservables are uncorrelated in order to estimate demand (MacKay and Miller (2019)). Columns (7)-(9) include linear domestic-market-fiscal year fixed effects in $W$. These controls allow for possible correlations between local preferences and costs for domestic products as a group, and cause conduct to be identified only from within-market-year cross-brewer/-product variation. We will also use these specifications to test the MSW model.

We reject Nash pricing after the JV in all nine specifications. This is, of course, consistent with MW's interpretation that there was collusion after the JV. All of the estimated pre-JV $\kappa$s are positive, and some are significant. The estimates in columns (1)-(6) are consistent with an increase in $\kappa$ after the JV, but the estimates with market-year controls suggest that conduct did not change, even though the $\kappa$ estimates are very precise.

The plausibility of these CI pricing models can also be assessed by looking at what they imply for marginal costs and synergies. Table 11 reports average implied marginal costs for ML 12-packs. Less elastic demand and higher $\kappa$ imply lower marginal costs, and the (1), (4) and (7)-(9) costs are implausibly/impossibly low. The remaining columns imply synergies for ML, which was being shipped the same distances before and after the JV in most markets,

[39] We have also estimated specifications using the two nest nested logit models, and specifications that estimate $\kappa$s based only on the pricing of the flagship brands. These estimates lead us to reject Nash pricing behavior before the JV, and the pre- and post-JV parameters are closer than those in columns (1)-(6).

[40] There are eight excluded distance instruments. For AB products in market $m$ and time $t$ before the JV, the $(m,t)$ distance measure for Miller and the $(m,t)$ distance measure for Coors are instruments. For pre-JV Miller products, the distance measures for AB and Coors are instruments. For pre-JV Coors products, the distance measures for Miller and AB are instruments. For AB (MC) products in market $m$ and time $t$ after the JV, the $(m,t)$ distance measure for MC (AB) is the instrument.

[41] Specifically, we calculate the average value of the demand residuals for the products sold by brewer $b$ in market $m$ either before or after the JV, and then construct eight instruments in the same way that we construct the instruments for distance. We average across periods because the demand unobservables are more variable than the distance measures.

that are higher than the 17.5% synergy for ML and CL that we assumed for the column (1) specification of our model. Controlling for market and time effects, the implied $\nu_{imt}$s are also serially correlated and quite volatile.[42] While cost volatility is certainly not inconsistent with CI, we view volatility as suggesting that a collusive interpretation of the data requires a very strong CI assumption: if CI is not satisfied, then, given that prices are volatile, collusion would be hampered by the difficulty of distinguishing cheating from a conforming price set by a low marginal cost firm.

The conduct model is not a fully-specified model of collusion because it does not specify why firms choose not to cheat. Some collusion models cannot be tested using the conduct framework, but the MSW supermarkup model can. MSW assume that, every fiscal year, both before and after the JV, a price leader suggests a "supermarkup" on top of Bertrand Nash prices that domestic brewers should charge. If a domestic firm fails to charge the supermarkup, a punishment phase ensues, but in a CI subgame perfect equilibrium, the suggested supermarkup will satisfy the incentive-compatibility constraints (ICCs). Prices may increase after a merger if the ICCs are relaxed. We can test this model by using an appropriately defined domestic product market-fiscal year fixed effect to control for the supermarkup. If the "supermarkup on Nash" theory is correct, estimates of conduct $\kappa$ parameters should be equal to zero once the fixed effects are included.

The columns (7)-(9) include linear domestic-market-fiscal year fixed effects. These specifications are not quite consistent with the MSW's exact theory[43], but they are simple to estimate. As already discussed, we can reject $\kappa = 0$ before or after the JV at any significance level. Column (10) tests MSW's exact model by allowing for non-linear domestic-market-fiscal year fixed effects (see Appendix E for details) using the quarterly RCNL model (most favorable to pre-JV Nash pricing in columns (3) and (6)). We estimate the model separately for each fiscal year to reduce the number of coefficients estimated simultaneously. Consistent with column (9), the reported conduct parameters are precisely estimated and are between 0.9 and 1, and, because estimated supermarkups are also positive, most of the implied marginal costs are neg-

[42]The rich fixed effects in columns (7)-(9) cause the $\nu_{imt}$s to jump across fiscal years, so the estimated serial correlation falls.

[43]Linear fixed effects would be consistent with a model where the leader suggested domestic firms set Nash prices "as if" all of their marginal costs had been raised by a common fixed amount, rather than suggesting a common dollar per 12-pack equivalent price addition to Nash prices.

ative. Therefore, we can clearly reject the MSW formulation of CI collusion, although, as we have emphasized, this does not imply that all models of collusion would be rejected.

While our model implies that the conduct parameter framework is misspecified, because it does not control for beliefs or signaling incentives, we have estimated conduct parameter models using data simulated from two and three-firm versions of our model with cross-firm heterogeneity. The estimated conduct parameters are typically between 0.3 and 1, and the implied marginal costs are usually significantly below their true levels.[44] The estimated conduct parameters can rise, fall or stay roughly unchanged after a merger. The results of our conduct analysis are therefore not inconsistent with what one would expect given our model.

# 5 Conclusion

We have developed a model where oligopolists simultaneously use prices to signal private information that is relevant for their future pricing decisions. Although the possibility that this type of behavior would raise equilibrium prices was identified in the theoretical literature over thirty years ago, we provide the first attempt to quantify the magnitude of these effects, both in examples and in an empirical application. We find that effects can be large, and that they can explain changes in price levels and price dynamics after a large horizontal transaction in the U.S. beer industry. While CI theories of tacit collusion can also explain an increase in price levels, our model provides a natural explanation of the period-to-period price changes observed in this data, and in data from other industries where tacit collusion has been suggested (Ordover (2007)). It is also consistent with how firms treat margin information as highly confidential.

We have often been asked how our model and our empirical analysis relate to theories of "coordinated effects" in merger analysis. There is no standard definition of coordinated effects: the presentation in Ordover (2007) is focused on variants of tacit collusion models, but Baker and Farrell (forthcoming) and Farrell and Baker (2021) use a much broader definition which includes both "purposive" theories of collusion and "non-purposive" theories, a group which includes the non-collusive Markov Perfect theories of Maskin and Tirole (1988). Our model

[44] If, conditional on controls for costs, firms tend to set higher prices when other firms have higher margins or there is more diversion to those rivals' products, then estimated $\kappa$s will be positive. As discussed previously, these features also tend to lead to stronger signaling effects in our model, so we tend to estimate positive $\kappa$s using simulated data from our model.

lies within this group. Non-purposive theories are valuable partly because they can explain why it may not be appropriate to rely exclusively on static CI unilateral effects models in industries that do not have the characteristics that economists typically believe favor tacit collusion (Stigler (1964)) or where, before a transaction, prices do not display the rigidity that collusive theories often predict (Athey, Bagwell, and Sanchirico (2004)). They can also explain why coordinated effects do not raise prices to joint-profit maximizing levels, an outcome that a tacit collusion model will predict if prices are set frequently and firms are patient. However, we also believe that combining tacit collusion and asymmetric information is likely to be a profitable direction for future research, building on the work of Kreps, Milgrom, Roberts, and Wilson (1982) and Athey and Bagwell (2008) who have examined the links in very stylized theoretical models.[45] In fact, one of our examples in Appendix B.2 illustrates how signaling could exacerbate the impact of small coordination incentives.

One could also ask what our model adds to existing non-purposive theories. Maskin and Tirole (1988) provide examples of price-setting games which lead to both price rigidity and price volatility without any underlying volatility in costs or asymmetries of information. We view our introduction of serially correlated asymmetric information as not only realistic, but, also potentially helpful in solving more complicated models, without assuming price changes are asynchronous or subject to potentially large menu costs (Maskin and Tirole (1988), Nakamura and Zerom (2010)), because it means that firms choose prices against a perceived continuous distribution of rivals' prices. This feature of asymmetric information models has long been appreciated in both the static and dynamic discrete choice games literatures (e.g., Seim (2006)), but there are also benefits when choices are continuous.

[45] Athey and Bagwell (2008) consider an example that is explicitly connected to the Mailath (1989) model.

# References


Ashenfelter, O., D. Hosken, and M. Weinberg (2014): "Did Robert Bork Understate the Competitive Impact of Mergers? Evidence from Consummated Mergers," *Journal of Law and Economics*, 57(S3), S67–S100.

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

Asker, J., C. Fershtman, J. Jeon, and A. Pakes (2020): "A Computational Framework for Analyzing Dynamic Procurement Auctions: The Market Impact of Information Sharing," *RAND Journal of Economics*, 51(3), 805–839.

Athey, S., and K. Bagwell (2008): "Collusion with Persistent Cost Shocks," *Econometrica*, 76(3), 493–540.

Athey, S., K. Bagwell, and C. Sanchirico (2004): "Collusion and Price Rigidity," *Review of Economic Studies*, 71(2), 317–349.

Azar, J., M. C. Schmalz, and I. Tecu (2018): "Anticompetitive Effects of Common Ownership," *Journal of Finance*, 73(4), 1513–1565.

Baker, J. B., and J. Farrell (forthcoming): "Oligopoly Coordination, Economic Analysis, and the Prophylactic Role of Horizontal Merger Enforcement," *University of Pennsylvania Law Review*.

Berry, S. T., and P. A. Haile (2014): "Identification in differentiated products markets using market level data," *Econometrica*, 82(5), 1749–1797.

Bonatti, A., G. Cisternas, and J. Toikka (2017): "Dynamic Oligopoly with Incomplete Information," *Review of Economic Studies*, 84(2), 503–546.

Bresnahan, T. F. (1982): "The Oligopoly Solution Concept is Identified," *Economics Letters*, 10(1-2), 87–92.

——— (1989): "Empirical Studies of Industries with Market Power," *Handbook of Industrial Organization*, 2, 1011–1057.

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

Caminal, R. (1990): "A Dynamic Duopoly Model with Asymmetric Information," *Journal of Industrial Economics*, 38(3), 315–333.

Cho, I., and D. Kreps (1987): "Signaling Games and Stable Equilibria," *Quarterly Journal of Economics*, 102(2), 179–221.

Corts, K. S. (1999): "Conduct Parameters and the Measurement of Market Power," *Journal of Econometrics*, 88(2), 227–250.

De Angelis, D., and Y. Grinstein (2014): "Performance Terms in CEO Compensation Contracts," *Review of Finance*, 19(2), 619–651.

Doraszelski, U., and J. Jaumandreu (2013): "R&D and Productivity: Estimating Endogenous Productivity," *Review of Economic Studies*, 80(4), 1338–1383.

Ericson, R., and A. Pakes (1995): "Markov-Perfect Industry Dynamics: A Framework for Empirical Work," *Review of Economic Studies*, 62(1), 53–82.

Farrell, J., and J. B. Baker (2021): "Natural Oligopoly Responses, Repeated Games, and Coordinated Effects in Merger Analysis: A Perspective and Research Agenda," *Review of Industrial Organization*, 58(2), 1–39.

Fershtman, C., and A. Pakes (2012): "Dynamic Games with Asymmetric Information: A Framework for Empirical Work," *Quarterly Journal of Economics*, 127(4), 1611–1661.

Garmon, C. (2017): "The Accuracy of Hospital Merger Screening Methods," *RAND Journal of Economics*, 48(4), 1068–1102.

Hubbard, T. P., and H. J. Paarsch (2013): "On the Numerical Solution of Equilibria in Auction Models with Asymmetries within the Private-Values Paradigm," in *Handbook of Computational Economics*, ed. by K. Schmedders, and K. L. Judd, vol. 3, pp. 37–115. Elsevier: North-Holland.

Judd, K. L. (1998): *Numerical Methods in Economics*. MIT Press.

Katz, M. L. (1991): "Game-Playing Agents: Unobservable Contracts as Precommitments," *RAND Journal of Economics*, 22(3), 307–328.

Kaya, A. (2009): "Repeated Signaling Games," *Games and Economic Behavior*, 66(2), 841–854.

Kreps, D. M., P. Milgrom, J. Roberts, and R. Wilson (1982): "Rational Cooperation in the Finitely Repeated Prisoners' Dilemma," *Journal of Economic Theory*, 27(2), 245–252.

Lau, L. J. (1982): "On Identifying the Degree of Competitiveness from Industry Price and Output Data," *Economics Letters*, 10(1-2), 93–99.

MacKay, A., and N. H. Miller (2019): "Estimating Models of Supply and Demand: Instruments and Covariance Restrictions," Discussion paper, Harvard University.

Mailath, G. J. (1987): "Incentive Compatibility in Signaling Games with a Continuum of Types," *Econometrica*, 55(6), 1349–1365.

Mailath, G. J. (1988): "An Abstract Two-Period Game with Simultaneous Signaling: Existence of Separating Equilibria," *Journal of Economic Theory*, 46(2), 373–394.

——— (1989): "Simultaneous Signaling in an Oligopoly Model," *Quarterly Journal of Economics*, 104(2), 417–427.

Mailath, G. J., and E.-L. von Thadden (2013): "Incentive Compatibility and Differentiability: New Results and Classic Applications," *Journal of Economic Theory*, 148(5), 1841–1861.

Maskin, E., and J. Tirole (1988): "A Theory of Dynamic Oligopoly, II: Price Competition, Kinked Demand curves, and Edgeworth Cycles," *Econometrica*, pp. 571–599.

Mester, L. J. (1992): "Perpetual Signalling with Imperfectly Correlated Costs," *RAND Journal of Economics*, 23(4), 548–563.

Milgrom, P., and J. Roberts (1982): "Limit Pricing and Entry Under Incomplete Information: An Equilibrium Analysis," *Econometrica*, 50(2), 443–459.

Miller, N. H., G. Sheu, and M. C. Weinberg (2020): "Oligopolistic Price Leadership and Mergers: The United States Beer Industry," Discussion paper, Georgetown University.

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

Murphy, K. J. (1999): "Executive Compensation," *Handbook of Labor Economics*, 3, 2485–2563.

Nakamura, E., and D. Zerom (2010): "Accounting for Incomplete Pass-Through," *Review of Economic Studies*, 77(3), 1192–1230.

Nevo, A. (1998): "Identification of the Oligopoly Solution Concept in a Differentiated-Products Industry," *Economics Letters*, 59(3), 391–395.

O'Brien, D. P., and S. C. Salop (1999): "Competitive Effects of Partial Ownership: Financial Interest and Corporate Control," *Antitrust Law Journal*, 67, 559–614.

Olley, G. S., and A. Pakes (1996): "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 64(6), 1263–1297.

Ordover, J. A. (2007): "Coordinated Effects in Merger Analysis: An Introduction," *Columbia Business Law Review*, (2), 411–436.

Pakes, A., and P. McGuire (1994): "Computing Markov-Perfect Nash Equilibria: Numerical Implications of a Dynamic Differentiated Product Model," *RAND Journal of Economics*, 25(4), 555–589.

Peters, C. (2009): "Evaluating the Performance of Merger Simulation: Evidence from the U.S. Airline Industry," *Journal of Law and Economics*, 49(2), 627–649.

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

Roddie, C. (2012): "Signaling and Reputation in Repeated Games, I: Finite Games," Discussion paper, University of Cambridge.

Rotemberg, J. J., and D. S. Scharfstein (1990): "Shareholder-Value Maximization and Product-Market Competition," *Review of Financial Studies*, 3(3), 367–391.

Seim, K. (2006): "An Empirical Model of Firm Entry with Endogenous Product-Type Choices," *RAND Journal of Economics*, 37(3), 619–640.

Shapiro, C. (1986): "Exchange of Cost Information in Oligopoly," *Review of Economic Studies*, 53(3), 433–446.

Sklivas, S. D. (1987): "The Strategic Choice of Managerial Incentives," *RAND Journal of Economics*, 18(3), 452–458.

Smith, A. (2008): "Indirect Inference," *The New Palgrave Dictionary of Economics*, 2nd Edition.

Stigler, G. J. (1964): "A Theory of Oligopoly," *Journal of Political Economy*, 72(1), 44–61.

Sweeting, A., J. W. Roberts, and C. Gedge (2020): "A Model of Dynamic Limit Pricing with an Application to the Airline Industry," *Journal of Political Economy*, 128(3), 1148–1193.

Toxvaerd, F. (2008): "Strategic Merger Waves: A Theory of Musical Chairs," *Journal of Economic Theory*, 140(1), 1–26.

——— (2017): "Dynamic Limit Pricing," *RAND Journal of Economics*, 48(1), 281–306.

Vives, X. (2011): "Strategic Supply Function Competition with Private Information," *Econometrica*, 79(6), 1919–1966.