Dynamic Oligopoly Pricing with Asymmetric Information: Implications for Horizontal Mergers*

Andrew Sweeting†
University of Maryland and NBER

Xuezhen Tao
Shanghai University of Finance and Economics

Xinlu Yao
University of Maryland

October 2019

Abstract

Almost all empirical models of competition in differentiated product markets assume that firms have complete information and set prices to maximize current profits. Building on a small theoretical literature, we develop a dynamic model where each firm has private information about a serially-correlated state variable, such as its marginal cost, and sets its price to signal information to rivals. We find that even limited amounts of private information can raise equilibrium prices significantly, and that failing to account for signaling effects can lead conventional merger simulation calculations to substantially underpredict post-merger price increases. We structurally estimate our model using data from the beer industry, and find that our model predicts observed changes in price levels and price dynamics following the 2008 MillerCoors joint venture.

JEL CODES: D43, D82, L13, L41, L93.
Keywords: signaling, strategic investment, asymmetric information, oligopoly pricing, dynamic pricing, horizontal mergers.

*We thank Dan Vincent, Nate Miller, Bob Majure, Joe Harrington, Kyle Bagwell and many seminar participants for valuable comments, which have been especially useful in framing the paper. Carl Mela and Mike Kruger helped to provide access to the SymphonyIRI Group Inc. data, and all analyses are by the authors, and not SymphonyIRI. We have benefited from Carl Mela and John Singleton's work with the beer data on another project. This research was supported by NSF Grant SES-1658670 (Sweeting). All errors are our own.

†Corresponding author. Contact: Department of Economics, 3135 Tydings Hall, University of Maryland, College Park, MD 20742, USA. Email: atsweet@umd.edu.

1 Introduction

Empirical, theoretical and antitrust analyses of competition in differentiated oligopoly markets typically assume that firms set prices to maximize their current profits (i.e., statically) in a complete information environment, i.e., all firms know their rivals' payoff functions exactly. These assumptions provide tractability, but while competitors are likely to be well-informed about their rivals, the assumption that information is really complete appears to be inconsistent with how firms rarely disclose information about the profitability of individual product lines and with how government agencies treat all firm-specific cost and margin data as highly confidential.[1]

This paper investigates what may happen when we relax the static and complete information assumptions. Specifically, each firm will have a payoff-relevant state variable (the level of marginal cost in our leading example) which is both unobserved by rivals and imperfectly serially-correlated over time. The serial correlation can give rise to dynamics where rivals draw inferences from a firm's chosen price about the value of its state variable, and the prices that they expect the firm to set in future periods. This, in turn, can lead to a firm choosing its current price in order to affect rivals' inferences. We will specifically consider fully separating equilibria where a firm's chosen price perfectly reveals the value of its state variable in equilibrium.

While a theoretical literature has modeled this type of oligopoly signaling in two or three period models with linear demand, we provide, as far as we are aware, the first analysis of the magnitude of the effects on prices, and we consider both infinite and finite horizon models, using nested logit demand to model substitution between differentiated products. Our main finding is that equilibrium prices can be much higher than in a static complete information model with similar parameters. Price effects can be large even when asymmetric information is limited: for example, we present one duopoly example where equilibrium prices increase by more than 40% when there is uncertainty about less than 1% of each firm's marginal cost.

To understand why effects can be large, suppose that the marginal cost of each firm is private information and that a firm's sales increase in its rivals' prices. In a fully separating equilibrium, all firm types that do not have the lowest possible marginal cost raise their prices above static best response levels to signal this information to their rivals. The incentive to signal a higher

[1] A Wall Street Journal article on May 23, 2016, describes how Anthem and Cigna, two health insurers who were attempting to merge refused to share cost information with each other as they would be competitors if the merger was blocked.

cost comes from the fact that rivals will then expect the firm to be more likely to have a higher cost and so be likely to set a higher price in the next period, causing the rivals to raise their next period prices as well. When all firms are signaling, static best response prices will increase as well, causing the signaling prices set by higher cost types to rise further, and this may also tend to increase the incentives of each firm to signal higher costs in earlier periods. Depending on the curvature of the profit functions and the degree to which firms lose demand to the outside good (of not purchasing) when they raise prices, these positive feedback effects can cause equilibrium prices to rise significantly. We show how features of demand, such as diversion ratios, as well as the serial correlation in the cost process, can affect both equilibrium prices and the existence of separating equilibria, as, with non-linear demand, these equilibria can cease to exist when prices rise too much.

Price increases in our model are strategic investments in raising rivals' future prices, and the incentive to invest is stronger when there are fewer firms. We therefore focus on what our framework predicts about the effects of horizontal mergers. Conventional merger simulation calculations, which assume complete information, will tend to underpredict post-merger price increases insofar as they ignore how the merger increases signaling incentives.

As an empirical application, we consider the US market for light beer, motivated by Miller and Weinberg (2017) (MW). MW show that following the 2008 joint venture (JV) between SAB Miller and Coors (forming MillerCoors, MC), which effectively merged their US brewing and marketing assets, the prices of both MC brands and Anheuser-Busch (AB) brands, which were not part of the JV, increased. MW show that this pattern is inconsistent with static, complete information pricing before and after the JV, leading them to suggest that the JV changed the nature of firm strategies (e.g., firms started to collude). We structurally estimate our model using pre-JV data and show that, for a range of alternative demand parameters, it can accurately predict the observed post-JV price increases for both MC and AB, as well as several qualitative changes in observed price dynamics. Therefore, if firms were always signaling, no change in the nature of behavior is required to explain what happened after the JV. As far as we are aware, we provide the first structural estimation of a model where multiple players simultaneously signal.

Before discussing the related literature, we should be clear about several limitations of our analysis. First, tractability limits us to specifications where each firm has at most one dimension of private information and one signal (price) that it can send. Second, we will only consider fully separating equilibria even though one can construct examples where firms engage in some type of pooling. Third, while theory allows us to characterize firms' unique separating best-response strategies, we cannot prove that fully separating equilibria are unique when using nonlinear demand. Fourth, we do not try to create an empirical horse-race between our model and tacit collusion models. This type of comparison is difficult because folk theorems show that many patterns of behavior can be supported in a collusive equilibrium if firms are patient enough. Instead, we restrict ourselves to illustrating that our model, estimated using only pre-JV data, is able to match several observed changes accurately.

Following a review of the related literature, the remainder of the paper is organized as follows. Section 2 lays out the model and the equilibrium concept. Section 3 presents duopoly examples to illustrate the mechanisms at work and the effects of signaling. Section 4 extends the analysis to more firms and presents a stylized merger analysis. Section 5 provides our empirical application. Section 6 concludes. The Appendices detail the computational algorithms and some additional results.

1.1 Related Literature

The paper is related to several literatures. Shapiro (1986) compares welfare in one-shot Cournot models where marginal costs are either public (complete) or private information. He shows that complete information increases expected total welfare, but reduces expected consumer surplus. Vives (2011) shows that private information about costs can raise prices in a one-shot game where firms submit supply functions (for example, in some types of multi-unit auction). Given the parameters that we consider, the welfare effects of private information with no dynamics would be small.

A large theoretical literature has examined signaling models where only one player has private information. The most well-known example in Industrial Organization is the Milgrom and Roberts (1982) limit pricing model, where an incumbent may lower its first period price to signal that entry will not be profitable in a two period game. Kaya (2009) and Toxvaerd (2017) consider games with more periods where the type of the informed firm is fixed. The equilibrium involves the informed firm signaling until its reputation is established.[2] Sweeting, Roberts, and

[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 each firm's action (quantity) is not observed

Gedge (forthcoming) develop and estimate a dynamic version of the limit pricing model where the incumbent monopolist's type can change over time, as in this paper, and they show that this leads to the incumbent lowering its prices persistently. The oligopoly market structures considered in this paper are much more common, and we show how feedback effects, that only exist in the oligopoly model, can lead to large price effects even when the privately known state variables have quite narrow supports.

There is a small theoretical literature on games where multiple players signal simultaneously. 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 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] The current paper extends the Mailath (1989) model to multiple periods, but with time-varying, serially correlated types and the type of nonlinear demand usually assumed when modeling differentiated products. Our focus will be on the magnitude, and potential policy implications, of the equilibrium price changes caused by signaling. Our extensions impose some loss of tractability: while Mailath could prove that only one fully-separating equilibrium exists, the proof does not extend to non-linear demand or the infinite horizon case. Mester (1992) considers a three-period quantity-setting version of the Mailath model but allowing for costs to change over time. Consistent with our results, she shows that signaling happens in the first two periods of the game, but, because quantities are usually strategic substitutes, the equilibrium involves firms increasing their output.

The literature on fully dynamic oligopoly models has followed Ericson and Pakes (1995) and Pakes and McGuire (1994) in assuming complete information up to payoff shocks that are independent over time, so they create no incentives to signal. Nakamura and Zerom (2010) use this framework to model oligopoly pricing by firms with menu costs. Departing from this

and cannot be exactly inferred because of unobserved demand shocks. We 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.

framework, Fershtman and Pakes (2012) propose the concept of (Restricted) Experience Based Equilibrium (EBE) for games with discrete states and discrete actions, some of which may be private information to particular firms. This approach has been used by Asker, Fershtman, Jeon, and Pakes (2018) to solve a duopoly repeated auction model. EBE is distinguished from Perfect Bayesian concepts by being defined in terms of expected payoffs from different actions, rather than beliefs about rivals' types, which simplifies computation. We use the more standard type of equilibrium concept, focusing on fully separating equilibria where the computational burden is reduced by the simple structure of beliefs on the equilibrium path. The resulting computational burden is small enough that we can structurally estimate our model.

Our paper is motivated by the finding of the empirical merger retrospectives literature (*inter alia* Borenstein (1990), Kim and Singal (1993), Peters (2009), Ashenfelter, Hosken, and Weinberg (2015) and MW) that prices often rise after completed mergers.[4] Given that the antitrust authorities try to prevent mergers that will raise prices, this is a striking empirical result, and it suggests that the conventional analysis of mergers may miss some incentive that can cause prices to rise, such as the one that we examine here. We will discuss the relationship between our model and existing "coordinated effects" theories of harm from mergers in the conclusion.

2 Model

In this section we present our general model. We will make more specific assumptions when we present our examples and application.

There are discrete time periods, $t = 1, ..., T$, where $T \leq \infty$, and a common and known discount factor $0 < \beta < 1$. There are a fixed set of N risk-neutral firms, and no entry and exit. Firms have unidimensional types, which are private information. We will consider cases where firm i's type can lie anywhere on a known compact interval $[\underline{\theta_i}, \overline{\theta_i}]$ (a "continuous type" model) and where it can only take two possible discrete values, $\underline{\theta_i}$ and $\overline{\theta_i}$ ("two-type"). In both cases, the types evolve exogenously, and independently, from period-to-period according to a

[4]The survey by Ashenfelter, Hosken, and Weinberg (2014) classifies 36 of 49 studies as finding significant post-merger price increases across a range of industries, although there are no systematic effects after banking mergers. Notably, the Mester (1992) analysis of oligopoly signaling in a Cournot framework, where signaling can lead to more competitive outcomes, was motivated by the observation that concentration and multi-market contact appear to lead to more competitive behavior in the banking industry, which is inconsistent with what standard models of tacit collusion under complete information would predict.

first-order Markov process, $\psi_i : \theta_{i,t-1} \to \theta_{i,t}$.[5] We will only consider processes with positive correlation in the sense that a higher value of $\theta_{i,t-1}$ will raise the probability of higher values of $\theta_{i,t}$. 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, in the sense that the type can transition to any value on the support in the following period.*

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

3. *for any $\theta_{i,t-1}$ there is some θ' 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 $\int_{\underline{\theta_i}}^{\overline{\theta_i}} \frac{\partial \psi_i(\theta_{i,t}|\theta_{i,t-1})}{\partial \theta_{i,t-1}} d\theta_{i,t} = 0$.*

In each period t of the game, timing is as follows.

Assumption 2 ***Stage Game Timing***

1. *Firms enter period t with their $t-1$ types, which then evolve according to the processes ψ_i. Firms observe their own new type, but not the types of other firms.*

2. *Each firm simultaneously sets a price, $p_{i,t}$, for period t. Once a firm sets its 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 $-i$.*

The assumptions on firm profits are consistent with firms selling substitute products and having no menu costs, and demand being static. The assumption that firms are committed to a price once it has been chosen is consistent with Shapiro (1986) and Mailath (1989). We assume that firms are able to observe the complete history of prices in the game but that past rival types are not observed.[6] We make the following assumption about what firms believe about rivals' types entering the first period of the game.

[5] While the assumption that marginal costs evolve exogenously may seem unrealistic, we note that some or all innovations in marginal cost are treated as exogenous throughout the existing empirical literature on horizontal mergers, as well as in most of the production function literature that has followed the approach of Olley and Pakes (1996) (see Doraszelski and Jaumandreu (2013) for an exception).

[6] The equilibria that we consider would still be equilibria in a game where types were publicly revealed with a lag of at least two periods.

Assumption 3 ***Initial Period Beliefs.*** *Firms know what their rivals' types were in a fictitious prior period, $t = 0$.*

Under this assumption, equilibrium separating strategies will be the same in $t = 1$ as in every other period of a stationary infinite horizon game where firms used fully separating strategies in the previous period, as in all periods firms will act as if they know the previous period types of all firms. This assumption simplifies the presentation of the results, because we do not need to caveat our statements by noting that strategies in the first period will be different. However, we could make other assumptions as long as all players have consistent beliefs at the start of the game and those beliefs support separating strategies in the first period.[7]

2.1 Separating Equilibrium in a Game with Finite Horizon and Continuous Types

We now explain the equilibrium for the finite horizon model with continuous types. For simplicity, we focus on the case of single product duopolists. Throughout the paper we will focus on equilibria that are fully separating, meaning that there is a one-to-one mapping between a player's type and the equilibrium price that it sets conditional on beliefs about firm types in the previous period.

Final Period (T). In the final period, each firm should maximize its expected payoff given its own type, its beliefs about the types of the other firms and their pricing strategies. This corresponds to the Bayesian Nash Equilibrium considered by Shapiro (1986). If firm j believes that firm i's $T-1$ type was $\widehat{\theta^j_{i,T-1}}$ and j's period T pricing function is $P_{j,T}(\theta_{j,T}, \theta_{j,T-1}, \widehat{\theta^j_{i,T-1}})$[8], then i will set a price

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

Earlier Periods ($1, .., T-1$). In earlier periods, there may be incentives to distort pricing choices in order to affect the rival's belief about pricing in the next period. The equilibrium

[7]For example, we might assume that each firm's first period type is drawn from the steady state distribution of its type.

[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.

concept that we use is Markov Perfect Bayesian Equilibrium (MPBE) (Roddie (2012), Toxvaerd (2008)). This requires, for each period:

- a period-specific pricing strategy for each firm as a function of its current type, its beliefs about the type of its rival in the previous period, and what it believes to be its rival's belief about its own previous type; and,

- a specification of each firm's beliefs about its rival's type given all possible observed histories of the game. These beliefs should be consistent with the application of Bayes Rule given equilibrium pricing strategies.[9]

While only current types and prices are directly payoff-relevant, history can matter in this Markovian equilibrium because history affects beliefs. We focus on fully separating equilibria, where, given any history, a firm's equilibrium pricing strategy will perfectly reveal its current type, and beliefs on the equilibrium path have a simple form. Mailath (1989) argues that when a fully separating equilibrium exists, it is the natural one to look at.

We follow Mailath (1989)'s characterization of fully separating best response functions. To do so, we define firm i's period-specific "signaling payoff function", $\Pi^{i,t}(\theta_{i,t}, \widehat{\theta^j_{i,t}}, p_{i,t})$, which is the present discounted value of firm i's expected current and future payoff when it has current type $\theta_{i,t}$, sets current price $p_{i,t}$ and the rival believes, *at the end of period* t, that the firm has type $\widehat{\theta^j_{i,t}}$. The function is implicitly conditional on the strategies that the rival uses in the current period and both players will use in future periods, as well as i's beliefs about their types at $t-1$ (this conditioning is made explicit in Appendix A where we describe our computational algorithm). As the rival's end-of-period t belief about firm i's type enter as a separate argument, the only effect of i's period t price, $p_{i,t}$, on $\Pi_{i,t}$ is through period t profits although, of course, in equilibrium, it will be $p_{i,t}$ that affects the rival's belief. A fully separating best response function can be characterized as follows, under some conditions on $\Pi^{i,t}$ that will be listed in a moment.

Characterization of Separating Best Response Pricing Functions in Period $t < T$. A firm will have a set of best response pricing functions reflecting beliefs about costs in the previous period. Each pricing function will be the solution to a set of differential equations

[9] When there are multiple rivals, they should all have the same beliefs given an observed history.

where

$$\frac{\partial p^*_{i,t}(\theta_{i,t})}{\partial \theta_{i,t}} = -\frac{\Pi_2^{i,t}\left(\theta_{i,t}, \widehat{\theta^j_{i,t}}, p_{i,t}\right)}{\Pi_3^{i,t}\left(\theta_{i,t}, \widehat{\theta^j_{i,t}}, p_{i,t}\right)} > 0, \tag{1}$$

where the subscript n in $\Pi_n^{i,t}$ denotes the partial derivative of the signaling payoff function with respect to the n^{th} argument, and a boundary condition. This characterization assumes that Π is continuous and differentiable in its arguments. Assuming that it is natural for lower types to set lower prices (e.g., a type corresponds to the firm's marginal cost), this boundary condition will be that $p^*_{i,t}(\underline{\theta_i})$ is the solution to

$$\Pi_3^{i,t}\left(\underline{\theta_i}, \widehat{\theta^j_{i,t}}, p_{i,t}\right) = 0 \tag{2}$$

i.e., the lowest type's price should be a static best response to j's pricing policy. The numerator in the differential equation represents the marginal future benefit from raising a rival's belief and the denominator is the effect of a price increase on current profit. Above the 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 best response function will be valid under four conditions on $\Pi_{i,t}$.

Condition 1 ***Shape of the Static Profit Function 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* $\Pi_{33}^{i,t}\left(\theta_{i,t}, \widehat{\theta^j_{i,t}}, p_{i,t}\right) > 0$*, there is some* $k > 0$ *such that* $\left|\Pi_3^{i,t}\left(\theta_{i,t}, \widehat{\theta^j_{i,t}}, p_{i,t}\right)\right| > k$*.*

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

Condition 3 ***Belief Monotonicity.*** $\Pi_2^{i,t}\left(\theta_{i,t}, \widehat{\theta^j_{i,t}}, p_{i,t}\right)$ *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.*** $\frac{\Pi_3^{i,t}\left(\theta_{i,t}, \widehat{\theta^j_{i,t}}, p_{i,t}\right)}{\Pi_2^{i,t}\left(\theta_{i,t}, \widehat{\theta^j_{i,t}}, p_{i,t},\right)}$ *is a monotone function of* $\theta_{i,t}$ *for all* $\widehat{\theta^j_{i,t}}$ *and for relevant* $p_{i,t}$ *above the static best response price.*

To interpret these conditions, suppose that types correspond to firm 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 depends on the assumed form of demand and the condition will hold for the multinomial logit and nested logit models that are commonly used for differentiated products. Type monotonicity requires that, when a firm increases its price, its expected current period profit will decrease by less if it has higher marginal costs. This will hold for linear 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. While it is natural to expect that rivals would respond to this type of belief by raising their prices, we will see examples where this is not the case, causing the belief monotonicty condition to fail. 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. If belief monotonicity fails, single crossing is also likely to fail.

This characterization uniquely defines separating best response functions with continuous types. These best response functions define strategies in a fully separating equilibrium in period t if each firm is making a best response to the pricing strategy of its rival given the beliefs that each firm has about its rival's type. Given these strategies, each firm will have a point belief about its rival's type at the end of the period, formed by inverting back from the observed price to the type using the pricing function. On the equilibrium path, these beliefs will be correct.

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.

2.2 Separating Equilibrium in Alternative Forms of the Game

While we use the finite horizon, continuous type model to develop intuition about the effects of signaling on prices, we use an infinite horizon model in our empirical application because it can be solved more quickly. We will also use a model where a firm's type can take on only two values to illustrate when separating equilibria may fail to exist. We briefly note here how equilibrium strategies change in these models.

Infinite Horizon Version of the Model. In the infinite horizon version of the model, fully separating strategies can be characterized in the same way as those in $t < T$ periods of the finite horizon game, but they will be stationary (i.e., independent of t). We will illustrate how these strategies reflect the limits of the strategies early in a finite horizon game as we increase the number of periods.

Two-Type Model. The type evolution will be represented by a single probability (greater than 0.5 but strictly less than 1) that a firm's type will remain the same. The best response pricing strategy in this model cannot be characterized using a differential equation, and in general there can be many separating best responses depending on how signals are interpreted. We therefore consider only best response pricing strategies that are consistent with refinements such as the "intuitive criterion" (Cho and Kreps (1987)), that identify the separating strategy that has the least cost, in terms of current profits lost by raising price, to the firm sending the signal, holding the conjectured pricing behavior of the other firm fixed. When marginal costs are private information, the low cost type's strategy will be the static best response, as in the continuous type model, and, under assumptions that appropriately map Conditions 1-4 to the two-type case, the high cost type's price will be the lowest price that the low type would be unwilling to set if this would result in rivals' perceiving it as a high type rather than a low type.

Complete Information Version of the Model. When comparing average equilibrium prices in our model to the averages under complete information, we assume that in the latter case firms would use static Nash equilibrium pricing strategies in every period. This is, of course, the only subgame perfect Nash equilibrium in the finite horizon version of the complete information game.

2.3 Computation

Appendix A details the computation of the equilibrium strategies. As belief monotonicity and single-crossing depend on the prices that firms charge, we verify these conditions during computation. When our algorithm finds a fully separating equilibrium, it may not be unique (without the linear demand and finite horizon assumptions of Mailath (1988), Mailath (1989) and Mester (1992)). In the continuous type model and the finite horizon two-type model we have been unable to find more than one equilibrium for a given set of parameters when we have varied starting values or the details of our computational procedure. In the two-type, infinite horizon duopoly model we have found examples where we can identify three equilibria that produce different average prices even when we use the "intuitive criterion" refinement on best responses.[10] In these examples, our algorithm converges to the equilibrium with the highest average prices and where strategies are the same as in the early periods of a long, finite horizon game.

3 Duopoly Examples

This section presents several examples to show how and why signaling affects equilibrium prices.

3.1 Demand Specification and Discount Factor

We assume that demand is determined by a discrete choice nested logit model, where each firm has a single product. All products are grouped into a single nest, with the outside good of not purchasing in its own nest. Our initial examples use an indirect utility function where consumer c's utility from buying the product of firm i is $u_{i,c} = 5 - 0.1p_i + \sigma\nu_c + (1-\sigma)\varepsilon_{i,c}$ where p_i is the price of firm i in dollars, ε is a draw from a Type I extreme value distribution, $\sigma = 0.25$, and ν_c is an appropriately distributed draw for c's nest preferences. For the outside good, $u_{0,c} = \varepsilon_{0,c}$. For these parameters, combined market shares will be high at static Nash equilibrium prices for the costs that we consider, implying limited diversion to the outside good, and equilibrium mark-ups will be substantial. As we will illustrate below, these properties will lead to large signaling effects. The assumed period-to-period discount factor, in both our examples and our application, is 0.99. We set market size equal to one, so profits, consumer surplus and welfare are reported on a "per potential consumer" basis.

3.2 Marginal Cost Types

Our primary example assumes that the level of each firm's linear marginal cost is private information. The duopolists are ex-ante symmetric with the marginal cost of each firm lying in the interval $[\underline{c}, \overline{c}] = [\$8, \$8.05]$. We interpret the level of uncertainty as being "small" because a firm's marginal cost can diverge by less than 0.3% from its median possible value. Costs evolve

[10] In the two-type model it is possible search for multiple equilibria systematically by solving systems of non-linear equations from several thousand starting points.

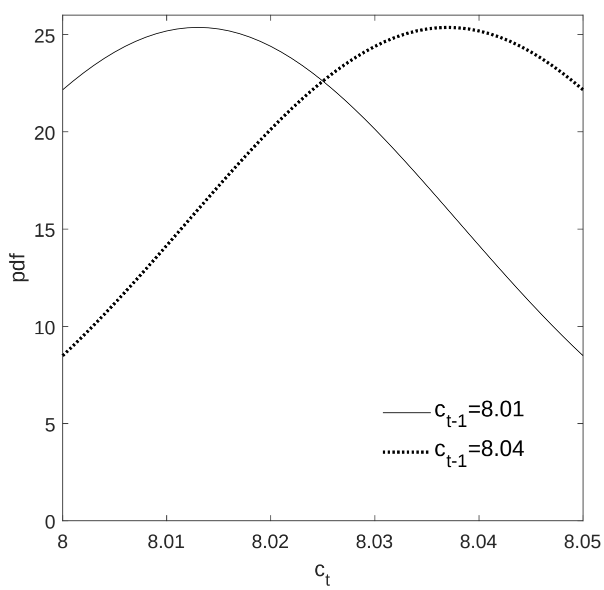
independently according to an exogenous truncated AR(1) process where

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

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

We assume $\sigma_c = 0.025$. As illustrated in Figure 1, the assumed values of $\underline{c}, \overline{c}$ and σ_c imply that a firm's marginal cost is quite likely to move from being relatively high to relatively low from one period to the next. For example, the probability that a firm with the highest marginal cost will have a marginal cost in the lower half of the support in the next period is 0.32. This implies that knowledge of a rival's previous period marginal cost should not have too much effect on the cost that a firm expects the rival to have, which tends to limit signaling incentives. We will see, however, that we can still generate large equilibrium effects on prices.

Figure 1: Conditional PDF of $c_{i,t}$ for Two Values of $c_{i,t-1}$ Given the Assumed Cost Process



3.3 Baseline Results: Finite Horizon Model

We begin by considering a finite horizon game.

Figure 2: Period T and $T-1$ Pricing Strategies in the Finite Horizon, Continuous Type Signaling Game

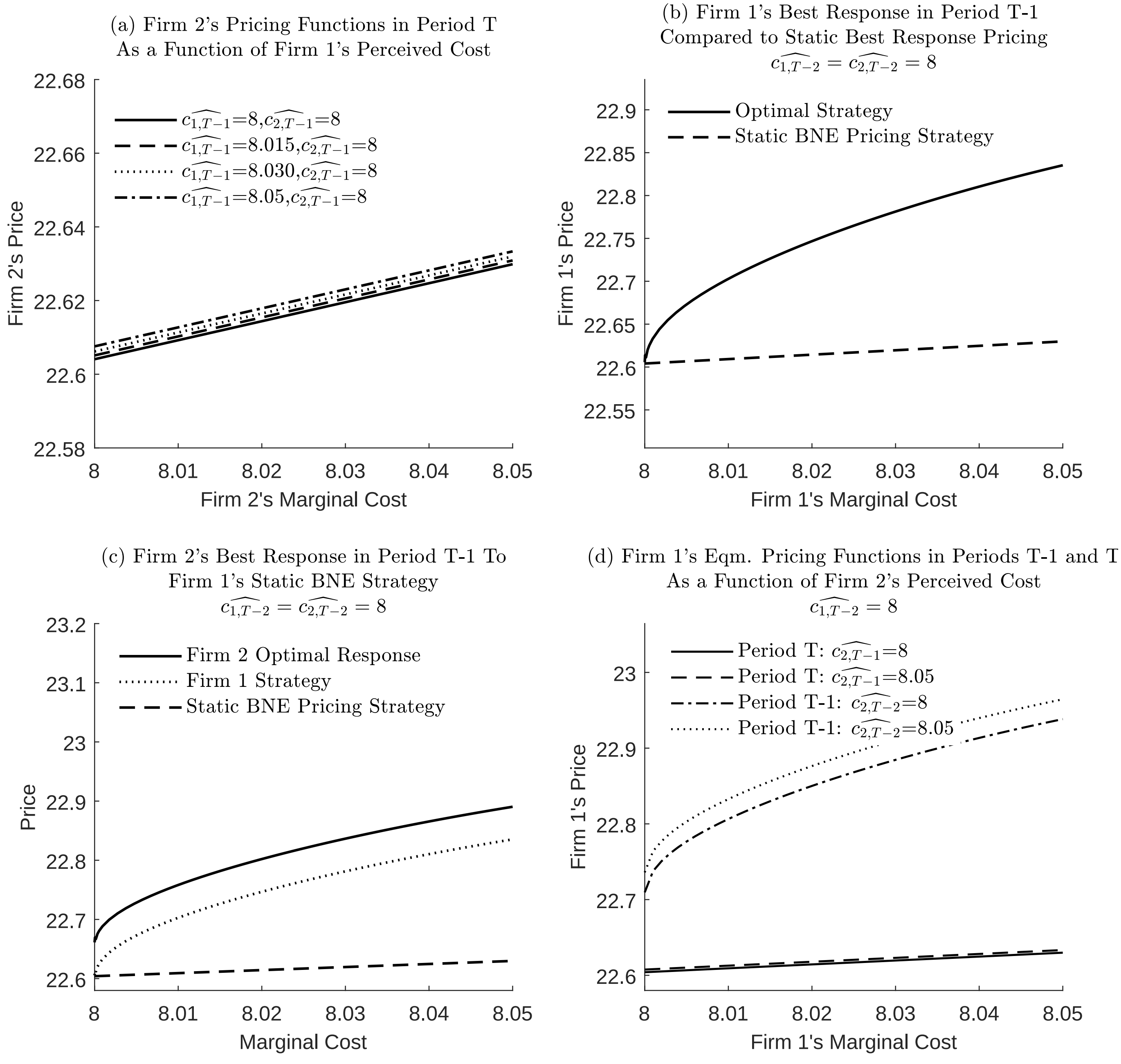
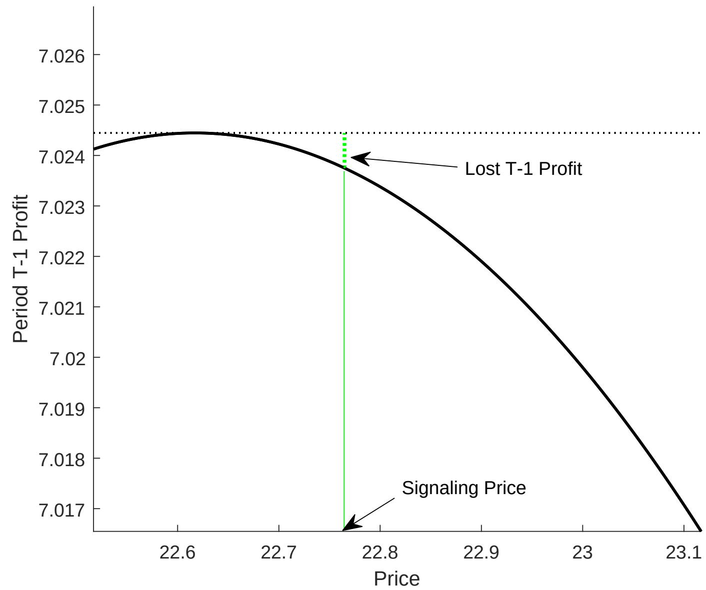


Figure 3: Expected $T-1$ Period Profit Function



Note: 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.

3.3.1 Pricing in Period T

Assume that the game has at least two periods, and that equilibrium play in period $T-1$ reveals $T-1$ costs. Figure 2(a) shows four static BNE (i.e., period T) pricing functions for firm 2, for different values of firm 1's $T-1$ marginal cost, when both firms know/believe that $c_{2,T-1}=8$. Firm 2's price is increasing in both its own marginal cost and in firm 1's prior marginal cost because this increases the price that firm 2 expects firm 1 to set in period T. However, the variation in firm 1's prior cost only changes firm 2's price by less than one cent. Averaging across all possible realizations of costs (using the steady state distribution of costs implied by the transition processes as weights), a firm's price has mean \$22.62 and standard deviation \$0.008. The corresponding values under complete information, where both period T costs are known, are \$22.62 and \$0.011, and expected producer and consumer surplus differ by less than \$0.0001 across the models. Therefore the existence of asymmetric information alone (i.e., when not combined with some form of dynamics) does not generate interesting welfare effects given our assumed parameters.

3.3.2 Pricing in Period T-1

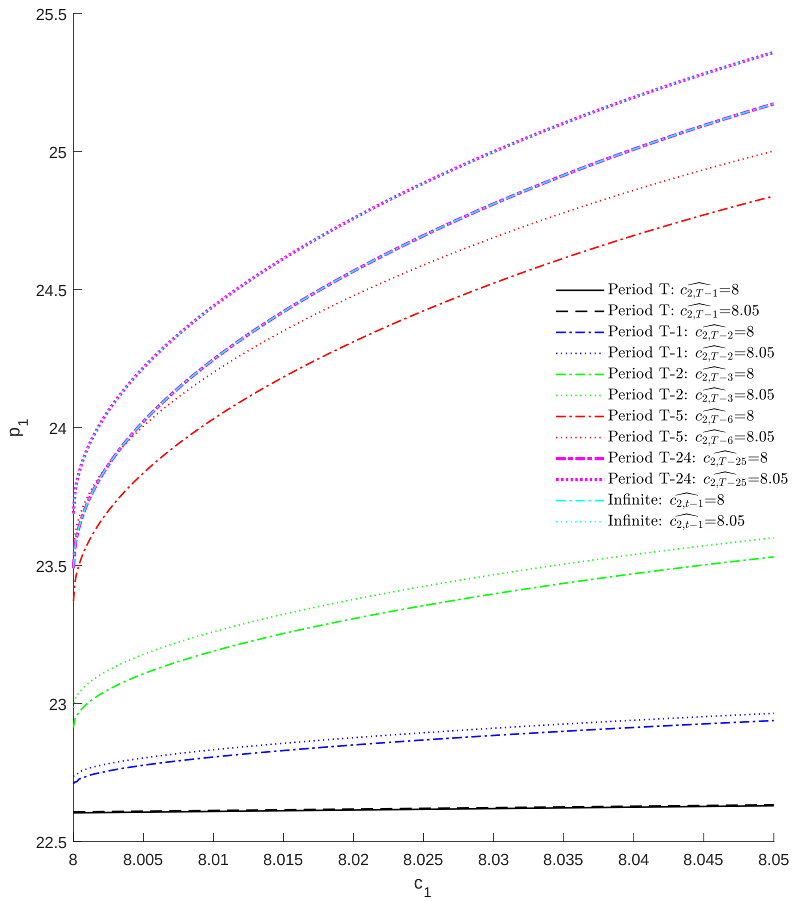
Now consider pricing in period $T-1$ where the ability to raise a rival's price in period T can lead a firm to want to signal that it has higher costs. Figure 2(b) shows firm 1's signaling best response pricing function (found by solving the differential equation in (1) given the boundary condition (2)) when firm 2 uses the pricing strategy that we predict it should actually use in period T, and compares it to firm 1's period T strategy for the same beliefs about previous costs. The figure is drawn assuming that both firms' previous marginal costs are known/believed to be equal to 8. For the lowest marginal cost, the pricing functions intersect because, as discussed in Section 2.1, firm 1's strategy in this case is a static best response, but for higher marginal costs, firm 1 may raise its price by up to 20 cents.

At first blush, this large price increase may seem surprising, because, as shown in Figure 2(a), alternative beliefs about firm 1's prior cost can raise firm 2's period T price by no more than 1 cent. However, as shown in Figure 3, drawn assuming that $c_{1,T-1} = 8.025$ and $c_{1,T-2} = c_{2,T-2} = 8$, firm 1's expected lost $T-1$ profit when it uses its signaling price of \$22.76, rather than the statically optimal $T-1$ price of \$22.61, is small (\$0.00070). This is less than the (discounted) expected period T profit gain of \$0.00079, so, in fact, firm 1 does prefer to raise its price by 16 cents in $T-1$.[11]

Figure 2(c) shows firm 2's best signaling response when firm 1 uses the strategy in Figure 2(b) (repeated in the new figure). As firm 1 charges higher prices than under static BNE prices, firm 2's static best response price increases and the pricing function shifts upwards. Of course, this type of positive feedback will also cause firm 1's pricing function to rise. Figure 2(d) shows a set of the equilibrium $T-1$ pricing functions found by iterating best responses. The $T-1$ functions imply prices that, relative to period T, are higher on average (\$22.88 vs. \$22.62), more volatile (standard deviation \$0.063 vs. \$0.008) and more sensitive to the rival's perceived prior cost (increased vertical spread). This latter property implies that signaling incentives at $T-2$ may be stronger than at $T-1$ because a higher signal can raise a rival's next period price by a larger absolute amount.

[11] If firm 1 set the price of \$22.61 in period $T-1$, it would be interpreted as having a $T-1$ marginal cost of \$8.0001 when firm 2 is expecting firm 1 to use the signaling pricing function shown in Figure 2(b).

Figure 4: Equilibrium Pricing Functions for Firm 1 in the Infinite Horizon Game and Various Periods of the Finite Horizon Game.



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

3.3.3 Pricing in Earlier Periods and in the Infinite Horizon Game

Figure 4 shows a selection of equilibrium pricing functions for $T-2$ and earlier periods. The pricing functions become more spread out in earlier periods, increasing signaling incentives in prior periods. As shown in Table 1, average equilibrium prices rise as one moves away from the end of the game, stabilizing around $T-13$ at 9.4% above static Bayesian Nash, or complete information Nash, levels, and price volatility increases. Expected firm profits are substantially higher, while consumer surplus, as well as total welfare, falls. The stationary pricing functions in the infinite horizon version of the game with the same parameters are also plotted in Figure 4, although they are indistinguishable from the $T-24$ pricing functions in the finite horizon game.

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	Signaling MPBE	24.76	0.47	30.91	15.96	46.87
T-13	Signaling MPBE	24.76	0.47	30.91	15.96	46.87
T-9	Signaling MPBE	24.74	0.47	30.93	15.94	46.87
T-8	Signaling MPBE	24.72	0.46	30.95	15.93	46.87
T-7	Signaling MPBE	24.68	0.45	30.98	15.89	46.88
T-6	Signaling MPBE	24.61	0.44	31.05	15.83	46.88
T-5	Signaling MPBE	24.48	0.41	31.18	15.72	46.89
T-4	Signaling MPBE	24.25	0.36	31.40	15.51	46.91
T-3	Signaling MPBE	23.88	0.28	31.75	15.19	46.94
T-2	Signaling MPBE	23.38	0.17	32.23	14.74	46.97
T-1	Signaling 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 Game	Signaling MPBE	24.76	0.47	30.91	15.96	46.87

Notes: values in all but the last line are based on the continuous type duopoly, finite horizon model with parameters described in the text. The last line reports results for the stationary strategies in the infinite horizon model with the same parameters.

In this sense, strategies and outcomes in the infinite horizon game are the limits of strategies in the early periods of a finite horizon game as the length of the game increases.

3.4 Stronger Signaling Incentives and the Existence of Perfectly Separating Equilibria

The first six columns of Table 2 show the time path of average prices in the finite horizon duopoly game for our previous specification and five variants (three with a larger range of costs, two with smaller variance innovations). The changes make a firm's current cost more informative about its next period cost, which strengthens signaling incentives and, as a result, increases equilibrium prices. For example, increasing $\overline{c}$ from 8.05 to 8.15 raises average $T-2$ prices from \$23.38 (3.3% above static BNE levels) to \$25.12 (11%) . However, the table also shows that the conditions required to characterize separating best response pricing strategies fail in earlier periods when

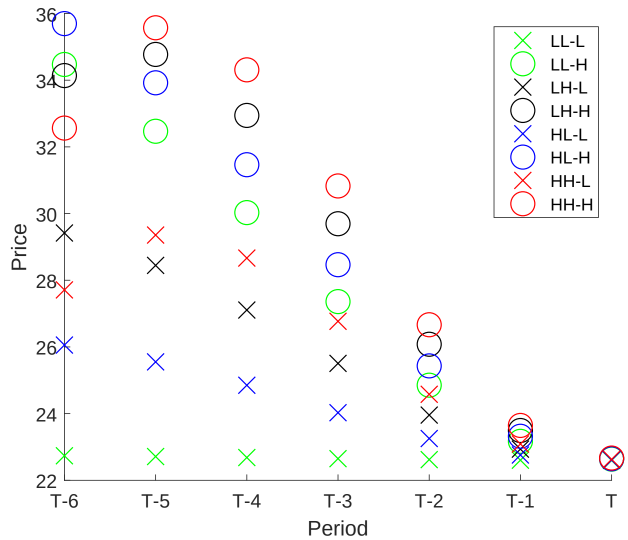
Table 2: Equilibrium Pricing with Stronger Signaling Incentives

	Baseline	Expand Range			Reduce Std. Deviation		Expand Range & Increase Std. Dev.
Interval	[8,8.05]	[8,8.075]	[8,8.15]	[8,8.3]	[8,8.05]	[8,8.05]	[8,8.50]
Std. Dev. of Cost Innovations	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-23	24.76	26.51	-	-	25.71	-	24.90
T-13	24.76	26.56	-	-	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-5	24.48	26.11	29.09	30.62	25.27	28.21	24.80
T-4	24.25	25.64	28.31	30.01	24.91	27.32	24.69
T-3	23.88	24.91	26.95	28.52	24.35	25.97	24.49
T-2	23.38	23.96	25.12	26.25	23.61	24.33	24.13
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
Infinite Horizon Game	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.

we strengthen signaling incentives too much.[12]

Figure 5: Equilibrium Prices in the Two-Type Marginal Cost Model



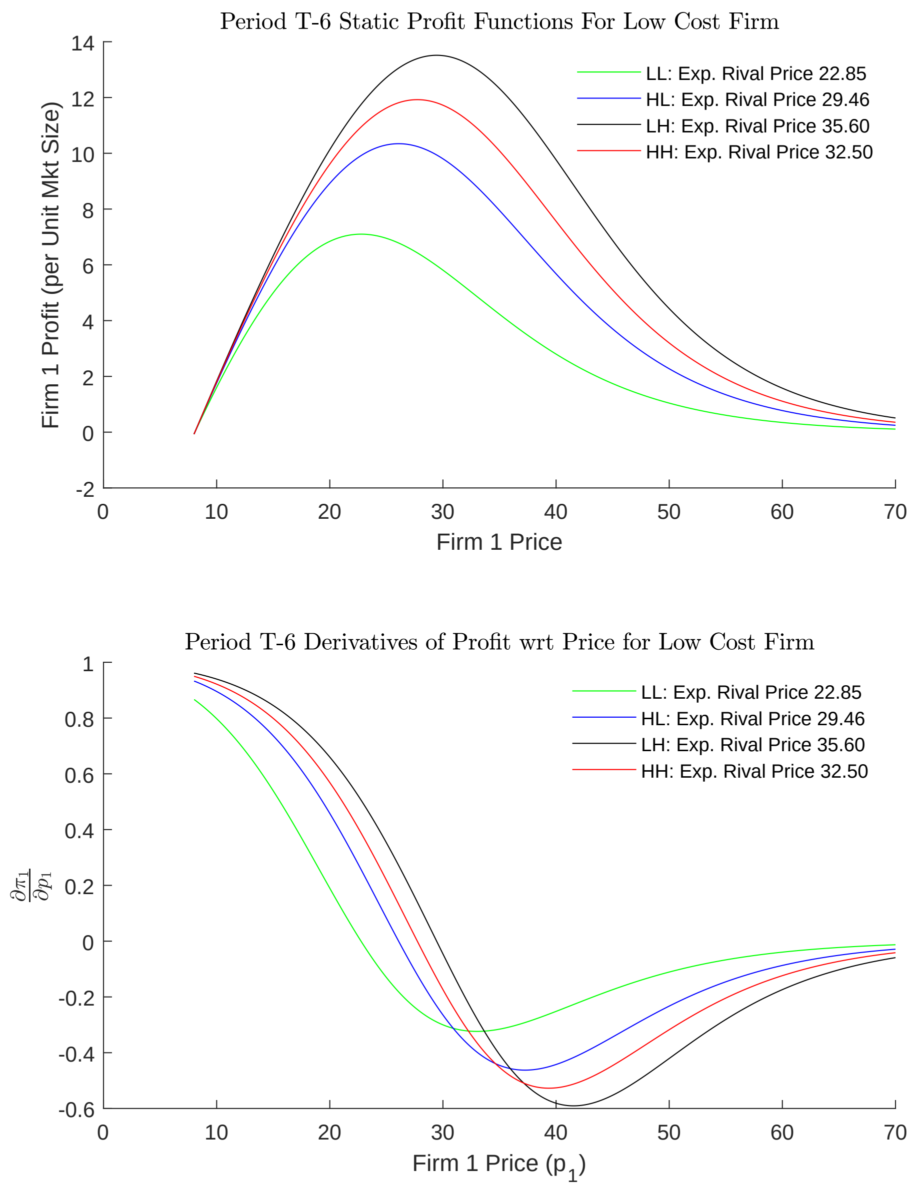
The reason for failure can be understood by considering a finite horizon duopoly game where each firm's marginal cost can only take on two values. Demand is the same as before, but each firm's marginal cost is either 8 (low) or 8.05 (high). We assume that the probability that a firm's marginal cost does not change across periods is 0.99 in order to generate very strong signaling incentives. Figure 5 shows the full set of eight equilibrium prices in each period as we move backwards from the end of the game. The legend denotes states by {"the firm's perceived cost in $t-1$", "its rival's perceived cost in $t-1$" - "the firm's realized marginal cost in t"} so blue indicates prices for a firm whose perceived marginal cost in the previous period was high, its rival's perceived previous period cost was low, and a cross (circle) indicates that the firm's current cost is low (high). In this game, every equilibrium price can be shown simultaneously on a single diagram.

The green crosses (LL-L) remain almost unchanged across periods, as they represent static best responses when both players know that their rival is very likely to be setting the same price, but, as we move earlier in the game, the remaining prices increase, because they involve either signaling (by a $\bar{c}$ firm) or a static best response to a rival who is likely to be raising its price to signal. In $T-6$ the order of the prices changes with the HH-H price (red circle) below the HL-H price (blue circle). This implies that in period $T-7$, a firm that believes its rival is likely to be

[12]It is noticeable that for specifications where conditions fail in period t, average prices are lower in period $t+1$ than in period $t+2$. We have also found cases where conditions can fail without this pattern.

high cost, is more likely to increase its rival's next period $(T-6)$ price if it (the firm) is believed to be *low* cost. As profits increase in the rival's price, this will lead belief monotonicity to be violated.

Figure 6: Period $T-6$ Profit Functions in the Two-Type Game



Why does the order of the red and blue circles switch? It reflects changes in both the incentive to signal (i.e., the possible effect on future prices) and the cost of signaling (i.e., the effect on current profits). Recall that in the two-type model the equilibrium price of the $\overline{c}$ type is determined by the lowest price that the low-cost firm would be unwilling to set even if choosing it would lead to it being perceived as high cost. Consider the cost, in terms of foregone $T-6$ profit, for a low-cost firm of raising its price. The upper panel of Figure 6 shows the period $T-6$ one-period profit functions for a low cost firm given different beliefs about previous firm types and the expected price of the rival.[13] The lower panel shows the corresponding derivatives of the profit function with respect to the firm's own price. For prices above \$34, the marginal loss in profit from a price increase is greater for a red firm than a blue firm so it is less costly for the blue firm to raise its price.[14]

Now consider the incentive of a low-cost firm to signal (i.e., to pretend to be high-cost). The incentive of an HL (blue) firm to signal a high cost at $T-6$ is that it is very likely to lead to its rival setting the black cross, rather than the green cross, price at $T-5$. This difference is large, so that the incentive to signal is strong. The incentive of an HH (red) firm to signal is that this will very likely lead to it facing the red, rather than the blue, circle price at $T-5$. These $T-5$ prices are more similar (than the black and green crosses) so the incentive to signal will be weaker. The cost and the incentive effects together lead to a reversal of the order of the $T-6$ equilibrium prices, causing belief monotonicity to fail at $T-7$.

It is important to note that while in all of the examples that we have considered so far the range of possible marginal costs has been small, it is possible to sustain separating equilibria for wider ranges, as long as the evolution of marginal costs is also changed to prevent signaling incentives from becoming too strong. This is illustrated in the final column of Table 2 where marginal costs have a range of \$0.50 and the standard deviation of the innovations is increased to \$0.25. In this case, the probability that a cost goes from one extreme of the support to the opposite half of the support is 0.32, similar to the baseline case. We can solve both the finite horizon and the infinite horizon models without conditions failing, and the increase in prices due to signaling, relative to the static incomplete information prices in period T, is slightly smaller than is observed in the baseline (as the elasticity of demand increases with the price level). In our empirical application we will estimate that the support of marginal costs is relatively wide.

[13] For example, an HL firm expects to face a low-cost LH firm (setting a black cross price) with probability 0.99, so the expected rival price is \$29.46.

[14] The crossing of the derivative functions reflects the failure of strategic complementarity (defined as $\frac{\partial^2 \pi_i}{\partial p_i \partial p_j} > 0$) for logit-based demand when prices are significantly above static profit-maximizing levels. The intuition is that, as a rival's price increases, the incentive for a firm to reduce its (high) price towards the static best response price can increase.

3.5 Signaling Incentives and Diversion

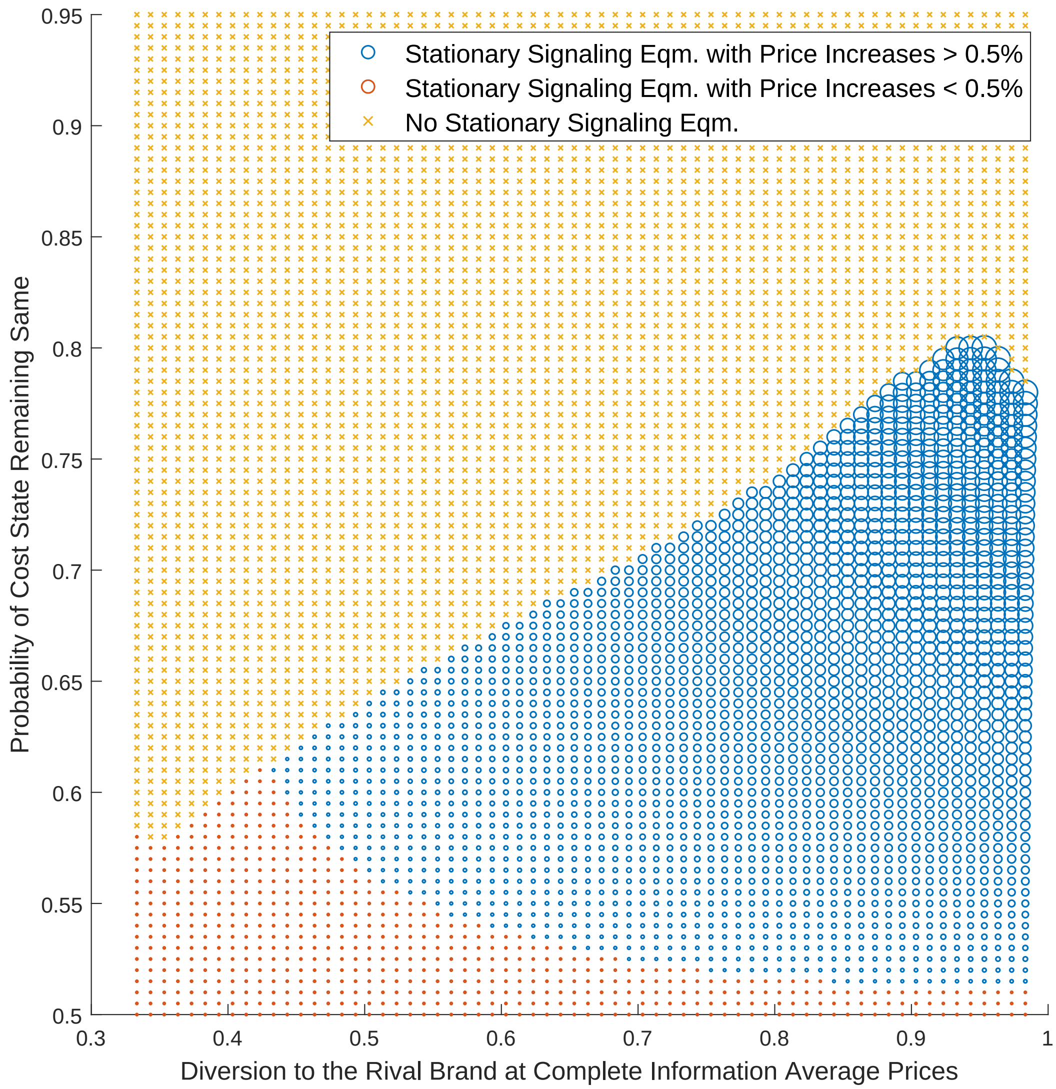
Figure 7 uses the infinite horizon version of the two-type duopoly model and a more flexible demand specification to illustrate how diversion and serial correlation affect both the existence of separating equilibria and how much signaling increases equilibrium prices (several alternative specifications produce qualitatively similar plots). As before, the two cost levels are 8 and 8.05. The generalized indirect utility function has the form $u_{i,c} = \beta - \alpha p_i + \sigma \nu_c + (1-\sigma)\epsilon_{i,c}$. We choose β, α and σ so that, for each combination of correlation and diversion that we consider, the complete information equilibrium prices (at average cost levels) are \$16 for each firm, the market share of each firm at these prices is 0.25, and the diversion, which measures the proportion of a product's demand that goes to the rival's product when its price increases from the complete information equilibrium price, has a value that we specify. Given assumed market shares, the lowest possible value of this diversion measure is $\frac{1}{3}$, which corresponds to multinomial logit demand. We also vary the (exogenous) probability that a firm's cost state remains the same as in the previous period from 0.5 (in which case there is no incentive to signal) to 0.99.

The orange crosses in Figure 7 indicate combinations where the conditions for characterizing best responses fail and we cannot find a separating equilibrium. For combinations where we can find a separating equilibrium the size and color of the circles indicates the percentage increase in average prices relative to average static Bayesian Nash equilibrium prices with the same demand and serial correlation parameters (these prices are always very close to \$16).

When serial correlation is very low, the price effects are always small whatever the level of diversion, and, for given diversion, the price effects become larger as serial correlation increases. For given serial correlation, higher diversion is associated with larger price effects, as it becomes more beneficial for a firm to increase its rival's price (because more of the demand that the rival loses will come to the firm), and the increase in a rival's price has a greater effect on the firm's best response. For moderate diversion, such as 0.6, an equilibrium cannot be sustained once serial correlation increases above 0.66. When diversion is very high, equilibria can be sustained with very large price effects: we find a maximum price increase of 44.8%.[15]

[15]In the diagram, the highest serial correlation for which we can find an equilibrium falls when we increase diversion above 0.95. This appears to reflect the fact that, at this level, small increases in diversion can increase signaling prices significantly, leading the conditions to fail. For each diversion value above 0.95 that we consider, we find maximum price effects of between 43.0% and 44.8%.

Figure 7: Equilibrium Average Price Increases in the Infinite Horizon Two-Type Duopoly Model as a Function of Diversion and Serial Correlation of Costs



Notes: red dots mark outcomes where there is a stationary separating equilibrium with average prices less than 0.5% above static BNE levels. The blue circles mark outcomes where there is a stationary separating equilibrium with larger average price increases relative to static BNE prices, and the size of the circle is linearly increasing in the percentage difference in prices (the largest effect shown has average prices increasing by 44.8%). Orange crosses mark outcomes where the conditions required to solve for best response functions fail. The diversion is measured by the proportion of demand that goes to the rival product when one product experiences a small increase in price at complete information Nash equilibrium prices given average costs.

3.6 Alternative Sources of Asymmetric Information

While it is plausible that, in many industries, firms have some private information about their marginal costs and that whatever is unobserved is likely to be serially correlated, other specifications can also generate significant price effects. We illustrate this point in Appendix B for three models where marginal costs are fixed and known but (i) rivals are uncertain about the serially-correlated weight that a firm's managers place on profits rather than revenues; (ii) rivals are uncertain about the serially-correlated weight that a firm's managers place on the firm's profits rather than the profits of other firms in the industry; or, (iii) rivals are uncertain about some serially-correlated element of a firm's demand that the firm's managers know. In each case we show that equilibrium prices can be significantly higher, and more volatile, than in the complete information or static incomplete information versions of the model.

4 Additional Firms and the Effects of Mergers

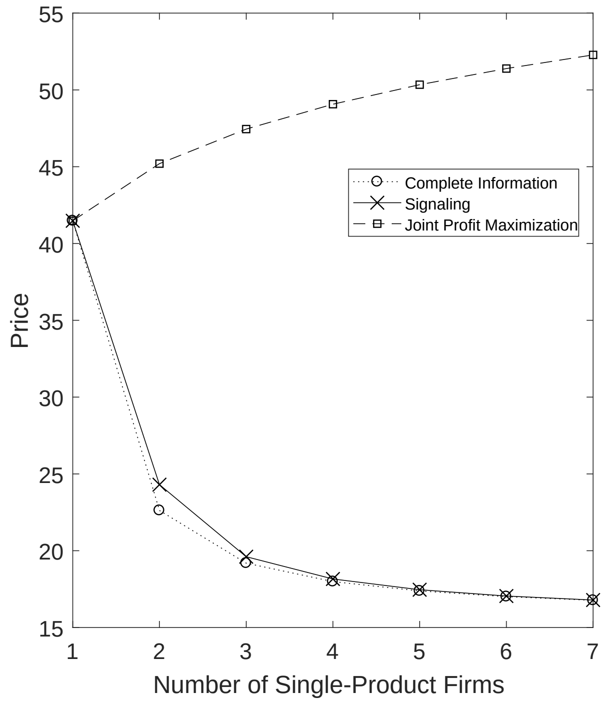
In our model signaling is a strategic investment to raise rivals' future prices. When there are more firms, price increases tend to be more costly because residual demand is more elastic and, all else equal, each firm will have less effect on rivals' pricing. In this section we illustrate the relationship between the number of firms and the difference between signaling and complete information prices, and we use a stylized example to illustrate what this relationship implies about the analysis of mergers. This analysis provides motivation for our empirical application where we consider whether our model can explain unexpected price increases after the MC JV.

4.1 Price Levels and the Number of Firms

Figure 8 shows average equilibrium prices under complete information and in our model in an infinite horizon game with two possible levels of marginal costs when the number of ex-ante symmetric single-product firms varies from one to seven. The demand specification is the same as used in our baseline example. Marginal costs are either 8 or 8.05, with a probability of $\frac{2}{3}$ of remaining at the same level in the next period.[16] Under duopoly, average signaling prices are

[16] We present results for the model with two types as only this model has a reasonable computational burden with five, six or seven firms. Results from a continuous cost model with comparable parameters gives similar price effects to the model used here for two, three or four firms.

Figure 8: Equilibrium Effects on Average Prices in the Infinite Horizon Two-Type Model with Different Numbers of Symmetric, Single Product Firms



7.4% higher than complete information prices in this example, with 2.2% and 0.9% differences with three and four firms. With seven firms the difference is just 0.1%. Under monopoly the prices are the same, as there is no role for signaling.

The figure also plots average prices when firms maximize their joint profits under complete information. These prices tend to increase with more firms because the number of products is increasing. While signaling prices are significantly above complete information prices, they are far below the prices that would be set if firms colluded to maximize their joint profits.

4.2 Merger Analysis

We now report the effect of mergers within our model, and compare them to predictions that an analyst would make using a complete information merger simulation model. We perform the analysis using the infinite horizon, continuous cost model with our baseline demand and marginal cost parameters, although we allow for more firms. This implicitly assumes that the merger considered is an unanticipated shock and the firms do not anticipate future mergers. We allow for post-merger asymmetries between the firms due to the merged firm owning more than one product and/or merger-specific synergies that lower marginal costs. Following the *Horizontal*

Merger Guidelines, a merger that significantly increases concentration is unlikely to be challenged if synergies are expected to be large enough to prevent prices from increasing.[17]

The first column of the upper section of Table 3 considers a merger that reduces the number of firms from 4 to 3 and eliminates a product (for example, the American Airlines/US Airways merger eliminated US Airways service from Washington National to Boston). If we assume that the firms play a signaling equilibrium before the merger and after the merger, then a 4-to-3 merger without a synergy will increase the merged firm's average price by 8.5%. As, under the no synergy assumption, the firms remain symmetric post-merger, the other firms' average prices increase by the same amount as their incentives to signal increase in the same way. To prevent its average price from rising, the merger would need to reduce the average marginal cost of the merging firm from \$8.03 to \$5.76, a 29% reduction, which is much larger than would usually be considered plausible.[18]

The next rows consider what an analyst would expect to happen if he assumes that firms set prices in a complete information Nash equilibrium. Using complete information first-order conditions and an accurate estimate of the demand system, the analyst would infer that average pre-merger marginal costs are equal to \$8.29 (i.e., that they are higher than they really are). A complete information merger simulation model would also imply that the merged firm's average marginal cost would need to fall to \$7.11 to prevent prices from rising. The required reduction in marginal costs is smaller than in the first part of the table because the analyst ignores that signaling incentives will be strengthened by the merger. If the merged firm's post-merger average marginal cost does fall to \$7.11, but firms use equilibrium signaling strategies after the merger, then average prices increase to \$19.17 (a 5% post-merger increase).

The second column performs the same analysis for a 3-to-2 merger. All of the effects identified in the first column become larger and, in the signaling equilibrium, the merged firm's marginal cost would need to be *negative* to prevent its average price from increasing.[19] The realization of the synergy identified by a complete information analysis would not prevent prices rising by 17%.

[17]In this example, we focus on the merged firm's average price rather than the expected level of consumer surplus.

[18]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.

[19]We assume that the firm cannot freely dispose of products rather than selling them, so that it will not choose to produce an infinite amount if it has negative marginal costs.

Table 3: Examples Illustrating the Effects of Mergers and the Effects on 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 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 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%)

The lower part of the table considers the case where after the merger, the merged firm has two products and so it engages in multi-product pricing, which increases the incentive to raise prices. This is the more standard case for mergers in differentiated product markets. We assume that both of the merged firm's products will have identical marginal costs in every period after the merger (evolving by the same process assumed previously) and that the merged firm must set the same prices for both of them. These assumptions maintain the tractable structure of a single type and single signal per firm. If no synergies are realized, then prices rise more than in the case when a product is dropped, under both signaling and complete information. The synergies required to prevent price increases under complete information become significantly larger, which slightly reduces the predicted price increase that occurs if this synergy is realized. However, for our parameters, mergers to triopoly or duopoly must still generate exceptionally large, and implausible, marginal cost synergies if prices are not to rise.

5 Empirical Application: The MillerCoors Joint Venture

We now turn to our empirical application. We estimate our model in order to examine if it can explain price increases following the MC JV. The JV was an effective merger of the US brewing, marketing and sales operations of the second and third largest US brewers. The Department of Justice decided not to challenge the JV in June 2008, after an eight month investigation, on the grounds that "large reductions in variable costs of the type that are likely to have a beneficial effect on prices" were expected.[20] Ashenfelter, Hosken, and Weinberg (2015) provide cross-market evidence that these efficiencies were indeed realized. However, MW show that, on average, the prices of leading MillerCoors brands, and also the main AB brands, which were not directly affected by the JV, increased after the JV, when compared to the prices of imported brands, such as Corona and Heineken, that they view as controls for industry-wide cost changes. A complete information Nash Bertrand model cannot easily rationalize why AB should increase its markups as much as MC. This feature has led MW and Miller, Sheu, and Weinberg (2019) to suggest that the JV initiated tacitly collusive behavior by AB and MC. In contrast, we analyze whether the data can be explained by a theory where the major firms engage in price signaling both before and after the JV.

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

We focus on the prices of Bud Light (BL), Miller Lite (ML) and Coors Light (CL), the leading light beer brands. The light beer segment is particularly dominated by domestic brewers (Miller, Coors and AB sold 90% of volume in 2007), and the estimates of Hausman, Leonard, and Zona (1994) and Toro-González, McCluskey, and Mittelhammer (2014) indicate that there is only limited substitution between light beers and other segments. However, as we report predictions for a range of demand parameters[21], and the prices of domestic full calorie brands, such as Budweiser, move in parallel with those of the leading light beer brands, our results can also be interpreted as applying to the broader beer market as long as there is limited substitution from the main domestic brands (75% of the whole market) to craft and imported brands, which are usually significantly more expensive and are sold in smaller packs. After briefly discussing the motivating changes in prices and market shares, we describe how we estimate the cost parameters by matching pre-JV price dynamics, before presenting estimates and counterfactuals.

5.1 Data

Following MW, we use scanner data from the IRI Academic Dataset (Bronnenberg, Kruger, and Mela (2008)). This dataset provides weekly UPC-level data on revenues and units sold from a large sample of retail stores. We use data from grocery stores in 41 IRI-defined geographic markets. We do not observe wholesale prices and we will instead treat retail prices as if brewers choose them directly[22], although, as we will describe below, we will exclude temporary store price promotions (sales) when estimating the model.

Table 4 shows the ownership, market shares (by volume) and average prices (total revenues divided by the volume of 12-pack equivalent cases) of the fifteen highest-selling light beer brands in 2007. Foreign and craft light beer brands have low market shares, both individually and combined. The fact that these brands also have much higher prices and are sold almost exclusively in smaller package sizes suggests that they may be poor substitutes to the main light beer brands

[21] A more standard approach would be to estimate a single preferred demand system and perform counterfactuals using this system. However, reported elasticities and diversion in the beer industry vary widely (for example, Hellerstein (2008) and Hausman, Leonard, and Zona (1994) report average brand own-price elasticities of around -6 and -5 respectively, whereas Rojas and Peterson (2008) finds estimates of around -1.3. Asker (2016) reports an average elasticity of -3.4. Our own estimation has also found that implied elasticities and diversion are sensitive to the specification, the aggregation of pack sizes and the choice of instruments, so we prefer our approach as being more transparent and robust.

[22] This is a common approach in the literature. MW consider a variant of their model where local monopolist retailers choose mark-ups, but they cannot reject the model where wholesale price changes are perfectly passed through.

Table 4: Leading US Light Beer Brands in 2007

Brand	Owner (2007)	Volume Market Share	Average Price Per 12 Pack Equiv.	Size/Package Forms	% Vol. Sold in Large Packs
Bud Light	Anheuser-Busch	31.5%	$8.36	9	71.4%
Miller Lite	SABMiller	20.5%	$8.12	9	74.9%
Coors Light	Molson Coors	16.4%	$8.40	9	74.6%
Natural Light	Anheuser-Busch	7.6%	$6.04	7	70.9%
Busch Light	Anheuser-Busch	6.0%	$6.06	9	80.1%
Keystone Light	Molson Coors	2.9%	$5.91	6	82.6%
Milwaukee's Best Light	SABMiller	2.6%	$5.36	6	68.1%
Corona Light	Grupo Modelo	2.5%	$14.17	3	2.5%
Miller High Life Light	SABMiller	1.5%	$5.97	9	64.6%
Michelob Light	Anheuser-Busch	1.3%	$9.72	7	29.4%
Heineken Premium Light	Heineken	1.2%	$14.17	5	2.2%
Labatt Blue Light	Labatt	1.1%	$7.47	5	78.6%
Miller Genuine Draft Light	SABMiller	0.9%	$7.72	9	68.2%
Amstel Light	Heineken	0.8%	$14.29	5	0.8%
Samuel Adams Light	Boston Beer Co	0.5%	$13.63	3	0.0%
	Combined Share	97.3%			

Notes: Statistics based on grocery stores in the IRI sample in 2007 and package sizes equivalent to 6, 12, 18, 24 and 30 packs of 12 oz. bottles or cans. Average retail price is calculated using total retail revenues divided by the number of 12-pack equivalent units sold. "Forms" counts the number of combinations of packaging (bottle or can) and package size observed in the sample (for the sizes listed above). The statistic in the final column is the proportion of total volume that is sold in package sizes containing at least 18 12oz. bottles or cans.

produced by Miller, Coors and AB.

5.1.1 Changes in Prices and Market Shares after the MillerCoors Joint Venture

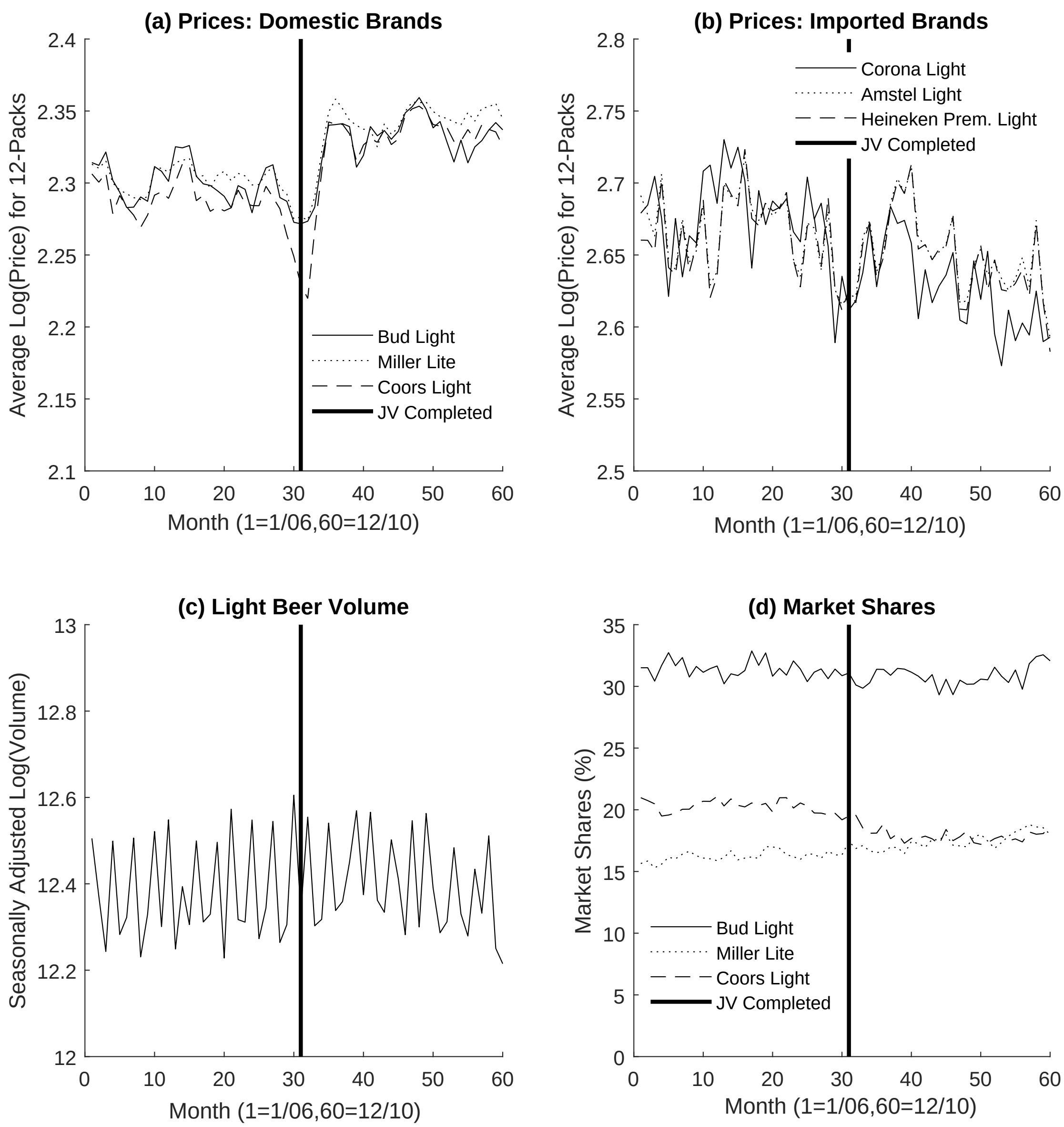
Figure 9(a) and (b) show the time-paths of monthly average log real (deflated to January 2010 prices using the CPI-U series from the BLS) retail prices for 12-packs (bottles or cans) for the three leading domestic light beers and three imported light beers. The real prices of the domestic brands rose after the completion of the JV, and remain relatively high through to the end of 2010. In contrast, the real prices of the imported brands declined. However, the jump in real prices during the second half of 2008, while striking, should not be interpreted as clear evidence of a large merger effect on prices because, just after the JV was competed, the CPI-U deflator fell from 220.0 in July 2008 to 210.2 in December 2008. Instead, it is the comparison between the domestic and imported prices in the years following the JV period which is informative.

Table 5 reports regression estimates of the post-JV price increase for domestic brands. Brand-store-week deflated prices (measured as the average price per 12-pack equivalent in specifications where all package sizes are used) are regressed on dummies that interact the main domestic brands with a dummy for the period after the JV was completed, store-brand fixed effects and week fixed effects. Standard errors are clustered on the geographic market.

The sample in the first four columns consists of the BL, ML, CL, Corona Light, Heineken Premium Light and Amstel Light, so that the reported coefficients are measured by price changes for domestic brands relative to these imported brands. The specification therefore follows the identification strategy of MW. Common time effects are controlled for using week dummies, and prices are calculated in either levels or logs, using either all package sizes or only 12-packs. Column (4) excludes any store-brand observations where there is a temporary price reduction on any package size. The coefficients on the post-JV brand dummies indicate that, after the JV, the prices of BL, ML and CL increased by 4-6% or 70 cents to one dollar per 12-pack.

Columns (5) and (6) include brand-specific linear time trends, to capture the fact that the real prices of the imported brands appear to be falling slightly more quickly than the real prices of the domestic brands prior to the JV. Controlling for these trends reduces the estimated post-JV price increases for domestic brands to 3-4% or 25-30 cents per 12-pack equivalent. It is also noticeable that the pattern of the earlier columns that BL prices increase slightly less than the prices of ML and CL disappears once time trends are controlled for. The remaining columns,

Figure 9: Light Beer Prices and Volumes/Market Shares Between 2006 and 2010



Notes: the total volume figure is drawn using the set of stores that were in the sample for at least 250 weeks from January 2006 to December 2010, the period shown. Seasonal adjustment is implemented by regressing the log monthly volumes on month-of-year dummies and then removing the estimated month-of-year effects from the series. Market shares are calculated using all stores in the sample.

Table 5: Regression Estimates of Post-Joint Venture Price Changes for the Leading Domestic Brands

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Pack Sizes in Sample	All	All	12-Packs	All	All	All	All	All
Units of Price Variable	Log	Dollar	Dollar	Dollar	Log	Dollar	Dollar	Dollar
Price Reductions	incl.	incl.	incl.	excl.	incl.	incl.	incl.	incl.
Post-JV*								
Anheuser-Busch/BL	0.0444	0.7086	0.9279	0.7555	0.0358	0.2930	0.8186	0.3512
	(0.0051)	(0.0557)	(0.0862)	(0.0803)	(0.0034)	(0.0441)	(0.0573)	(0.0485)
SAB Miller/ML	0.0595	0.8504	1.0383	0.9157	0.0373	0.3011	0.9948	0.3360
	(0.0063)	(0.0621)	(0.0749)	(0.0766)	(0.0045)	(0.0530)	(0.0630)	(0.0567)
MolsonCoors/CL	0.0499	0.7603	1.0336	0.8377	0.0324	0.2531	0.8379	0.2602
	(0.0047)	(0.0523)	(0.0805)	(0.0743)	(0.0040)	(0.0522)	(0.0564)	(0.0597)
Fixed Effects	Store-Brand Week	Store-Brand Week	Store-Brand Week	Store-Brand Week	Store-Brand Week	Store-Brand Week	Store-Brand Week	Store-Brand Week
Brand Time Trends	N	N	N	N	Y	Y	N	Y
Domestic Brands	BL,ML,CL	BL,ML,CL	BL,ML,CL	BL,ML,CL	BL,ML,CL	BL,ML,CL	All light brands of AB, M and C	All light brands of AB, M and C
Observations	2,190,468	2,190,468	2,065,349	817,071	2,190,468	2,190,468	4,749,181	4,749,181
R-squared	0.9152	0.9071	0.8235	0.9386	0.9153	0.9078	0.9367	0.9377

Notes: observations are store-week-brand observations from 2003 to 2011 on either deflated prices of bottles or cans sold in packages containing 144 oz. or average prices when total revenues are divided by the total volume sold in 144 oz-pack equivalents. All regressions include Corona Light, Heineken Premium Light and Amstel Light, as well as the listed selection of domestic brands (BL = Bud Light, ML = Miller Lite, CL=Coors Light, AB=Anheuser-Busch, M=Miller and C=Coors). Standard errors are clustered on the IRI geographic market. In column (4) observations are excluded where a store-brand is indicated as having a price reduction on any pack size.

with and without brand time trend controls, include observations for other light beer brands produced by the domestic brewers (e.g., Busch Light), so that the coefficients measure average changes in the prices of each domestic firm's light beer portfolio relative to the imported brands. The estimated price increases are slightly larger in magnitude than the equivalent estimates using only the largest brands. Our take-away from the table is that we should be interested in whether our model predicts post-JV price increases in the range of 30-90 cents per 12-pack equivalent given reasonable assumptions on demand.[23]

To help to understand substitution patterns, Figure 9(c) and (d) show how the seasonally adjusted total volume of light beer sold and the market shares of the leading brands vary over time. There is no clear effect of post-JV price increases on either measure, although the market shares of CL (ML) increased (fell) consistent with MC marketing CL more aggressively in markets where transportation costs fell (Chandra and Weinberg (2018)). Correspondingly, the combined market shares of the three imported brands considered previously do not increase (these three brands accounted for 4.1% of light beer volume in 2006 and 4.0% in 2010) even though their relative prices are falling. These patterns suggest that there is at most limited substitution to either the outside good of not purchasing or imported/craft brands when the prices of domestic brands rise.[24]

5.2 Empirical Specification and Estimation

We estimate an infinite horizon, continuous marginal cost three-firm version of our model using pre-JV data. Implicitly we are assuming that firms did not anticipate the JV and that smaller brewers, whose products are included in the outside good, did not respond to signaling behavior by the incumbents.[25] We allow for demand and cost asymmetries between BL and ML and CL, while treating the latter two brands as being symmetric with each other, as we will impose, in

[23]We have also estimated specifications where observations from between the announcement and the completion of the JV are excluded. The coefficients are very similar to those reported in the table.

[24]The survey of Fogarty (2010) reports that aggregate industry demand is very inelastic, indicating that there is little diversion to the outside good when prices rise. On the other hand, the data shows that brand demand is elastic. For example, when log(sales)-log(price) regressions are used to estimate brand own-price elasticities, using a variety of different aggregation and control/instrument strategies, the estimated elasticities for the major brands are usually between 2.5 and 3.5, including for the leading light beer brands. These values are in the middle of the wide range estimated in the prior literature (see footnote 21).

[25]To the extent that other firms would increase their prices in response to signaling or signal themselves, this would tend to generate larger price increases. An earlier version of the paper did include Heineken and Corona light beer brands in a counterfactual, and found that their signaling incentives raised their prices by no more than two cents.

our counterfactual, that MC has to charge the same price for both products.[26]

We use an indirect inference estimation procedure (Smith (2008)). For given parameters, we solve our signaling model, simulate data and calculate moments of the distribution of prices and estimate AR(1) regressions for the price of each brand. The moments and the coefficients of the AR(1) regressions are matched to their empirical counterparts for each market, where we treat each market as a separate observation on a game with common parameters across markets.[27] A computationally-light two-step approach to estimation (as pioneered by Bajari, Benkard, and Levin (2007) for dynamic games with continuous choices) is not feasible in our setting because firm strategies are functions of unobserved beliefs.

5.2.1 Demand

We perform our estimation and counterfactuals for thirty different sets of assumptions about average diversion and own-price elasticities. Demand for the three named brands is determined by a nested logit demand model where all of the brands are in a single inside nest. The choice of not purchasing or choosing other brands (and here we have in mind, in particular, brands that are not controlled by Miller, Coors or AB) are subsumed into the outside good, which is in its own nest. We choose the parameters (nesting parameter, price coefficient and qualities for BL and ML/CL) so that, at average observed pre-JV prices, our demand system matches pre-specified values of the average own price elasticity and average (across brands) diversion to the outside good, and the average market share of each product.[28] For example, when we assume an average own price elasticity of -3 and that, when the price of one of the brands increases that, on average, 20% of lost demand goes to the outside good (equivalently 80% of lost demand goes to the two other brands), the selected price coefficient is -0.124, the nesting parameter is 0.689 and the mean qualities (i.e., mean utilities excluding the effect of price) are 1.317 and 1.079 for

[26]It is, of course, the case that observed retail prices for ML and CL can be different after the JV. However, their retail prices do become more correlated. For example, examining the brand-market-week average price of ML and CL 12-packs, the serial correlation is 0.89 before the announcement of the JV and 0.96 after its completion. Measures of within-market correlation of ML and CL prices also increase substantially (for example, a cross-market average of 0.54 to 0.74).

[27]In principle, we could estimate the model allowing for differences in costs and consumer preferences across markets, but this would scale the computational burden of estimation by the number of markets in the sample. Estimation takes around 12 hours for a given set of demand parameters, so it would likely take around 20 days if we used market-specific parameters.

[28]We use unweighted average market shares across the geographic markets and rescale the market shares to allow for some presence of not purchasing. This leads to us matching market shares of 28% for BL and 13.5% for the two other brands.

BL and ML/CL respectively.

5.2.2 Cost Parameters

We use our model to estimate five cost parameters: $\underline{c^{BL}}$ (the lower bound on costs for BL), $\underline{c^{ML/CL}}$, the length of the cost interval (assumed the same for all brands), the AR(1) serial correlation parameter for the cost process, and the standard deviation of the (non-truncated) marginal cost innovations. This means that the marginal cost innovation process is the same for all three firms, with AB allowed to have a different cost level. We could allow for additional differences across firms, but this raises the computational burden while only marginally improving the fit of the model. We initially estimate a version of the model where firms use static Bayesian Nash Equilibrium strategies, i.e., they do not signal, to find starting values.

5.2.3 Objective Function

We match the predictions of our model to features of prices taken from the retail data in the IRI database from January 2003 up to the announcement of the JV in October 2007. The estimates reported here use weekly market average deflated prices for each pack size for each of the three brands, with all prices converted to 12-pack equivalents. We only use market-week observations where at least five stores are observed so that changes in average prices may be a reasonable indicator of what is happening to market wholesale prices.[29] As temporary retail sales may provide a separate source of price volatility to the one that we are modeling we exclude store-pack size-week observations where IRI indicates that the store had a temporary price reduction (sale). We construct the following values from this data which will be matched during estimation:

- the average deflated price for each brand for each market, averaging the 12-pack equivalent prices across pack sizes;[30]
- measures of the interquartile range of prices. For each brand, we regress the weekly market average price for the brand-pack size on dummies for the specific set of stores observed in each market (interacted with pack size) and dummies for the week (also interacted with

[29] The five store criterion reduces the number of markets to 38.

[30] We use the unweighted average price across pack sizes, so that the average prices that we match (see Table 6) are higher than the total revenue divided by total volume price reported in Table 4. We find similar estimates if we match moments based on the average prices of 12-packs alone.

pack size), to control for fixed differences between retail stores and national promotions. The interquartile range of the residuals for each market for BL and the average for ML and CL are our dispersion measures; and,

- the coefficients from AR(1) regressions of prices where, for each market and brand separately, the market average price of the brand-pack size is regressed on the equivalent lagged price of each brand, a linear time trend and dummies for the specific set of observed stores (interacted with pack size). We collect the lagged price coefficients, and also calculate the standard deviation of the price innovations (residuals) from the regression.

Appendix C reports the coefficients from national level versions of the AR(1) regressions (i.e., specifications where the coefficients are the same across markets), and compares the estimates based on our chosen selection of data to alternatives where we aggregate data to the market-month level (rather than week) and we use data only on 12-packs (rather than all pack sizes). The patterns are qualitatively similar using these alternatives.

For a given value of the cost parameters, we solve the model[31] and simulate a long time-series of data from the solved model and calculate the average price, the interquartile range of prices, and the lagged price coefficients and the standard deviation of the innovations when we estimate a similar AR(1) regression.[32] The Wald-type indirect inference objective function is formed by measuring the difference between the relevant averages, ranges and coefficients. The specific form of the estimator 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 $\widehat{\tau_{k,m}^{data}}$ is one of the auxiliary model coefficients or averages estimated using the data and $\widehat{\tau_k(\theta)}$ is the equivalent coefficient estimated using simulated data from the model solved using parameters θ. W is a weighting matrix. The reported results use an identity weighting matrix.[33] The objective

[31]As described in Appendix A.2.1, we speed computation by solving for pricing functions for a relatively coarse subset of beliefs about prior marginal costs and interpolating them for other values so that we can calculate expected profits accurately. This exploits the feature seen in Figure 4 that when a firm's belief about a rival's prior marginal cost rises, its equilibrium function is very close to being simply translated upwards.

[32]Of course, our model does not have different pack sizes, varying sets of stores or time trends, so the regression using simulated data do not control for these factors.

[33]We have also estimated the model for a subset of the alternative demand parameters using an estimated inverse covariance matrix of the various moments as the weighting matrix and using a diagonal weighting matrix where the values on the diagonal are proportional to the inverse variance of each moment. The latter approach

function is minimized using `fminsearch` in MATLAB (version 2018a).

An issue with estimating a game using a nested fixed point approach, like ours, is that we have not shown that the equilibrium in our continuous cost, infinite horizon model is unique. This issue leads to two possible concerns. The first concern is that an algorithm that tries to minimize the objective function may not work well if a small change in the parameters causes the solution algorithm to jump between different equilibria that generate different predictions. This is not a problem in practice, as for all of the demand parameters that we consider, `fminsearch` is able to find parameters where the moments are matched almost exactly, and, when we calculate derivatives of the moments to calculate standard errors, we find that the predictions of the solved model change smoothly with the parameters.

The second concern is that if a model has multiple equilibria then it may be that equilibria supported by different parameters could produce very similar predictions so that the parameters are not econometrically identified even if our algorithm produces coefficient estimates with small standard errors because our algorithm selects a particular equilibrium (for example, the equilibrium with the highest prices). Here we have to rely on what we reported in Section 2.3: we have not found multiple equilibria in the continuous type model, and, in the two-type model, where we have identified some multiplicity in the infinite horizon model, our algorithm picks out strategies that are the limits, as the number of periods grows, of what appear to be the unique equilibrium strategies in the early periods of a finite horizon game.[34]

If we assume that equilibria are unique, then the heuristic intuition for identification is fairly straightforward. Given the assumed demand parameters, the observed price levels and the markups implied by the model will identify the lower bounds on brand marginal costs. The estimated AR(1) coefficients will identify the innovation process for costs, while the observed dispersion of prices, as measured by the interquartile range, should identify the range of costs given the cost process. Of course, when we estimate the parameters we are imposing the full equilibrium structure of the model, and the value of any moment can affect all parameters. This is illustrated in Appendix D where we report the sensitivity of the coefficient estimates to the twelve moments

produces similar estimates and fit to the results using that identity matrix that are reported below. Estimates and standard errors using the full covariance matrix are more sensitive to the exact selection of moments being matched, possibly because the estimates of covariances are noisy given the small number of markets.

[34] We recognize that this argument could be refuted by an alternative computational approach. An interesting direction for future work would be to try to establish the existence, or not, of multiplicity in a more systematic way or the identification of criteria that would allow a particular equilibrium to be selected.

using the parameter sensitivity measure suggested by Andrews, Gentzkow, and Shapiro (2017).

5.3 Parameter Estimates and Model Fit

The upper section of Table 6 reports the estimated cost parameters for three different combinations of average own-price elasticities and diversion values. Less elastic demand implies higher margins (in both our model and more standard models) and lower diversion to the outside good implies stronger signaling incentives, so that estimated marginal costs are lowest in the second column. BL is estimated to have a significantly lower marginal cost than ML and CL, to rationalize why the observed prices of the three brands are roughly equal. The cost intervals and the variance of the innovations in marginal cost are both larger than in our example.[35]

The middle section of the table reports the fit of the model. All of the moments/coefficients that are used in estimation are matched quite accurately, although in trying to match both the dispersion in prices and the standard deviation of week-to-week price changes, the latter tend to be underpredicted. The table also reports the skewness of the distributions of the innovations in prices, which we do not try to match in estimation. The innovations tend to be left (negative) skewed for both the data and the model. Given that temporary store price reductions, which could also generate skewness, have been removed from the data price series, the fact that we also match this feature of the data lends some additional support to the model.

5.4 Counterfactuals

We now report what our model predicts when we simulate the effect of the JV. We re-solve the model assuming that ML and CL are produced by a single firm that always has the same marginal cost for each of these brands, and sets the same price for both brands in every period (see footnote 26 describes how the prices of these brands became more correlated after the JV). The width of the cost interval and the innovation process are assumed to be the same as before the joint venture for both AB and MC. We assume that MC benefits from the synergy that would have prevented average prices from rising in a complete information model with the

[35]The estimates in column (1) imply that the probability that a firm with the highest marginal cost will have a marginal cost in the lower half of the cost interval in the next period is 0.2, which compares to 0.32 in our baseline example in Section 3.

Table 6: Cost Parameter Estimates and Model Fit for Three Different Demand Parameterizations

	(1)	(2)	(3)
Price Elasticity	-3	-3	-3.5
Mean Diversion to Outside Good	0.25	0.15	0.15
Implied Demand Coefficients (not estimated)			
Nesting Parameter	0.603	0.770	0.770
Price Coefficient	-0.159	-0.095	-0.111
Estimated Cost Parameters			
$\underline{c}^{BL}$	\$5.552	\$5.290	\$6.007
	(0.005)	(0.039)	(0.012)
$\underline{c}^{ML/CL}$	\$6.547	\$6.470	\$6.985
	(0.004)	(0.030)	(0.013)
Cost Interval Length ($\overline{c^i} - \underline{c^i}$)	\$0.799	\$0.890	\$0.784
	(0.058)	(0.055)	(0.012)
Cost Serial Correlation, ρ	0.802	0.609	0.899
	(0.058)	(0.110)	(0.005)
Std. Dev. of cost innovations, σ_c	\$0.263	\$0.238	\$0.280
	(0.006)	(0.039)	(0.004)
Value of Matched Moments/Coefficients			
Mean Price: BL (data \$10.38)	\$10.39	\$10.39	\$10.39
Mean Price: ML (\$10.28)	\$10.25	\$10.26	\$10.25
Mean Price:CL (\$10.26)	\$10.27	\$10.27	\$10.27
IQR of Prices BL (\$0.22)	\$0.22	\$0.21	\$0.21
...... Average for ML and CL (\$0.25)	\$0.30	\$0.30	\$0.31
AR(1) Regression Coefficients			
$\rho^{BL,BL}$(0.459)	0.430	0.425	0.428
Average of $\rho^{BL,ML}$ and $\rho^{BL,CL}$ (0.057)	0.059	0.073	0.075
...... $\rho^{ML,ML}$ and $\rho^{CL,CL}$ (0.421)	0.448	0.450	0.452
...... $\rho^{ML,BL}$ and $\rho^{CL,BL}$ (0.041)	0.032	0.035	0.039
...... $\rho^{ML,CL}$ and $\rho^{CL,ML}$ (0.052)	0.030	0.035	0.035
SD of innovations			
...... from BL AR(1) regression (0.193)	0.124	0.123	0.121
...... from ML & CL AR(1) regressions (0.212)	0.171	0.179	0.175
Skewness of Innovations (unmatched)			
BL AR(1) regression (-0.423)	-0.269	-0.263	-0.271
..... Average for ML & CL (-0.282)	-0.272	-0.255	-0.251
Complete Information ML/CL Synergy to Prevent Price Increases			
Marginal Cost Reduction	-\$1.06	-\$1.22	-\$1.04

Notes: BL = Bud Light, ML = Miller Lite and CL=Coors Light. Moments for the data reported in parentheses in the left column are averages across 38 markets, and are based on market average-brand-week retail prices of all package sizes in 12 pack equivalents, excluding temporary store price reductions, for market-weeks where at least five stores are observed in the IRI data. The assumed form of the cost transition is $c_{i,t} = \rho c_{i,t-1} + (1-\rho)\frac{\overline{c}+\underline{c}}{2} + \eta_{i,t}$ where $\eta \sim TRN(0, \sigma_c^2, \underline{c} - \rho c_{i,t-1} - (1-\rho)\frac{\overline{c}+\underline{c}}{2}, \overline{c} - \rho c_{i,t-1} - (1-\rho)\frac{\overline{c}+\underline{c}}{2})$ where TRN denotes a truncated normal distribution. The estimates of ρ and σ_c are reported in the table. "....." indicates a measure of the same type as the one reported in the previous row. Standard errors in parentheses for the coefficient estimates.

same parameters. Examples of these synergies are reported in the last row of Table 6.[36] These assumptions are reasonable in our context as MC produced ML and CL in the same breweries, and the Department of Justice expected that the resulting synergies would prevent prices from rising. MW suggest that the expectation would have been correct if the firms engaged in static Bertrand Nash pricing before and after the JV.

The results are reported in Table 7 for the thirty demand parameterizations that we consider. When demand is least elastic and there is more diversion to the outside good, the conditions required to characterize best responses fail when we try to solve the post-JV model. This is consistent with our investigation of when the conditions fail for the two-type duopoly model, although we are able to estimate the parameters using the three-firm model even for these cases. For the other twenty-eight cases we predict non-trivial average price increases for both of the MC brands and BL, with slightly larger increases for BL because our demand parameters imply BL has a slightly greater market share than ML and CL at pre-merger prices (see footnote 28), which tends to give BL a greater signaling incentive. The increases are larger with less elastic demand or less diversion to the outside good.

Figure 10 shows the equilibrium pricing strategies for BL when $c_{t-1}^{BL} = \underline{c^{BL}}$ and $c_{t-1}^{ML} = c_{t-1}^{CL} = \underline{c^{ML/CL}}$, and $c_{t-1}^{BL} = \overline{c^{BL}}$ and $c_{t-1}^{ML} = c_{t-1}^{CL} = \overline{c^{ML/CL}}$ for the static Bayesian Nash 3-firm model, the estimated signaling 3-firm model and the counterfactual when the elasticity is -3 and diversion to the outside good is 0.15. When all firms' past costs and BL's present cost are at their lowest possible levels, the BL price is similar in the three models, but, consistent with our earlier results, the range of prices that may be charged when any of the firms have higher costs increases when we allow signaling, especially in the counterfactual.

The estimates and the counterfactuals assume a discount factor of 0.99. This may be regarded as too low for weekly price changes. We have also re-performed the estimation and counterfactual computations for a weekly discount factor of 0.998 (implying an annual discount factor of 0.9) for a subset of the demand parameters. As more patience leads to larger price effects and more volatility in the three-firm model, the parameters change in such a way that the effect on the counterfactual predictions is very small. For example, with an elasticity of -3 and diversion ratio

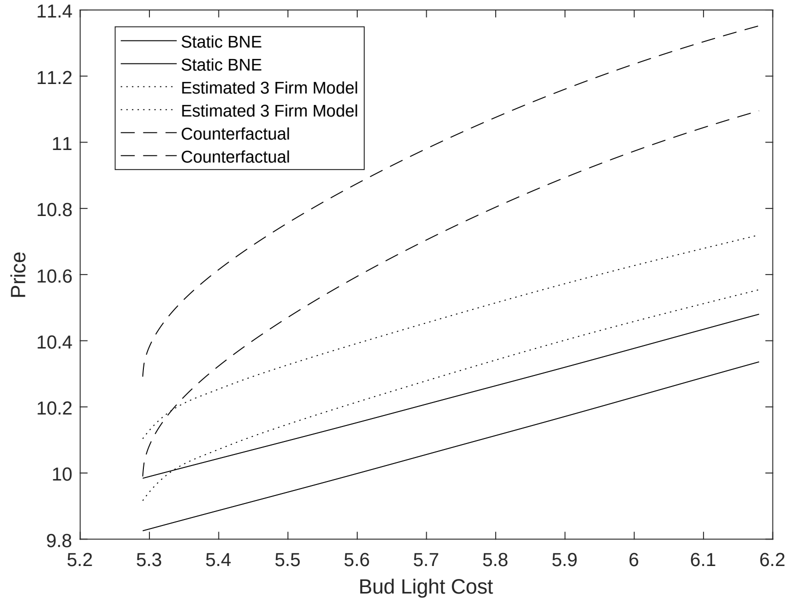
[36]The required synergy is larger when the brands are closer substitutes (lower diversion to the outside good) and, in each case, the synergy would make MC's marginal costs very similar to, and slightly lower than, those of BL. This reflects the fact that, with no price increases, ML and CL had a similar combined market share to BL but were sold at slightly lower average prices in the pre-JV data.

Table 7: Counterfactual Predictions of Average Prices After the MillerCoors Joint Venture

Mean Diversion to the Outside Good	0.3		0.25		0.2		0.15		0.1		0.05	
	BL	ML/CL	BL	ML/CL	BL	ML/CL	BL	ML/CL	BL	ML/CL	BL	ML/CL
Pre-JV Price	$10.38	$10.27	$10.38	$10.27	$10.38	$10.27	$10.38	$10.27	$10.38	$10.27	$10.38	$10.27
Elasticity												
-2	Conditions		Conditions		$11.05	$10.91	$11.30	$11.15	$11.50	$11.35	$12.21	$12.05
	Fail		Fail		+$0.67	+$0.64	+$0.92	+$0.88	+$1.12	+$1.08	+$1.83	+$1.78
					+6.5%	+6.2%	+8.9%	+8.6%	+10.8%	+10.5%	+17.6%	+17.3%
-2.5	$10.63	$10.49	$10.74	$10.60	$10.90	$10.77	$11.12	$10.98	$11.39	$11.22	$11.66	$11.49
	+$0.25	+$0.22	+$0.36	+$0.33	+$0.52	+$0.50	+$0.74	+$0.71	+$1.01	+$0.95	+$1.28	+$1.22
	+2.4%	+2.1%	+3.5%	+3.2%	+5.0%	+4.9%	+7.1%	+6.9%	+9.7%	+9.3%	+12.3%	+11.9%
-3	$10.59	$10.46	$10.70	$10.55	$10.82	$10.67	$11.00	$10.85	$11.21	$11.06	$11.60	$11.43
	+$0.21	+$0.19	+$0.32	+$0.28	+$0.44	+$0.40	+$0.62	+$0.58	+$0.83	+$0.79	+$1.22	+$1.16
	+2.0%	+1.9%	+3.1%	+2.7%	+4.2%	+3.9%	+6.0%	+5.6%	+8.0%	+7.7%	+11.8%	+11.3%
-3.5	$10.57	$10.43	$10.64	$10.50	$10.76	$10.61	$10.92	$10.77	$11.11	$10.96	$11.37	$11.21
	+$0.19	+$0.16	+$0.26	+$0.23	+$0.38	+$0.34	+$0.54	+$0.50	+$0.73	+$0.69	+$0.99	+$0.94
	+1.8%	+1.6%	+2.5%	+2.2%	+3.7%	+3.3%	+5.2%	+4.9%	+7.0%	+6.7%	+9.5%	+9.2%
-4	$10.54	$10.41	$10.61	$10.47	$10.71	$10.56	$10.83	$10.69	$11.02	$10.87	$11.25	$11.10
	+$0.16	+$0.14	+$0.23	+$0.20	+$0.33	+$0.29	+$0.45	+$0.42	+$0.64	+$0.60	+$0.87	+$0.83
	+1.5%	+1.4%	+2.2%	+1.9%	+3.2%	+2.8%	+4.3%	+4.1%	+6.2%	+5.8%	+8.4%	+8.1%

Notes: BL = Bud Light, ML = Miller Lite and CL=Coors Light. For each {elasticity, diversion} combination the table reports the average predicted price of each brand following the JV, and, below the price, the change (in dollars and percent) from the average pre-JV price.

Figure 10: Bud Light Equilibrium Pricing Strategies



Note: figure shows the equilibrium strategies for Bud Light implied by the estimated parameters (shown in Table 6) when average own-price elasticities, at observed prices, are -3 and diversion to the outside good is 0.15.

of 0.15, the estimated AR(1) parameter for costs falls to 0.51 (from 0.61) and the predicted post-JV prices for BL and ML/CL are \$11.00 and \$10.85, which, to the nearest cent, are the same as reported in Table 7.

We can compare these predictions to the post-JV price changes implied by the regression estimates in Table 5. Recall that the most robust estimates suggested that the prices of BL, ML and CL rose by between 30 and 90 cents per 12-pack equivalent, or 3% and 5%, wih similar increases for each of these brands. These magnitudes are consistent with the predictions of our model for the full range of considered own price elasticities if we assume that diversion to the outside good (or brands not controlled by the leading domestic brewers) would constitute between 10% and 25% of the demand lost when a brand raises its price. Given that the domestic brands do not appear to have lost significant volume (Figure 9 (c) and (d)) when their relative prices rose after the JV, these low rates of diversion to the outside good appear to be plausible.

Table 8 compares price dynamics before and after the JV in the data and for two of the estimated models that predict post-JV price increases in the observed range. In the data we continue to observe three prices, whereas our model only allows two prices (one of which is the

Table 8: Comparison of Pre- and Post-JV Price Dynamics in the Data and the Estimated Model

	Data			Estimated Model Elasticity=-3 Diversion=.15			Estimated Model Elasticity=-3 Diversion=.25		
	Pre-JV Averages	Post-JV Averages	Change	Fitted	Prediction	Change	Fitted	Prediction	Change
Interquartile Range of BL Price ($)	0.22	0.26	0.04	0.21	0.35	0.14	0.22	0.3	0.08
Interquartile Range of ML and CL Price ($)	0.25	0.28	0.03	0.3	0.34	0.04	0.30	0.30	0
AR(1) Regression Coefficients									
$\rho^{BL,BL}$	0.459	0.550	0.091	0.425	0.462	0.037	0.430	0.447	0.017
Average of $\rho^{BL,ML}$ and $\rho^{BL,CL}$	0.057	0.100	0.043	0.073	0.151	0.078	0.059	0.122	0.063
Average of $\rho^{ML,ML}$ and $\rho^{CL,CL}$	0.421	0.542	0.121	0.450	0.443	-0.007	0.448	0.414	-0.034
Average of $\rho^{ML,BL}$ and $\rho^{CL,BL}$	0.041	0.094	0.053	0.035	0.139	0.104	0.032	0.150	0.118
SD of innovations									
..... from BL AR(1) regression	0.193	0.210	0.017	0.123	0.209	0.086	0.124	0.168	0.044
..... Average from ML & CL AR(1) regressions	0.212	0.220	0.008	0.179	0.208	0.029	0.171	0.166	-0.005

Notes: BL = Bud Light, ML = Miller Lite and CL=Coors Light. Values for the data are based on market average-brand-week retail 12-pack equivalent prices of all package sizes, excluding temporary store price reductions, for market-weeks where at least five stores are observed in the IRI data. Pre-JV values are averages of ML and CL values (where relevant) in both the data and the estimated models. Post-JV we report the values for the data using only the ML price, as our model assumes that ML and CL prices are the same. "....." indicates a measure that is related of the same type as the one reported in the previous row.

same for two products). Therefore, for comparison purposes, we report post-JV patterns for the data using the ML price (results are similar if we use the CL price), and, to be consistent with our analysis of model fit in Table 6, the reported data values are averages of the values estimated for each sample market. In both the data and simulations from the model, the dispersion of prices, the dispersion of price innovations and the own-price AR(1) coefficients increase after the JV, and BL prices become more sensitive to the lagged prices of rivals, although our model does not accurately match the relative sizes of each of these changes. Our model also correctly predicts that ML prices become more sensitive to lagged BL prices. The fact that we can match several qualitative changes in price dynamics after the JV, as well as the magnitude of changes in average price levels, even though no post-JV data was used when estimating the model, provides some additional support for the relevance of this type of model in this context.

6 Conclusion

We have developed a model where oligopolists simultaneously use prices to signal private information that is relevant for their future pricing decisions. While the possibility that this type of behavior would raise equilibrium prices was identified in the theoretical literature thirty years ago, we believe that we provide both the first attempt to quantify how large these effects may be and the first structural estimation of a model with multi-sided asymmetric information.

We show that even small amounts of private information - in the sense that the range that the privately observed state variable can take is narrow - can lead to large equilibrium price increases. Our empirical application examines pricing behavior in the US beer industry around the MillerCoors joint venture. Our model is able to match pre-joint venture price dynamics, and, for a range of plausible demand parameters, generate a pattern of price increases and changes in price dynamics which are qualitatively and quantitatively similar to those observed once the joint venture was completed, even when we allow for a marginal cost synergy that would have kept their prices from rising under complete information. Importantly, and in comparison with other explanations of why prices increased after the joint venture, we are able to explain what happened with a model of firm behavior that is the same both before and after the transaction.

While we focus on mergers, our result that equilibrium prices can diverge substantially from complete information levels could be important for any counterfactual in a setting where a few firms selling differentiated products are assumed to set prices. These effects are likely to be especially large when there is limited diversion to the outside good, and this is likely to be a feature of many markets for consumer staples as well as necessarily being a feature of markets where purchase is required by law (e.g., some insurance markets).

A natural question is whether, both in general and in our empirical setting, our model generates predictions that are different to models of tacit collusion developed in complete information frameworks. Tacit collusion models currently form the basis of almost all discussions of "coordinated effects" in the antitrust literature on mergers. The models are similar in the sense that, in both cases, firms raise their current prices in order to raise the future prices of rivals, although the mechanisms (signaling information in our model, avoiding a price war in the tacit collusion model) are different.

There are several additional differences. First, because the threat of a price war is potentially a much more powerful deterrent to lowering prices than a misinterpreted cost, tacit collusion can usually sustain higher equilibrium prices. As a result, it may only be possible to explain why firms do not collude or why, when they are believed to collude, they only raise prices by relatively modest amounts, which is a common finding in the merger retrospectives literature (Ashenfelter, Hosken, and Weinberg (2014)), by assuming that firms are quite impatient or are unable to punish quickly.[37] In contrast, our model can generate price increases of the observed size with frequent price setting, which is an attractive feature given that price changes occur frequently in most datasets. Second, our model provides a natural explanation for why observed prices may be fairly volatile. Ordover (2007) notes that price volatility is a feature of many settings where economists have claimed coordinated effects, but volatility cannot easily be explained by commonly considered types of collusive equilibria (such as trigger strategies) without quite large, temporary and commonly observed shocks to demand or costs. Third, our model would predict quite different effects if we assumed that firms set quantities or capacities, rather than prices, because in a quantity-setting game firms will typically want to indicate that their future quantities will be high so that their rivals reduce their output, and they will do this by acting more competitively. In contrast, tacit collusion models predict higher prices whatever strategic variables firms use, and, indeed, the literature on tacit collusion is often interpreted as suggesting

[37]For example, Miller, Sheu, and Weinberg (2019)'s tacit collusion model rationalizes the observed price increases after the MillerCoors joint venture in a framework that assumes prices are set annually with discount factors of around 0.7.

that anti-competitive coordinated effects will be more likely when products are homogeneous, which is where quantity-setting is typically assumed.[38]

While we have highlighted the differences between our asymmetric information model and traditional tacit collusion models, we believe that combining elements of both approaches would be valuable, as we recognize that threats of vigorous responses to price cuts, as well as asymmetric information, are plausible features of many markets. Kreps, Milgrom, Roberts, and Wilson (1982) and Athey and Bagwell (2008), who explicitly consider an example with elements of the Mailath (1989) model, have examined these links in stylized theoretical models. Applying our type of approach, where we calculate how large the effects on prices would be given realistic substitution patterns, to these models is a natural direction for future research.

[38]See footnote 4 for an example of how this observation influenced the early theoretical literature.

APPENDICES TO "DYNAMIC OLIGOPOLY PRICING WITH ASYMMETRIC INFORMATION AND HORIZONTAL MERGERS" FOR ONLINE PUBLICATION

A Computational Algorithms

This Appendix describes the methods used to solve our model. We describe the continuous type, finite horizon model in detail, before noting what changes in other cases. Our discussion will assume that there are two ex-ante symmetric duopolists. When firms are asymmetric, all of the operations need to be repeated for each firm.

A.1 Finite Horizon Model

A.1.1 Preliminaries

We specify discrete grids for the actual and perceived marginal costs of each firm, which will be used to keep track of expected per-period profits, value functions and pricing strategies. For example, when each firm's marginal cost lies on $[8, 8.05]$ and we use 8-point equally spaced grids, the points are $\{8, 8.0071, 8.0143, 8.0214, 8.0286, 8.0357, 8.0429, 8.0500\}$.[39] We use interpolation and numerical integration to account for the fact that realized types will lie between these isolated points. The discount factor is β.

It is useful to define several functions that we will use below:

- $P_{i,t}\left(\widehat{c^j_{i,t-1}}, c_{j,t-1}\right)$ is firm i's pricing function in period t. This is a function of the marginal cost that j believes that i had in the previous period, $\widehat{c^j_{i,t-1}}$ (which, when j is forming equilibrium beliefs, will reflect that cost that i signaled in the previous period). It will also depend on the marginal cost that i believes that j had in the previous period, but we solve the game assuming that j is using its equilibrium strategy, so that i assumes that its

[39] The examples reported in Section 3 use 12 gridpoints, although we have experimented with as many as 20 gridpoints in each dimension to make sure that this does not have a material effect on the reported results.

perception of j's prior cost is correct, so we use the argument $c_{j,t-1}$. The actual price set will depend on $c_{i,t}$, and, when we need to integrate over the values that $p_{i,t}$ may take (e.g., to calculate expected profits) we will include $c_{i,t}$ as an argument in the function.

- $\pi_i(p_{i,t}, p_{j,t}, c_{i,t})$ is firm i's one-period profit when it sets price $p_{i,t}$ and has marginal cost $c_{i,t}$, and its rival sets price $p_{j,t}$. As demand is assumed to be static and time-invariant this function does not depend on t.

- $V_{i,t}\left(c_{i,t-1}, \widehat{c^j_{i,t-1}}, c_{j,t-1}\right)$ is the value function for firm i defined at the beginning of period t, before firm types have evolved to their period t values. It reflects the expected payoffs of firm i in period t and the discounted value of expected payoffs in future periods given equilibrium play in both t and future periods. It depends on the true value of each firm's type in $t-1$, and the rival's perception of i's $t-1$ type (reflecting any deviation that i made in $t-1$). In the case of an 8-point grid, $V_{i,t}$ is a 512x1 vector.

- $\Pi_{i,t}\left(c_{i,t}, \widehat{c^j_{i,t}}, p_{i,t}, \widehat{c^j_{i,t-1}}, c_{j,t-1}\right)$ is the intermediate signaling payoff function of firm i when it knows its current marginal cost $c_{i,t}$, and is deciding what price to set. It does not know the period t type of its rival, but it reflects the pricing function that i expects j to use, $P_{j,t}\left(c_{j,t-1}, \widehat{c^j_{i,t-1}}\right)$. $\widehat{c^j_{i,t}}$ is the perception that j will have about i's cost at the end of period t. When the rival sets price $P_{j,t}\left(c_{j,t}, c_{j,t-1}, \widehat{c^j_{i,t-1}}\right)$,

$$\Pi_{i,t}\left(c_{i,t}, \widehat{c^j_{i,t}}, p_{i,t}, \widehat{c^j_{i,t-1}}, c_{jt-1}\right) = \int_{\underline{c}_j}^{\bar{c}_j} \left(\begin{array}{c} \pi_i\left(p_{i,t}, P_{j,t}\left(c_{j,t}, c_{j,t-1}, \widehat{c^j_{i,t-1}}\right), c_{i,t}\right) + \\ \beta V_{i,t+1}\left(c_{i,t}, \widehat{c^j_{i,t}}, c_{j,t}\right) \end{array} \right) \psi_j(c_{j,t}|c_{j,t-1}) dc_{j,t}.$$

where we note that $p_{i,t}$ only enters through current profits, and $\widehat{c^j_{i,t}}$ only enters through the discounted continuation value. We will also make use of the functions $\widetilde{\pi_i}$ and $\widetilde{V_{i,t}}$, where

$$\widetilde{\pi_i}\left(p_{i,t}, P_{j,t}\left(c_{j,t-1}, \widehat{c^j_{i,t-1}}\right), c_{i,t}\right) = \int_{\underline{c}_j}^{\bar{c}_j} \pi_i\left(p_{i,t}, P_{j,t}\left(c_{j,t}, c_{j,t-1}, \widehat{c^j_{i,t-1}}\right), c_{i,t}\right) \psi_j(c_{j,t}|c_{j,t-1}) dc_{j,t}$$

and

$$\widetilde{V_{i,t}}\left(c_{i,t}, \widehat{c^j_{i,t}}, c_{j,t-1}\right) = \int_{\underline{c}_j}^{\bar{c}_j} \beta V_{i,t+1}\left(c_{i,t}, \widehat{c^j_{i,t}}, c_{j,t}\right) \psi_j(c_{j,t}|c_{j,t-1}) dc_{j,t}$$

so that $\Pi_{i,t} = \widetilde{\pi}_i + \widetilde{V_{i,t}}$. Given a set of fully separating pricing functions $P_{i,t}\left(\widehat{c^j_{i,t-1}}, c_{j,t-1}\right)$, the relationship between Π and V is that

$$V_{i,t}\left(c_{i,t-1}, \widehat{c^j_{i,t-1}}, c_{j,t-1}\right) = \int_{\underline{c}_i}^{\overline{c}_i} \Pi_{i,t}\left(c_{i,t}, c_{i,t}, P_{i,t}\left(c_{i,t}, \widehat{c^j_{i,t-1}}, c_{j,t-1}\right), \widehat{c^j_{i,t-1}}, c_{jt-1}\right) \psi_i(c_{i,t}|c_{i,t-1}) dc_{i,t}$$

where we recognize that, in equilibrium, i's period t pricing function will reveal its cost to j, implying $\widehat{c^j_{i,t}} = c_{i,t}$.

A.1.2 Period T

Assuming that play at $T-1$ was fully separating, we solve for BNE pricing strategies for each possible combination of beliefs (on our grid) about $T-1$ marginal costs. A strategy for each firm is an optimal price given the realized value of its own T-period cost, given the pricing strategy of the rival, its prior marginal cost and the rival's belief about the firm's $T-1$ cost. Trapezoidal integration is used to integrate over the realized cost/price of the rival using a discretized version of the pdf of each firm's cost transition, and we solve for the BNE prices using the implied first-order conditions (i.e., those associated with maximizing static profits). With symmetric duopolists and 8-point grids, we find 512 equilibrium prices.

We use the equilibrium prices to calculate the beginning of period value function

$$V_{i,T}\left(c_{i,T-1}, \widehat{c^j_{i,T-1}}, c_{j,T-1}\right) = ...$$

$$\int_{\underline{c}_i}^{\overline{c}_i}\int_{\underline{c}_j}^{\overline{c}_j} \pi_i\left(P^*_{i,T}\left(c_{i,T}, \widehat{c^j_{i,T-1}}, c_{j,T-1}\right), P^*_{j,T}\left(c_{j,T}, c_{j,T-1}, \widehat{c^j_{i,T-1}}\right), c_{i,T}\right) \psi_j(c_{j,T}|c_{j,T-1})\psi_i(c_{i,T}|c_{i,T-1}) dc_{i,T} dc_{j,T}.$$

A.1.3 Period $T-1$

Firms choose prices in $T-1$ recognizing that their prices will affect rivals' prices in period T. We solve for $T-1$ strategies, assuming separating equilibrium pricing and interpretation of beliefs at $T-2$, so that each firm has a point belief about its rival's $T-2$ marginal cost. We then use the following steps to compute equilibrium strategies.

Step 1. (a) Compute

$$\widetilde{V}_{i,T-1}\left(c_{i,T-1}, \widehat{c^j_{i,T-1}}, c_{j,T-2}\right) = \beta \int_{\underline{c}_j}^{\overline{c}_j} V_{i,T}\left(c_{i,T-1}, \widehat{c^j_{i,T-1}}, c_{j,T-1}\right) \psi_j(c_{j,T-1}|c_{j,T-2}) dc_{j,T-1}.$$

$\widetilde{V}_{i,T-1}$ is the expected continuation value (i.e., not including $T-1$ payoffs) for i when it is setting its period $T-1$ price, without knowing the $T-1$ realization of c_j (but knowing that, in equilibrium, it will be revealed by $p_{j,T-1}$).

(b) Compute $\beta \frac{\partial \widetilde{V}_{i,T-1}\left(c_{i,T-1}, \widehat{c^j_{i,T-1}}, c_{j,T-2}\right)}{\partial \widehat{c^j_{i,T-1}}}$ using numerical differences at each of the gridpoints (one-sided as appropriate). This array provides us with a set of values for the numerator in the differential equation (1). These derivatives do not depend on $T-1$ prices, so we do not repeat this calculation as we look for equilibrium strategies.

(c) Verify belief monotonicity using these derivatives.

Step 2. We use the following iterative procedure to solve for equilibrium fully separating prices.[40] Use the BNE prices (i.e., those calculated in period T) as initial starting values. Set the iteration counter, $iter = 0$.

(a) Given the current guess of the strategy of firm j, $P_{j,T-1}\left(c_{j,T-1}, c_{j,T-2}, \widehat{c^j_{i,T-2}}\right)$, which is equal to the pricing functions solved for in the previous iteration, calculate $\frac{\partial \widetilde{\pi_{i,T-1}}\left(p_{i,T-1}, P_{j,T-1}\left(c_{j,T-2}, \widehat{c^j_{i,T-2}}\right), c_{i,T-1}\right)}{\partial p_{i,T-1}}$ for a grid of values $\left(p_{i,T-1}, \widehat{c^j_{i,T-2}}, c_{i,T-1}\right)$ where

$$\widetilde{\pi_{i,T-1}}\left(p_{i,T-1}, P_{j,T-1}\left(c_{j,T-2}, \widehat{c^j_{i,T-2}}\right), c_{i,T-1}\right) =$$

$$\int_{\underline{c}_j}^{\overline{c}_j} \pi_i\left(p_{i,T-1}, P_{j,T-1}\left(c_{j,T-1}, c_{j,T-2}, \widehat{c^j_{i,T-2}}\right), c_{i,T-1}\right) \psi_j(c_{j,T-1}|c_{j,T-2}) dc_{j,T-1}$$

i.e., the derivative of i's expected profit with respect to its price, given that it does not know what price j will charge because it does not know $c_{j,T-1}$. The derivatives are evaluated on a fine

[40] We do not claim that this iterative procedure is computationally optimal, although it works reliably in our examples. There are some parallels between our problem and the problem of solving for equilibrium bid functions in asymmetric first-price auctions where both the lower and upper bounds of bid functions are endogenous. Hubbard and Paarsch (2013) provides a discussion of the types of methods that are used for these problems.

grid (steps of one cent) of prices.[41] This vector will be used to calculate the denominator in the differential equation (1).

For each $\left(\widehat{c_{i,T-2}^{j}}, c_{j,T-2}\right)$,

(b) Solve the lower boundary condition equation $\frac{\partial\widetilde{\pi}\left(p_{i,T-1}^{*}, P_{j,T-1}\left(c_{j,T-1}, c_{j,T-2}, \widehat{c_{i,T-2}^{j}}\right), \underline{c_i}\right)}{\partial p_{i,T-1}} = 0$ for $p_{i,T-1}^{*}$, using a cubic spline to interpolate the vector calculated in (a). This gives the static best response price and the lowest price on i's pricing function.

(c) Using this price as the initial point[42], solve the differential equation, (1), to find i's best response signaling pricing function. This is done using `ode113` in MATLAB, with cubic spline interpolation used to calculate the values of the numerator and the denominator between the gridpoints.[43] Interpolation is then used to calculate values for the pricing function for the specific values of $c_{i,T-1}$ on the cost/belief grid $\left(c_{i,T-1}, \widehat{c_{i,T-2}^{j}}, c_{j,T-2}\right)$.

(d) update the current guess of i's pricing strategy using

$$P_{i,T-1}^{iter=k+1}\left(c_{i,T-1}, \widehat{c_{i,T-2}^{j}}, c_{j,T-2}\right) = (1-\tau)P_{i,T-1}^{iter=k}\left(c_{i,T-1}, \widehat{c_{i,T-2}^{j}}, c_{j,T-2}\right) + \ldots$$
$$\tau P_{i,T-1}^{'}\left(c_{i,T-1}, \widehat{c_{i,T-2}^{j}}, c_{j,T-2}\right) \quad \forall c_{i,T-1}, \widehat{c_{i,T-2}^{j}}, c_{j,T-2}$$

where $P_{i,T-1}^{'}$ are the best response functions that have just been computed. $\tau = 1$, i.e., full updating, works effectively unless we are close to prices where the conditions required to characterize the unique best response fail to hold, in which case we also try using $\tau = \frac{1}{1+iter^{\frac{1}{6}}}$.

(e) Check if the maximum difference between $P_{i,T-1}^{iter=k}$ and $P_{i,T-1}^{'}$, across all gridpoints, is less than 1e-6. If so, terminate the iterative process, else update the iteration counter to $iter = iter + 1$, and return to step 2(a).

(f) verify that the solved pricing functions are monotonic in a firm's own marginal costs, and

[41] A fine grid is required because it is important to evaluate the derivatives accurately around the static best response, where the derivative will be equal to zero.

[42] In practice, the exact value of the derivative will be zero at the static best response, so that the differential equation will not be well-defined if this derivative is plugged in. We therefore begin solving the differential equation at the price where $\Pi_3^{i,T-1} + 1e-4 = 0$. Pricing functions are essentially identical if we add 1e-5 or 1e-6 instead.

[43] We use an initial step size of 5e-5 and a maximum step size of 0.001.

that, given the pricing functions of the rival, that the single-crossing condition holds for the full range of prices used in the putative equilibrium.

Step 4. Compute i's value $V_{i,T-1}$,

$$V_{i,T-1}\left(c_{i,T-2},\widehat{c_{i,T-2}^{j}},c_{j,T-2}\right)=...$$

$$\int_{\underline{c}_i}^{\overline{c}_i}\int_{\underline{c}_j}^{\overline{c}_j}\left\{\begin{array}{c}\pi\left(P^{*}_{i,T-1}\left(c_{i,T-1},\widehat{c_{i,T-2}^{j}},c_{j,T-2}\right),P^{*}_{j,T-1}\left(c_{j,T-1},c_{j,T-2},\widehat{c_{i,T-2}^{j}}\right),c_{i,T-1}\right)\\ +\beta V_{i,T}\left(c_{i,T-1},c_{j,T-1},c_{i,T-1}\right)\end{array}\right\}*...$$

$$\psi_j(c_{j,T-1}|c_{j,T-2})\psi_i(c_{i,T-1}|c_{i,T-2})dc_{i,T-1}dc_{j,T-1}$$

where we are recognizing that equilibrium play at $T-1$ will reveal i's true cost to j. Note that this is the case even if $\widehat{c_{i,T-2}^{j}}\neq c_{i,T-2}$ (i.e., j was misled at $T-2$) because i should find it optimal to use its equilibrium signaling strategy given its new cost $c_{i,T-1}$ in response to j using a strategy based on its $\widehat{c_{i,T-2}^{j}}$ belief.

A.1.4 Earlier Periods

This process is then repeated for earlier periods, with an appropriate changing of subscripts. Given our assumption that first period beliefs reflect actual costs in a fictitious prior period this procedure will also calculate strategies in the first period of the game.

A.2 Infinite Horizon Model

We use an infinite horizon model for some of our examples and the empirical application. We find equilibrium pricing functions in the continuous type model using a modification of the procedure described above: in particular, we follow the logic of policy function iteration (Judd (1998)) to calculate values given a set of strategies.

The equilibrium objects that we need to solve for are a set of stationary pricing functions, $P_i^{*}\left(\widehat{c_{i,t-1}^{j}},c_{j,t-1}\right)$ and value functions $V_i\left(c_{i,t-1},\widehat{c_{i,t-1}^{j}},c_{j,t-1}\right)$ which are consistent with each other given the static profit function and the transition functions for firm types.

We start by solving the $T-1$ game described previously (i.e., assuming that there is a one more period of play where firms will use static Bayesian Nash Equilibrium strategies) to give an initial set of signaling pricing functions ($P_i^{*,iter=1}$). We then calculate firm values in each state $\left(c_{i,t-1}, \widehat{c^j_{i,t-1}}, c_{j,t-1}\right)$ if these pricing functions were used in every period of an infinite horizon game. This is done by creating a discretized form of the state transition process and calculating

$$\widehat{V}_i^{iter=1} = [I - \beta T]^{-1} \pi_i' \left(c_{i,t-1}, \widehat{c^j_{i,t-1}}, c_{j,t-1}\right)$$

where

$$\pi_i'\left(c_{i,t-1}, \widehat{c^j_{i,t-1}}, c_{j,t-1}\right) = \int_{\underline{c}_i}^{\bar{c}_i}\int_{\underline{c}_j}^{\bar{c}_j} \left\{ \pi_i \left(\begin{array}{c} P_i^{*,iter=1}\left(c_{i,t}, \widehat{c^j_{i,t-1}}, c_{j,t-1}\right), \\ P_j^{*,iter=1}\left(c_{j,t}, c_{j,t-1}, \widehat{c^j_{i,t-1}}\right), c_{i,t} \end{array} \right) \right\} \psi_j(c_{j,t}|c_{j,t-1})\psi_i(c_{i,t}|c_{i,t-1}) dc_{i,t} dc_{j,t}$$

and T is a transition matrix that reflects the transition probabilities for both firms' types and the behavioral assumption that equilibrium play in t (and future periods) will reveal period t costs. $P_j^{*,iter=1}\left(c_{j,t-1}, \widehat{c^j_{i,t-1}}\right)$ will reflect $P_i^{*,iter=1}$, applied to the states of the rival, when the firms are symmetric.

$\widehat{V}_i^{iter=1}$ is then used to compute a new set of pricing functions, $P_i^{*,iter=2}$, and the process is repeated until prices converge (tolerance 1e-4). Even though policy function iteration procedures do not necessarily converge, we find they work very well in our setting. As illustrated in the text, the pricing functions found by this method are essentially identical to the pricing functions found for the early periods of long finite horizon games where the exact value of t has almost no affect on equilibrium pricing strategies. The computational advantage of this procedure comes from the fact that we do not perform the iterative procedure described above for every period of the game: instead there is a single iterative procedure where we solve for a single set of pricing strategies for the entire game.

A.2.1 Speeding Up Solutions By Interpolating Pricing Functions

When we consider more than two firms and allow for asymmetries, the solution algorithm laid out above becomes slow, with most of the time spent solving differential equations. For example, with 8-point cost/belief grids, three asymmetric firms and 50 iterations, we would have to solve 25,600 differential equations. This would make estimation of the model using a nested fixed point procedure slow. On the other hand, reducing the number of gridpoints can lead to inaccurate calculations of expected payoffs, and therefore strategies.

Examination of the equilibrium pricing functions (see, for example, Figure 4) shows that as we vary rivals' prior types, a firm's pricing functions look like they are translated without (noticeably) changing shape. We exploit this fact by solving for pricing functions for only a subset of the points $\left(\widehat{c_{i,t-1}^{j}}, c_{j,t-1}\right)$ and using cubic splines to interpolate the remaining values.[44] This allows us to achieve a substantial speed increase, while continuing to calculate expected values accurately on a finer grid.

A.3 Two-Type Model

We use a model where each firm can have one of two types when we want to examine all strategies simultaneously or to consider a large number of alternative demand parameters. An additional advantage is that because prices, profits and values can be calculated for each possible type, we avoid small inaccuracies that result from numerical integration.

The key difference to the solution algorithm is that we no longer use solve differential equations to solve the pricing functions. Recall that in the continuous type model, the differential equations characterize the unique separating best response when the signaling payoff function satisfies several conditions. In the discrete type model, one can construct multiple separating pricing functions that can be supported for different beliefs of the rival firm. To proceed we therefore need to choose a particular pricing function.

To be as consistent with the continuous type model as possible, we use the prices that allow the two types to separate at the lowest cost, in terms of foregone current profits taking the current guess of the pricing function of the rival as given, to the signaling firm (i.e., "Riley" signaling strategies, which would also be those that satisfy application of the intuitive criterion).[45]

The amended computational procedure is as follows (described for the infinite horizon case). Suppose that we are looking to find the pricing strategy of firm i in period t when it believes that j's previous cost was $c_{j,t-1}$ and j believes that i's previous cost was $\widehat{c_{i,t-1}^{j}}$ (i.e., we will repeat this process for each $\left(\widehat{c_{i,t-1}^{j}}, c_{j,t-1}\right)$ combination, of which there will be four in the duopoly model. We need to solve for two prices: i's price when its cost is $\underline{c_i}$ and its price when its cost is $\overline{c_i}$.

[44] For example, when we estimate our model in Section 4, we use a seven-point cost grid ({1,..,7}) for the profits and values of each firm. We solve for pricing functions for the full interaction of gridpoints {1,3,5,7} and then interpolate the pricing functions for the remaining gridpoints.

[45] Of course, in the game we are considering it can be advantageous to the firms to use higher signaling prices, because of how this raises rivals' prices in equilibrium. This equilibrium consideration is ignored when selecting the Riley best response.

Step 1. Find $p^*_{i,t}(\underline{c_i})$, which will be the static best response, as the solution to $\frac{\partial \widetilde{\pi}\left(p_{i,t}, P_{j,t}\left(c_{j,t}, c_{j,t-1}, \widehat{c^j_{i,t-1}}\right), \underline{c_i}\right)}{\partial p_{i,t}} = 0$ where

$$\widetilde{\pi}_i\left(p_{i,t}, P_{j,t}\left(c_{j,t}, c_{j,t-1}, \widehat{c^j_{i,t-1}}\right), \underline{c_i}\right) =$$
$$\pi\left(p_{i,t}, P_{j,t}\left(\underline{c_j}, c_{j,t-1}, \widehat{c^j_{i,t-1}}\right), \underline{c_i}\right) \Pr(c_{j,t} = \underline{c_j}|c_{j,t-1}) + ...$$
$$\pi\left(p_{i,t}, P_{j,t}\left(\overline{c_j}, c_{j,t-1}, \widehat{c^j_{i,t-1}}\right), \underline{c_i}\right) \Pr(c_{j,t} = \overline{c_j}|c_{j,t-1})$$

Step 2. Find $p^*_{i,t}(\overline{c_i})$. This is done by finding the price, p', higher than $p^*_{i,t}(\underline{c_i})$, which would make the low cost firm indifferent between setting price $p^*_{i,t}(\underline{c_i})$ and being perceived as a low cost type, and setting price p' and being perceived as a high cost type, i.e.,

$$\widetilde{\pi}\left(p^*_{i,t}(\underline{c_i}), P_{j,t}\left(c_{j,t}, c_{j,t-1}, \widehat{c^j_{i,t-1}}\right), \underline{c_i}\right) + \beta\widetilde{V_{i,t+1}}(\underline{c_i}, \underline{c_i}, c_{j,t-1}) = ...$$
$$\widetilde{\pi}\left(p', P_{j,t}\left(c_{j,t}, c_{j,t-1}, \widehat{c^j_{i,t-1}}\right), \underline{c_i}\right) + \beta\widetilde{V_{i,t+1}}(\underline{c_i}, \overline{c_i}, c_{j,t-1})$$

where

$$\beta\widetilde{V_{i,t+1}}\left(c_{i,t-1}, \widehat{c^j_{i,t-1}}, c_{j,t-1}\right) = V_{i,t+1}\left(c_{i,t-1}, \widehat{c^j_{i,t-1}}, \underline{c_j}\right) \Pr(c_{j,t} = \underline{c_j}|c_{j,t-1}) + ...$$
$$V_{i,t+1}\left(c_{i,t-1}, \widehat{c^j_{i,t-1}}, \overline{c_j}\right) (1 - \Pr(c_{j,t} = \underline{c_j}|c_{j,t-1})).$$

We verify that, consistent with single-crossing, the $\overline{c_i}$ type prefers to set the price p' rather than setting its static best response price. We also verify belief monotonicity when we calculate the value functions. As illustrated in Section 3, there are parameters for which belief monotonicity fails.

B Locating Asymmetric Information at Alternative Places in the Model

In this Appendix we describe several variants of the model that also generate significant increases in equilibrium prices. In common with our baseline model, each variant assumes that each firm has a single piece of private information that it can signal using its choice of price, with differences coming from which piece of the profit function is private information. The fact that other formulations generate similar results is not surprising, but we perform the calculations in order to emphasize the point that we are not limited to the assumption that marginal costs are uncertain. In all cases we assume single product duopolists, as in Section 3, and use the continuous type, infinite horizon version of our model. The demand parameters take on their baseline values from Section 3 and marginal cost of each firm is held fixed at 8.

Variant 1: Weights on Profits and Revenues. In the first variant, we allow for there to be uncertainty about the weight that each firm places on profits rather than revenues. There are large theoretical and empirical literatures studying the question of whether managers want to maximize profits or alternative outcome variables, and whether shareholders might strategically choose to incentivize managers to deviate from profit maximization (Sklivas (1987), Katz (1991), Murphy (1999), De Angelis and Grinstein (2014)). The empirical literature suggests that managers are affected by a variety of incentives that may be complicated for outsiders to evaluate and which may vary over time, depending on oversight from shareholders or corporate boards, and financial constraints.

Without assuming a particular theory of governance, we suppose that the weight placed on profits by firm i in period t is $\tau_{i,t}$ and that this variable lies on the interval $[0.89, 0.9]$, with the remaining weight on firm revenues. As before, we suppose that the variable evolves according to a truncated AR(1) process, with $\rho = 0.8$. The standard deviation of the innovations is chosen so that, as for our baseline model where marginal costs are private information, the probability that a type will transition from the highest point of the support to below the median is 0.32.

As reported in the first panel of Table B.1 the average complete information price when both firms (are known to) maximize profits is 22.59. When a firm places some weight on revenues, it will tend to set a lower price, and the average static BNE or complete information price when the profit weight lies on $[0.89, 0.90]$ is 21.78. However, with signaling, average prices

Table B.1: Price Effects in Models Where Alternative Elements of Firm Objective Functions are Private Information

	Static BNE		Infinite Horizon Signaling Model	
Model	Mean Price	Std. Dev. Price	Mean Price	Std. Dev. Price
Weight on Firm Profits vs. Revenues is Private Information	21.79	0.01	24.45	0.59
Weight on Profits on [0.89,0.9]				
$\rho = 0.8$, std. dev. innovation 0.0044				
(eqm. prices if firms are known to maximize profits, i.e., if weight on profit equals 1, are 22.59)				
Weight on Firm vs. Industry Profits is Private Information	22.66	0.02	25.34	0.59
Weight on Industry Profit [0.00,0.02]				
$\rho = 0.8$, std. dev. innovation 0.0088				
(eqm. prices if firms are known to maximize own profits are 22.59)				
Size of Each Firm's Loyal Market is Private Information	25.17	0.07	27.56	0.59
Size of Loyal Market [0.10,0.12] (as fraction of duopoly market)				
$\rho = 0.8$, std. dev. innovation 0.0088				
(eqm. prices when loyal markets known to be 11% of duopoly markets are 25.17)				

increase significantly: in this example, the average Markov Perfect Bayesian Equilibrium price is 8.2% above the average price level *when both firms are known to maximize profits*, with profits increasing by 18%. This example suggests there may be some advantage to shareholders if they keep managers' incentives opaque to rivals even in markets where firms set prices for differentiated products.[46]

Variant 2: Weight on Profits of Other Firms in the Industry. In the empirical Industrial Organization literature it is common to model tacitly collusive behavior in a reduced-form way by generalizing static first-order conditions to allow for each firm to place some weight on the profits of other firms in the same market (Porter (1983), Bresnahan (1989), Miller and Weinberg (2017)). This type of formulation could also rationalized by models where participants in financial markets will improve their view of a firm's prospects when rivals' announce high profits (Rotemberg and Scharfstein (1990)) or if shareholders hold stock in competitors (O'Brien and Salop (1999), Azar, Schmalz, and Tecu (2018)).

We consider a model where rivals have some limited uncertainty about the weight that a firm places on its own profit rather than the profit of the industry. Formally we consider that each firm places a weight $\tau_{i,t}$ of $[0.98, 1]$ on its own profits, and $1 - \tau_{i,t}$ on the profits of the industry as a whole (of course, its own profits also contribute to industry profits). We assume that the transition process has ρ=0.8 and $\sigma = 0.0088$, which means that the probability of a type transitioning from the highest point of the support to below the median is 0.32, as in the first example. As can be seen in the second panel of Table B.1, the effect is, once again, to raise prices substantially in the dynamic game with asymmetric information.

Variant 3: Demand Shocks. In our experience, many economists are attracted to the idea that firms have private information about some aspect of demand rather than marginal costs. Some formulations of demand uncertainty give rise to signaling incentives that would be qualitatively different from the ones modeled within our framework. For example, suppose that demand has a logit structure and that each firm has private information about the serially correlated and unobserved quality of its product. Duopolist firms observe each other's prices but not quantities, so that prices are informative about quality. A firm with higher quality will want

[46] The usual explanation for why shareholders might want to commit to incentivizing their managers to place some weight on revenues comes from quantity-setting models where other firms will reduce their output when a firm's managers are committed to increase their output. In our model it is uncertainty about what firms are trying to maximize that causes equilibirum prices to rise, through the mechanism of signaling.

to charge a higher price, but its rival's optimal price will likely decrease in the firm's quality, so it is unclear whether a firm would want to be perceived as high quality or as low quality.

Here we consider a simple example where firms do have incentives to raise prices to signal that their demand is high. Suppose that each firm sells its products in two markets. In one market, the firms compete as duopolists, but in the other market the firm is a monopolist (so for example, both firms are in market A, firm 1 is the only firm in market B, and firm 2 is the only firm in market C). Due to the possibility of arbitrage, or some other constraint, each firm can only set one price across the markets. One rationalization of this setup would be that each firm has some loyal or locked-in customers, but that additional consumers are competed for. Product quality is known, but firms are uncertain about the size of their rival's loyal market. Normalizing the size of the common market to 1, the size of the other markets lie between $[0.1, 0.12]$. The utility specification is the same as before except loyal customers only choose between a single product and the outside good. The transition assumptions are the same as in variant 2. In this formulation, firms will set prices based on the weighted average marginal revenues from the two markets, and when the size of their monopoly market is larger they will prefer higher prices. A firm will therefore have incentive to raise its price to signal that its monopoly market is larger.

The results are presented in the third panel of Table B.1. The addition of the loyal market, where a firm's demand is less elastic, raises prices under all information structures, but the average signaling equilibium prices are 10% higher than the prices under complete information or in a static incomplete information game.

C Pre-Joint Venture Price Dynamics Under Alternative Definitions

As explained in Section 5, we estimate our model by matching predicted price levels, price dispersion and price dynamics, summarized using the coefficients from AR(1) regressions of prices, to values from retail price data from before the MillerCoors joint venture. The values from the data are calculated using a dataset of average brand-pack size-market-week prices, calculated excluding stores that have temporary price reductions. In the Appendix we provide some evidence on how the AR(1) coefficients (and the dispersion of the residuals) that we are trying to match change if we restrict attention to one particular pack size (12-packs) and/or aggregate our data to the monthly level.

Panel (b) of Table C.1 reports results of AR(1) regressions where we pool brand-pack size-market-week observations across geographic markets, including brand-pack size-week fixed effects so that we control for national price variation. This contrasts slightly with the process used to calculate the coefficients that are matched in estimation where we estimate separate AR(1) regressions for each geographic market separately. However, as can be seen by comparing the coefficients in panel (b) of Table C.1 and the average of the market-specific coefficients reported in parentheses in the left-hand column of Table 6, the cross-market average coefficients are similar to those from the pooled regressions.

Panel (b) (weekly, all pack sizes) is the closest to the specification used to estimate the model. The pattern is that a brand's prices are serially correlated, while there are positive, but sometimes statistically insignificant, correlations with the lagged prices of other brands. The distribution of the residuals from these price regressions consistently display negative skewness for all brands. Panel (d) uses monthly rather than weekly prices. The estimated own brand serial correlation coefficients tend to fall, but there are some (statistically insignificant) increases in the coefficients on the lagged prices of rival brands. As one might expect, averaging across months lowers the estimated variances of the residuals from the regressions.

Panels (a) and (c) use data on 12-packs, rather than all pack sizes, and either weekly or monthly average prices. The estimated own-brand serial correlation coefficients increase slightly, and the estimated residuals have smaller variance, but the other patterns remain unchanged.

Table C.1: Price Dynamics Under Alternative Definitions

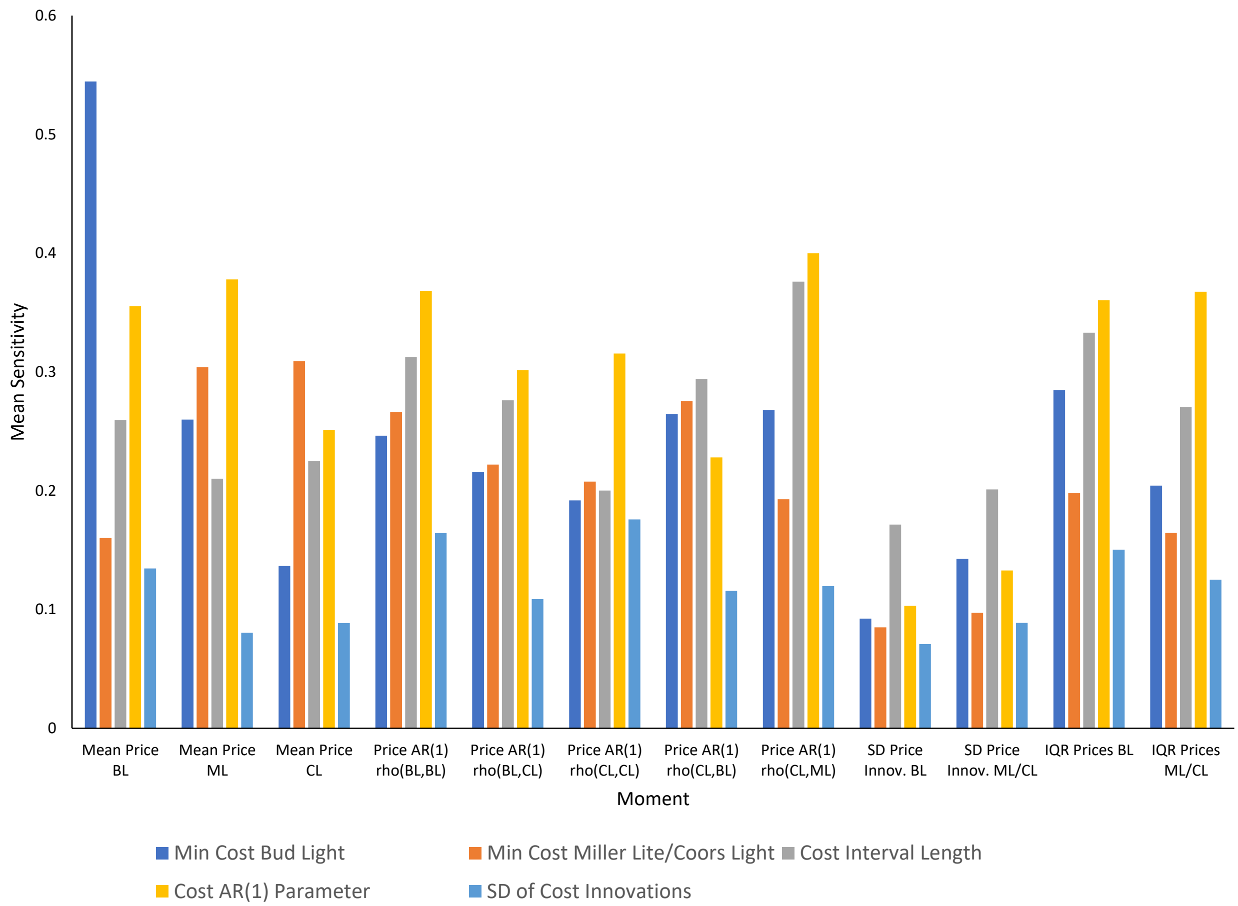
	(a) 12-pack price, weekly			(b) All pack sizes, weekly		
Dep. Var.	p_t^{BL}	p_t^{ML}	p_t^{CL}	p_t^{BL}	p_t^{ML}	p_t^{CL}
p_{t-1}^{BL}	0.510	0.071	0.022	0.447	0.040	0.036
	(0.036)	(0.027)	(0.045)	(0.041)	(0.019)	(0.013)
p_{t-1}^{ML}	0.042	0.476	0.022	0.016	0.383	0.005
	(0.019)	(0.044)	(0.014)	(0.012)	(0.043)	(0.018)
p_{t-1}^{CL}	0.001	0.033	0.572	0.019	0.012	0.438
	(0.016)	(0.018)	(0.043)	(0.014)	(0.017)	(0.047)
Mean Price	10.58	10.50	10.48	10.45	10.38	10.37
Innovations						
Std. Dev.	0.159	0.201	0.175	0.216	0.250	0.229
Skewness	-0.600	-0.987	-0.648	-0.799	-0.555	-0.428
Fixed Effects	Brand-Week Obs. Stores	Brand-Week Obs. Stores	Brand-Week Obs. Stores	Brand-Size-Week Obs. Stores-Size	Brand-Size-Week Obs. Stores-Size	Brand-Size-Week Obs. Stores-Size
Numb. of Obs.	8,046	8,045	8,044	29,818	29,853	29,860
Within R2	0.96	0.94	0.96	0.98	0.98	0.98

	(c) 12-pack price, monthly			(d) All pack sizes, monthly		
Dep. Var.	p_t^{BL}	p_t^{ML}	p_t^{CL}	p_t^{BL}	p_t^{ML}	p_t^{CL}
p_{t-1}^{BL}	0.347	0.130	0.081	0.290	0.058	0.037
	(0.047)	(0.047)	(0.044)	(0.044)	(0.026)	(0.022)
p_{t-1}^{ML}	0.126	0.330	0.119	0.062	0.286	0.056
	(0.029)	(0.033)	(0.032)	(0.020)	(0.062)	(0.033)
p_{t-1}^{CL}	0.027	0.073	0.342	0.082	0.107	0.377
	(0.021)	(0.049)	(0.069)	(0.016)	(0.031)	(0.058)
Mean Price	10.58	10.51	10.48	10.44	10.37	10.36
Innovations						
Std. Dev.	0.132	0.158	0.159	0.179	0.185	0.183
Skewness	-0.774	-0.775	-0.755	-0.754	-0.189	-0.317
Fixed Effects	Brand-Month Obs. Stores	Brand-Month Obs. Stores	Brand-Month Obs. Stores	Brand-Size-Month Obs. Stores-Size	Brand-Size-Month Obs. Stores-Size	Brand-Size-Month Obs. Stores-Size
Numb. of Obs.	1,826	1,826	1,826	6,945	6,945	6,941
Within R2	0.97	0.96	0.97	0.98	0.98	0.98

D Sensitivity Analysis for the Model Parameters

In this Appendix we report the results of an analysis of the sensitivity of the model parameters to the twelve moments used in estimation, following the approach proposed by Andrews, Gentzkow, and Shapiro (2017). This approach can provide a convenient way to summarize identification of the parameters in parametric structural models where changing a single parameter can affect multiple observed outcomes.

Figure D.1: Mean Absolute Values of the Parameter Sensitivity Statistics Calculated Following Andrews, Gentzkow, and Shapiro (2017)



Note: reported values are average values of the sensitivity measure for each moment/parameter combination across the thirty demand specifications considered in the main text. Sensitivity is calculated as $-(\widehat{G}'W\widehat{G})^{-1}\widehat{G}'W$ where $\widehat{G}$ is the Jacobian of the moments with respect to the parameters evaluated at the parameter estimates and W is the weighting matrix.

Figure D.1 shows the average of the absolute values of each sensitivity parameter (a parameter-moment combination) when we average across the 30 sets of demand parameters that we consider in our empirical estimation.[47] The different color columns represent different parameters, while the different groups of columns reflect different moments.

The minimum cost of Bud Light parameter is most sensitive to the average price of Bud Light moment, while, as one would expect, the minimum cost of Miller Lite/Coors Light (constrained to be the same in our model) is more (and equally) sensitive to the average prices of these brands.

The length of the cost interval parameter is most sensitive to several of the moments that measure the serial correlation in prices. This may seem counter-intuitive, in the sense that one might have expected it to be most closely related to the dispersion in observed prices, but it reflects the nature of equilibrium strategies in our model. When there is greater serial correlation in prices, signaling incentives increase and the range of equilibrium prices increases for a given length of cost interval. Therefore, we should expect the observed serial correlation in prices to have an effect on the estimated cost interval. Similar reasoning explains why the standard deviation of cost innovation parameter is also sensitive to the AR(1) coefficient moments.

[47] While many of the patterns are similar across the various demand specifications, the averages provide a clearer, and more intuitive, picture.

References

Andrews, I., M. Gentzkow, and J. M. Shapiro (2017): "Measuring the Sensitivity of Parameter Estimates to Estimation Moments," *Quarterly Journal of Economics*, 132(4), 1553–1592.

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. (2016): "Diagnosing Foreclosure Due to Exclusive Dealing," *Journal of Industrial Economics*, 64(3), 375–410.

Asker, J., C. Fershtman, J. Jeon, and A. Pakes (2018): "A Computational Framework for Analyzing Dynamic Procurement Auctions: The Market Impact of Information Sharing," Discussion paper, Boston University.

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

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

Bajari, P., C. L. Benkard, and J. Levin (2007): "Estimating Dynamic Models of Imperfect Competition," *Econometrica*, 75(5), 1331–1370.

Bonatti, A., G. Cisternas, and J. Toikka (2017): "Dynamic oligopoly with incomplete information," *The Review of Economic Studies*, 84(2), 503–546.

Borenstein, S. (1990): "Airline Mergers, Airport Dominance, and Market Power," *American Economic Review*, 80(2), 400–404.

Bresnahan, T. F. (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.

Chandra, A., and M. Weinberg (2018): "How Does Advertising Depend on Competition? Evidence from US Brewing," *Management Science*, 64(11), 5132–5148.

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

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.

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

Fogarty, J. (2010): "The Demand for Beer, Wine and Spirits: A Survey of the Literature," *Journal of Economic Surveys*, 24(3), 428–478.

Hausman, J., G. Leonard, and J. D. Zona (1994): "Competitive Analysis with Differentiated Products," *Annales d'Economie et de Statistique*, (34), 159–180.

Hellerstein, R. (2008): "Who Bears the Cost of a Change in the Exchange Rate? Pass-through Accounting for the Case of Beer," *Journal of International Economics*, 76(1), 14–32.

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.

Kim, E. H., and V. Singal (1993): "Mergers and Market Power: Evidence from the Airline Industry," *American Economic Review*, 83(3), 549–69.

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.

Mailath, G. (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.

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 (2019): "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.

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

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," *Colum. Bus. L. Rev.*, p. 411.

PAKES, A., AND P. MCGUIRE (1994): "Computing Markov-Perfect Nash Equilibria: Numerical Implications of a Dynamic Differentiated Product Model," *RAND Journal of Economics*, pp. 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.

ROJAS, C., AND E. B. PETERSON (2008): "Demand for Differentiated Products: Price and Advertising Evidence from the US Beer Market," *International Journal of Industrial Organization*, 26(1), 288–307.

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

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): *New Palgrave Dictionary of Economics, 2nd Edition*chap. Indirect Inference. Palgrave Macmillan.

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

Toro-González, D., J. J. McCluskey, and R. C. Mittelhammer (2014): "Beer Snobs Do Exist: Estimation of Beer Demand by Type," *Journal of Agricultural and Resource Economics*, 39(2), 174.

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.