Online Appendices for
"DYNAMIC OLIGOPOLY PRICING WITH ASYMMETRIC INFORMATION: IMPLICATIONS FOR HORIZONTAL MERGERS"

Sweeting, Tao and Yao

June 2023

A Computational Algorithms for Continuous Type Models

This Appendix describes the methods used to solve continuous type models. The methods used for discrete type models are detailed in Appendix B. We describe the methods used for our finite horizon continuous type examples in detail, before noting differences when we solve for infinite horizon games for our examples and in estimation. For simplicity, our discussion will assume that there are two ex-ante symmetric single-product 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, if each firm's marginal cost lies on $[8, 8.05]$ and we use 8-point equally spaced grids, the points are $\{8,\ 8.0071,\ 8.0143,\ 8.0214,\ 8.0286,\ 8.0357,\ 8.0429,\ 8.0500\}$.[1] We use interpolation and numerical integration to account for the fact that realized types will lie between these isolated points. The discount factor is $\beta = 0.99$.

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 perception

[1] Our symmetric duopoly examples in Section 3 use 20 gridpoints so that we can be as accurate as possible.

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 explicit argument in the function.

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

- $V_{i,t}\left(c_{i,t-1}, \widehat{c^j_{i,t-1}}, c_{j,t-1}\right)$ 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. In practice, it is useful to split $\Pi^{i,t}$ into two parts: $\Pi^{i,t} = \widetilde{\pi}_i + \widetilde{V_{i,t}}$, 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

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

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

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

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 in period $T-1$ was fully separating, we solve for BNE pricing strategies for each possible combination of beliefs (on our grid) about period $T-1$ marginal costs. A strategy for each firm is an optimal price given the realized value of its own period T cost, given the pricing strategy of the rival, its prior marginal cost and the rival's belief about the firm's period $T-1$ 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 would find 512 equilibrium prices.

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

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

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

A.1.3 Period $T-1$.

Firms choose prices in period $T-1$ recognizing that their prices will affect rivals' prices in period T. We solve for period $T-1$ strategies, assuming separating equilibrium pricing and interpretation of beliefs in period $T-2$, so that each firm has a point belief about its rival's period $T-2$ 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 period $T-1$ payoffs) for i when it is setting its period $T-1$ price, without knowing the period $T-1$ 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 (2). These derivatives do not depend on period $T-1$ 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.[2] 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 grid (steps of one cent) of prices.[3] This vector will be used to calculate the denominator in the

[2]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) provide a discussion of the types of methods that are used for these problems.

[3]A fine grid is required because it is important to evaluate the derivatives accurately around the static best

differential equation (2).

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

(b) Solve the lower boundary condition equation $\frac{\partial\tilde{\pi}\left(p^*_{i,T-1}, P_{j,T-1}\left(c_{j,T-1}, c_{j,T-2}, \widehat{c^j_{i,T-2}}\right), \underline{c_i}\right)}{\partial p_{i,T-1}} = 0$ 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[4], solve the differential equation, (2), 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. 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 $\left(c_{i,T-1}, \widehat{c^j_{i,T-2}}, c_{j,T-2}\right)$.

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

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

where $P'_{i,T-1}$ are the best response functions that have just been computed. For our examples, we use $\tau = 1$, i.e., full updating.

(e) Check if the maximum difference between $P^{iter=k}_{i,T-1}$ and $P'_{i,T-1}$, across all gridpoints, is less than 1e-4. 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 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.

response, where the derivative will be equal to zero.

[4] 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 $\Pi_3^{i,T-1} + 1\text{e-}4 = 0$. Pricing functions are essentially identical if we add 1e-5 or 1e-6 instead.

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

$$V_{i,T-1}\left(c_{i,T-2},\widehat{c^{j}_{i,T-2}},c_{j,T-2}\right)=...$$
$$\int\limits_{\underline{c}_i}^{\overline{c}_i}\int\limits_{\underline{c}_j}^{\overline{c}_j}\left\{\begin{array}{c}\pi\left(P^{*}_{i,T-1}\left(c_{i,T-1},\widehat{c^{j}_{i,T-2}},c_{j,T-2}\right),P^{*}_{j,T-1}\left(c_{j,T-1},c_{j,T-2},\widehat{c^{j}_{i,T-2}}\right),c_{i,T-1}\right)\\ +\beta V_{i,T}\left(c_{i,T-1},c_{j,T-1},c_{i,T-1}\right)\end{array}\right\}\times ...$$
$$\psi_j(c_{j,T-1}|c_{j,T-2})\psi_i(c_{i,T-1}|c_{i,T-2})dc_{j,T-1}dc_{i,T-1}$$

where we are recognizing that equilibrium play at period $T-1$ will reveal i's true cost to j. Note that this is the case even if, hypothetically, $\widehat{c^{j}_{i,T-2}}\neq c_{i,T-2}$ (i.e., j was misled in period $T-2$) 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 $\widehat{c^{j}_{i,T-2}}$ 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, $P_i^{*}\left(\widehat{c^{j}_{i,t-1}},c_{j,t-1}\right)$ and value functions $V_i\left(c_{i,t-1},\widehat{c^{j}_{i,t-1}},c_{j,t-1}\right)$ which are consistent with each other given the static profit function and the transition functions for firm types.

We start by solving the period $T-1$ 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

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

where

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

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. $P_j^{*,iter=1}\left(c_{j,t-1}, \widehat{c_{i,t-1}^j} \right)$ 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, when the conditions for separation hold, although it is sometimes necessary to update the pricing function to be a linear combination of the previous guess and the newly calculated best response. 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.3 Changes to Algorithm Used to Calibrate the Parameters.

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 very slow. On the other hand, reducing the number of gridpoints too much can lead to inaccurate calculations of expected payoffs.

Examination of the equilibrium pricing functions (see, for example, Figure 3) 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

$\left(\widehat{c^j_{i,t-1}}, c_{j,t-1}\right)$ gridpoints and using cubic splines to interpolate the remaining values.[5] This allows us to achieve a substantial speed increase, while continuing to calculate expected values accurately on a finer grid.

A.3.1 Tolerances and Updating Rules Used for the Calibration of the Cost Parameters Using the Infinite Horizon Model.

In Section 3 we calibrate the cost parameters using a nested fixed point algorithm, which means that both speed and accuracy are important. After considerable experimentation, we use the following tolerances:

- for the parameter search using `fminsearch`, we set the tolerance for the parameter values at 1e-5 and the tolerance on changes to the objective function at 1e-5. The value of the minimized objective function is typically less than 0.0002. In comparison, calibrated parameters assuming firms that use static Bayesian Nash strategies give an objective function value of around 0.2.

- the tolerance for evaluating whether the pricing functions have converged when solving the model is $10^{-6} * (iter + 1)^{0.642}$. Convergence usually happens within 45 iterations, when the tolerance is around 1e-5.

- for the differential equation solver, the initial step size is 5e-5 and the maximum step size is 0.003 for the first ten iterations of the algorithm, but we then use an initial step size of 1e-5 and a maximum step size of 0.001.

- we update the pricing function to be the best response for the first 15 iterations, and then use a linear combination of the best response and the current guess where the weight on the best response changes linearly from 1 (iteration 16) to 0.1 (iteration 115).

When we use these tolerances, the infinite horizon game is typically solved once using somewhere between 12 and 45 iterations, taking between 3 and 20 minutes. Calibration of the five parameters can take up to 250 function evaluations, although the objective function and parameters are usually very close to their final values within 100 evaluations.

[5] For example, when we estimate our model in Section 3, 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.

B Discrete (Two-)Type Examples.

A model where each firm can have one of two types has a much lower computational burden than the continuous type model. In this Appendix we explain how we solve two-type models and then present two additional analyses. One analysis explores the relationship between the serial correlation of costs, the diversion of demand to the outside good, the existence of separating equilibria and the size of equilibrium price effects. The second analysis considers an example where no equilibrium exists to illustrate why belief monotonicity can fail in a multi-period game.

In all of them we assume that firms are symmetric and that in any period $c_i = \underline{c} = 8$ or $c_i = \overline{c} = 8.05$. The probability that the cost remains the same as in the previous period is $0.5 \leq \rho < 1$. There are no signaling incentives when $\rho = 0.5$.

B.1 A Necessary Refinement.

A disadvantage of the two-type model is that for a given pricing strategy of firm j, firm i's separating best response pricing function is not unique in the sense that it depends on how firm j will interpret the signal.

We therefore impose a refinement that is consistent with the logic of the "intuitive criterion" (Cho and Kreps (1987)), which has often been applied as a refinement in discrete-type signaling games where only one player is signaling. Specifically, when we are looking for a player's separating best response, we find the low cost type's strategy will be the static best response, as in the continuous type model, and the high cost type's best response price will be the lowest price that the low cost type would be unwilling to set even if this would result in rivals' perceiving it as a high cost type rather than a low cost type.

While this refinement uniquely defines the best response, it does not guarantee a unique equilibrium in the oligopoly signaling game, and we have identified several examples in the infinite horizon version of the two-type model where there are multiple equilibria. Also, while the best response we identify is the lowest cost separating equilibrium given our current guess of the separating strategy of the rivals, this does not imply that it would be the lowest cost separating strategy if we also allowed the rival's strategy to change.

B.2 Solution Method.

The basic steps of the algorithms are the same as described in Appendix A, except that:

- we do not need to interpolate, as we can use a gridpoint for each possible state.

- we no longer solve differential equations to find best response pricing functions.

The computational for finding best response functions 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 $\widehat{c_{i,t-1}^{j}}$. 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 $\overline{c_i}$.

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

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

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.,

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

where

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

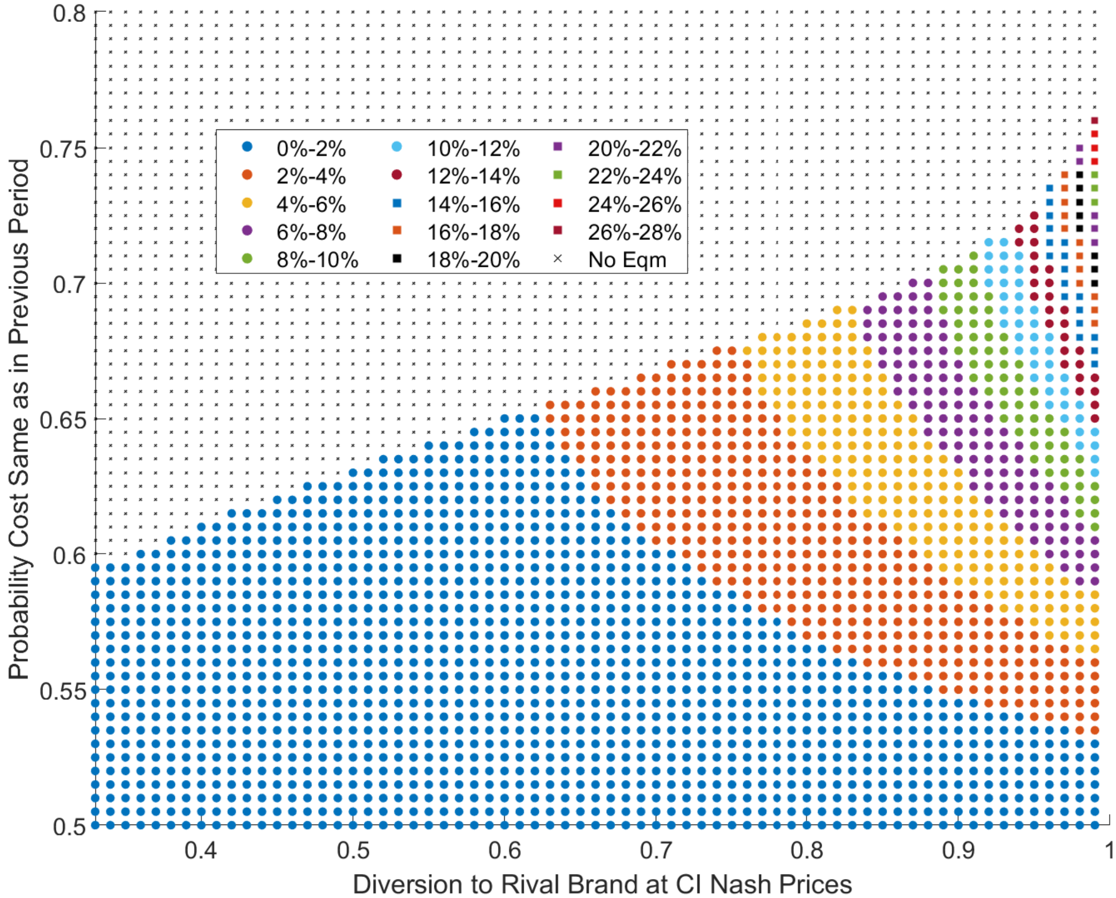
We verify that the $\overline{c_i}$ type prefers to set the price p' rather than setting its static best response price. We also verify belief monotonicity when we calculate the value functions.

B.3 Outcomes for Alternative Serial Correlation and Demand Parameters.

We assume nested logit demand where the indirect utility function for consumer c has the form $u_{i,c} = \beta - \alpha p_i + 0.25\nu_c + (1-0.25)\epsilon_{i,c}$. We choose β and α so that, for each combination of parameters that we consider, the CI equilibrium prices (at average cost levels) are \$16 for each firm and the diversion, which measures the proportion of a product's lost demand that goes to the rival's product, rather than the outside good, when its price increases from the CI equilibrium price, has a value that we specify. We focus on this diversion because when more demand goes to the outside good, which is like a competitor that always offers a fixed utility and does not respond to a signal, firms have less incentive to signal and, as we will show, the belief monotonicity and single-crossing conditions become harder to satisfy. We vary ρ from 0.5 (in which case there is no incentive to signal) to 0.85. We solve an infinite horizon version of our model.

Figure B.1 shows the results for a fine grid of values of diversion and ρ. The grey crosses indicate (diversion, serial correlation) combinations where we cannot find a separating equilibrium. For combinations where we can find a separating equilibrium, the colors indicate the percentage increase in average prices relative to average static Bayesian Nash equilibrium prices with the same demand and serial correlation parameters (BNE prices are always very close to the CI Nash prices of \$16). When serial correlation is very low, the price effects are always small whatever the level of diversion, consistent with how there is little incentive to signal when signals are uninformative. For given demand diversion, price effects become larger as serial correlation increases, until serial correlation reaches a level at which a signaling equilibrium cannot be sustained. For given serial correlation, higher diversion to the rival product is associated with the existence of separating equilibria and larger price effects, reflecting how 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 how an 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 0.65, and the effects of signaling are not very large. However, when diversion to the rival is very high, equilibria exist with price effects of over 20%.

Figure B.1: Equilibrium Average Price Increases in the Infinite Horizon Two-Type Duopoly Model as a Function of Diversion and Serial Correlation of Costs



Notes: the colored circles and squares are cases where we find a signaling equilibrium, and the colors represent the increase in prices relative to a static BNE. The grey crosses mark cases where we cannot find a separating equilibrium.

B.4 Failure of the Conditions Required for Existence of a Separating Equilibrium.

We use a finite horizon example to explain why a separating equilibrium may fail to exist in a sufficiently long game. Demand is the same as in our examples in the text (i.e., indirect utility is $u_{i,c} = 5 - 0.1p_i + 0.25\nu_c + (1-0.25)\epsilon_{i,c}$), and each firm's marginal cost is either 8 (low) or 8.075 (high). We assume that $\rho = 0.99$ so a signal is very informative about next period's marginal costs and signaling incentives are strong. Given these parameters, a separating equilibrium exists only in a game with a small number of periods.

Figure B.2: Equilibrium Prices in the Two-Type Marginal Cost Model (parameters described in the text)

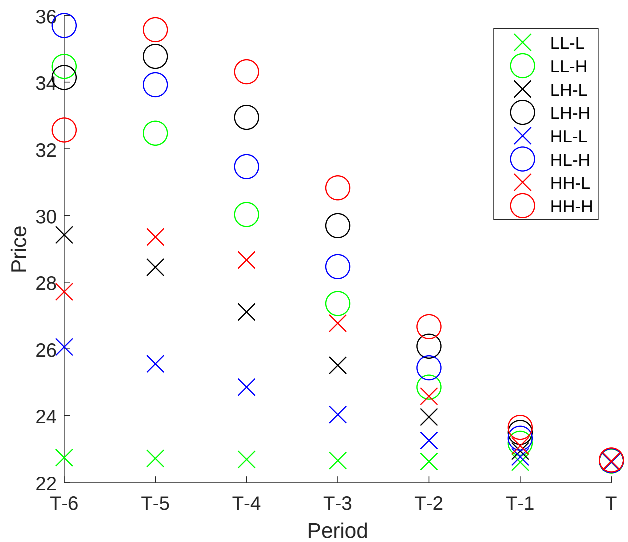


Figure B.2 shows the full set of eight equilibrium prices in each period as we move backwards from the end of the game. The legend denotes states by {"the firm's perceived cost in $t-1$", "its rival's perceived cost in $t-1$" - "the firm's realized marginal cost in t"} so blue indicates prices for a 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).

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 period $T-6$ 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 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 than if it is believed to be high cost. As profits increase in the rival's price, this will lead belief monotonicity to be violated.

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).

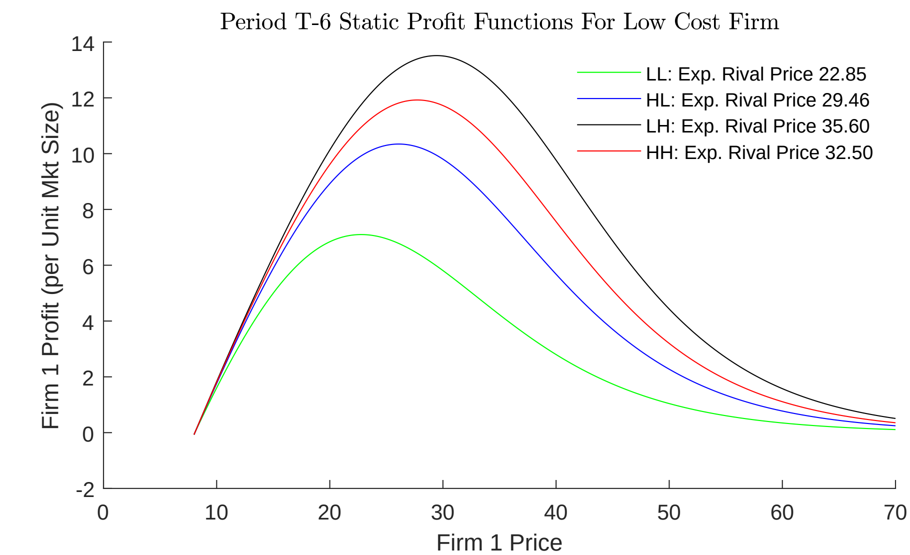
Recall that in the two-type model the equilibrium price of the $\overline{c}$ type is determined by the lowest price that the low-cost firm would be unwilling to set even if choosing it would lead to it being perceived as high cost. Consider the cost, in terms of foregone period $T-6$ profit, for a low-cost firm of raising its price. The upper panel of Figure B.3 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.[6] 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 (i.e., a firm likely to face a high cost rival) than a blue firm (i.e., a low cost rival) so it is less costly for the blue firm to raise its price.[7]

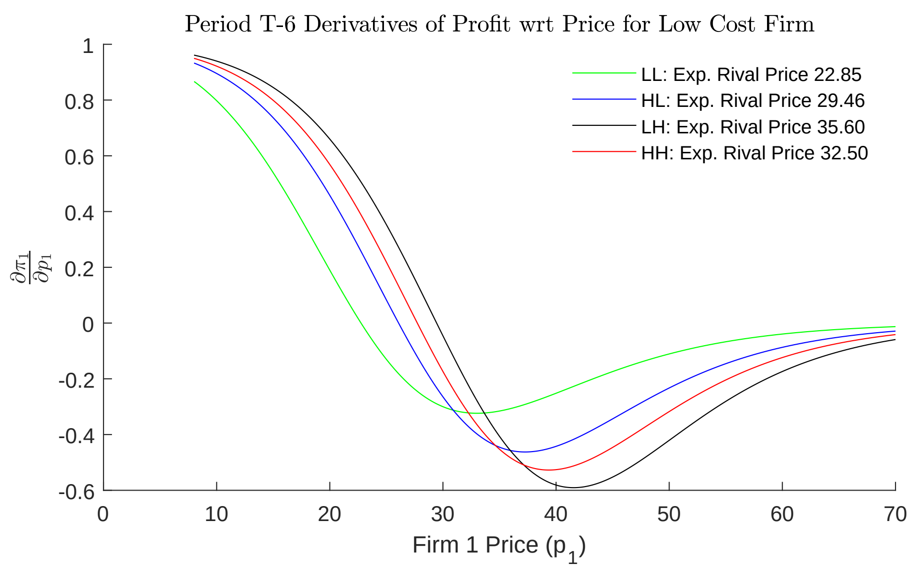
Now consider the incentive of a low-cost firm to signal (i.e., to pretend to be high-cost). The incentive of an HL (blue) firm to signal a high cost in period $T-6$ is that it is very likely to lead to its rival setting the black cross, rather than the green cross, price in period $T-5$. This difference is large, so that the incentive to signal is strong (so a high cost firm must set a higher price in order to not be copied by a low cost firm). 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 in period $T-5$. These period $T-5$ prices are closer together (than the black and green crosses) so the incentive to signal will tend to be weaker. Therefore, the cost of signaling and the incentive to signal combine to lead to a reversal of the order of the period $T-6$ equilibrium prices, causing belief monotonicity to fail in period $T-7$.

[6] 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.

[7] 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.

Figure B.3: Period $T-6$ Profit Functions in the Two-Type Game





C Supporting Materials for our Empirical Application: The Effects of the MillerCoors Joint Venture (MCJV)

This Appendix describes the data used in our analysis, as well as additional analyses that support the assumptions and interpretations presented in the text. The MCJV, announced in October 2007, effectively merged the U.S. brewing, marketing and sales operations of SABMiller (Miller) and MolsonCoors (Coors), the second and third largest U.S. brewers. The Department of Justice (DOJ) decided not to challenge the transaction in June 2008 because it expected "large reductions in variable costs of the type that are likely to have a beneficial effect on prices".[8] For example, production of Coors at Miller breweries around the country would lower the JV's transportation costs (Ashenfelter, Hosken and Weinberg (2015)). Readers are referred to Miller and Weinberg (2017) for more background on the JV.

C.1 IRI Data.

The data comes from the beer category of the IRI Academic Dataset (Bronnenberg, Kruger and Mela (2008)). The underlying data is at the weekly UPC-store-level from 2001 to 2011. We only use data from grocery stores. We now describe the data that we use when calibrating the model, noting where there are differences to the selection used by MW.

- **time period:** the IRI data stretches from January 2001 to December 2011. Our calibration will use the period up to October 2007, when the JV was announced, as the pre-JV period. MW use the period from Janaury 2005 to December 2011, excluding June 2008 to May 2009 (i.e., the period immediately after consummation) to estimate demand.

- **selection of markets:** the data comes from 45 geographic (IRI defined) regional markets where we observe the flagship brands (Bud Light, Miller Lite and Coors Light) being sold in at least 5 stores in some weeks. MW use 39 of these markets, excluding some markets where only a small number of observed grocery stores sell beer, in their analysis.

- **brands:** the calibration uses data on the flagship brands BL, ML and CL. However, MW include additional brands in their analyses, specifically domestic brands Budweiser, Michelob Ultra, Michelob Light, Miller Genuine Draft, Miller High Life and Coors, and imported brands Corona Extra, Corona Light, Heineken and Heineken Premium Light.

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

- **products and pack sizes:** the most popular packages are of cans and glass bottles in the equivalent of 6, 12, 18, 24, 30 and 36 12oz. servings, with 36 servings the least common and hardly observed at all at the start of the data. Our calibration will use data on 6, 12, 18, 24 and 30 12oz. serving pack sizes, with prices and quantities always transformed into 12-pack equivalents, and cans and bottles aggregated together. MW use data on 6, 12, 24 and 30 packs, aggregating 24 and 30 into a common "large" pack size. We keep 18 packs as over 20% of flagship sales came in this pack size in 2007, immediately before the JV.

- **deflator:** when using real prices, we deflate to January 2010 levels using the CPI-U All Urban Consumers-All Items price index.

- **distance and demographics:** our demand estimates and regressions that use distances between breweries/ports of entry and each market, use the variables that MW construct for the distribution of household income and transportation distances.

C.2 Summary Statistics Around the JV.

Table C.1 lists the 20 brands with the largest sales by volume in 2007, together with additional brands that MW include in their analysis. The table lists market shares and average nominal prices (per 144 oz, the volume in a standard 12-pack) in 2007 and 2011.

Relative to imported brands, the prices of domestic brands increased after the JV. However, it is noticeable that domestic brands tend to sell at lower prices and in larger package sizes than imported brands, and, consistent with imported brands being relatively poor substitutes to the domestic brands for price sensitive consumers, the change in relative prices is not associated with the imported brands gaining significant market share. This is also true at the brewer level. For example, AB's volume share was 41.3% in 2007 (before the JV), 41.5% in 2009 (after the JV) and 39.6% in 2011, with AB's light beer shares were 50.0%, 50.8% and 50.6% respectively.

C.3 Effects of the Joint Venture on Prices.

MW present estimates of the effects of the joint venture on prices. We present complementary estimates here, which can be compared to the price increases predicted by our calibrated model, assuming, in both cases, that the JV was an unanticipated and exogenous change in market structure.

Table C.1: Highest-Selling Beer Brands in 2007 with Ownership, Share and Average Nominal Retail Prices per 12-Pack.

Brand	2007 Company	2007 Packs	2007 % 18+	2007 Mkt. Share	2007 Price	2011 Mkt. Share	2011 Price
Bud Light*,†	AB	10	0.73	15.6%	$8.26	15.7%	$8.87
Miller Lite*,†	M	10	0.75	9.8%	$8.08	8.1%	$8.63
Coors Light*,†	C	10	0.75	8.3%	$8.29	9.3%	$8.93
Budweiser†	AB	10	0.72	7.9%	$8.23	6.7%	$8.93
Corona Extra$^{\dagger,\diamond}$	GM	5	0.15	4.0%	$13.83	3.7%	$13.45
Natural Light*	AB	7	0.69	3.9%	$6.00	3.1%	$7.11
Busch Light*	AB	9	0.79	2.8%	$6.04	2.5%	$6.92
Miller High Life†	M	9	0.56	2.4%	$6.32	2.2%	$7.19
Heineken$^{\dagger,\diamond}$	H	7	0.13	2.3%	$13.96	2.3%	$13.76
Miller Genuine Draft†	M	10	0.67	2.2%	$8.25	1.3%	$8.91
Michelob Ultra*,†	AB	9	0.29	2.1%	$10.02	2.5%	$10.35
Busch	AB	9	0.70	1.9%	$6.08	1.6%	$7.03
Keystone Light*	C	6	0.82	1.5%	$5.78	1.5%	$6.93
Budweiser Select	AB	9	0.62	1.3%	$8.36	0.7%	$8.73
Milwaukee's Best Light*	M	6	0.67	1.2%	$5.36	0.8%	$6.16
Corona Light$^{*,\dagger,\diamond}$	GM	3	0.03	1.2%	$14.13	1.3%	$13.72
Tecate$^{\diamond}$	H	7	0.66	1.1%	$8.68	1.0%	$9.06
Natural Ice	AB	7	0.52	1.1%	$5.93	0.9%	$7.15
Pabst Blue Ribbon	SP	9	0.50	1.0%	$6.25	1.4%	$7.51
Milwaukee's Best	M	6	0.61	0.8%	$5.46	0.4%	$6.42
Coors†	C	10	0.74	0.8%	$8.38	0.9%	$8.80
Michelob Light*,†	AB	9	0.35	0.8%	$9.69	0.4%	$10.07
Heineken Prem. Light$^{*,\dagger,\diamond}$	H	5	0.03	0.6%	$14.15	0.5%	$14.03

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

Table C.2: Estimates of the Effects of the Joint Venture on Flagship Brand Prices.

Dep. Var.	(1) $ Price/ 12 Pack	(2) Log(Price/ 12 Pack)	(3) $ Price/ 12 Pack	(4) Log(Price/ 12 Pack)	(5) $ Price/ 12 Pack	(6) Log(Price/ 12 Pack)
Price Reductions	incl.	incl.	incl.	incl.	excl.	excl.
Post-JV Brand Dummies						
Bud Light	0.9047	0.0461	0.4754	0.0488	0.4794	0.0436
	(0.0555)	(0.0052)	(0.0808)	(0.0062)	(0.1208)	(0.0090)
Miller Lite	1.0594	0.0628	0.4265	0.0444	0.4574	0.0417
	(0.0581)	(0.0056)	(0.0827)	(0.0066)	(0.1131)	(0.0084)
Coors Light	1.0435	0.0616	0.5285	0.0566	0.6007	0.0581
	(0.0631)	(0.0060)	(0.0830)	(0.0070)	(0.1095)	(0.0081)
Brand Time Trends	N	N	Y	Y	Y	Y
Observations	126,872	126,872	126,872	126,872	126,769	126,769
R^2	0.9578	0.9599	0.9588	0.9601	0.9643	0.9651

Notes: the reported coefficients are on domestic brand × post-JV interactions. The brands included are those listed, plus Corona Extra and Heineken. Observations at the market-week-brand level, aggregating across packages containing the equivalent of 6, 12, 18, 24, 30 or 36 12oz. containers in cans or glass bottles, and prices are converted into 12-pack equivalents. All specifications include market-brand and week fixed effects. The period from the announcement of the JV to its consummation is excluded. Standard errors in parentheses clustered on the market.

An observation in these regressions is a market-week-brand, where real prices are calculated at the brand level by adding up the total sales in package sizes equivalent to packs of 6, 12, 18, 24, 30 or 36 12oz. containers. The sample contains the following brands: BL, ML and CL (i.e., the domestic flagship brands), Corona Extra and Heineken which, following MW, we will treat as providing controls for industry-wide shocks. The sample runs from 2001 to 2011, but excludes the period between announcement and consummation of the JV. We consider prices defined using all store-UPC-week observations in the appropriate sizes, and prices that are defined excluding store-UPC-week observations that are identified as being sold at temporary price reduction prices.

Table C.2 presents the results from six specifications that differ in whether prices are measured in levels or logs, and whether brand-time trends and price reductions are included. The reported coefficients are the coefficients on post-JV dummies for the domestic flagship brands, so that they measure the increase in real prices relative to the two imported brands. The estimated price increases vary across the columns, but lie in the range from just over 40 cents to just over one dollar, or 4% to 6%.

C.4 Demand Assumptions for the Calibration.

Our baseline demand specification for our calibration assumes that, at observed average pre-merger prices, the representative domestic flagship brands have an average own-price elasticity of -3 and that a small price increase of one of the brands will lead, on average, to 85% of the lost demand going to the other two brands. In this section we explain why these values, and, in particular, the assumed high rate of diversion between the domestic brands, are justified.

One approach is to estimate demand models and see the elasticities and diversions that they imply. We estimate a number of demand models using the same data and instruments as MW, appropriately adjusted when we use alternative nesting specifications.[9] Observations are market-time period-brand-pack size where the pack sizes are 6, 12 and "large", which combines 24 and 30-packs, and there are 13 included brands. The demand models are estimated using the pyBLP program (Conlon and Gortmaker (2020)).

Table C.3 reports eight sets of demand estimates. The first four columns use monthly data, which is the shortest time period considered by MW. Two specifications are nested logit and two are random coefficient nested logit where preferences can vary with income. One nesting structure follows MW in placing all brands in one nest (with the outside good in a separate nest) and our new alternative places the flagship brands (the focus of our calibration) in their own nest, with the other 10 brands in a second nest, with possibly different nesting coefficients. The last four columns report estimates for weekly specifications. The table also reports statistics on diversion and elasticities.

The implied elasticities and diversions are quite heterogeneous, leading to the question of which estimates should be preferred. We choose to use estimates that imply that there is limited diversion from domestic/flagship brands as this captures what we observe happening to brand shares after the JV, as matching this pattern is obviously what is most relevant for firm incentives in our application.

Figure C.1 shows volume-based market shares of the different brands included in the demand analysis (for this purpose, we define market shares based on the shares of **all** beers in the IRI data, not just the ones that MW include in their demand model). We aggregate non-flagship brands based on their pre-JV ownership. While there is some shift in share from ML to CL (a change seemingly not driven by price changes, as the regressions in Appendix C.3 indicate that ML and

[9] For example, when we divide brands into different nests, we define instruments that are specific to the products in the nest.

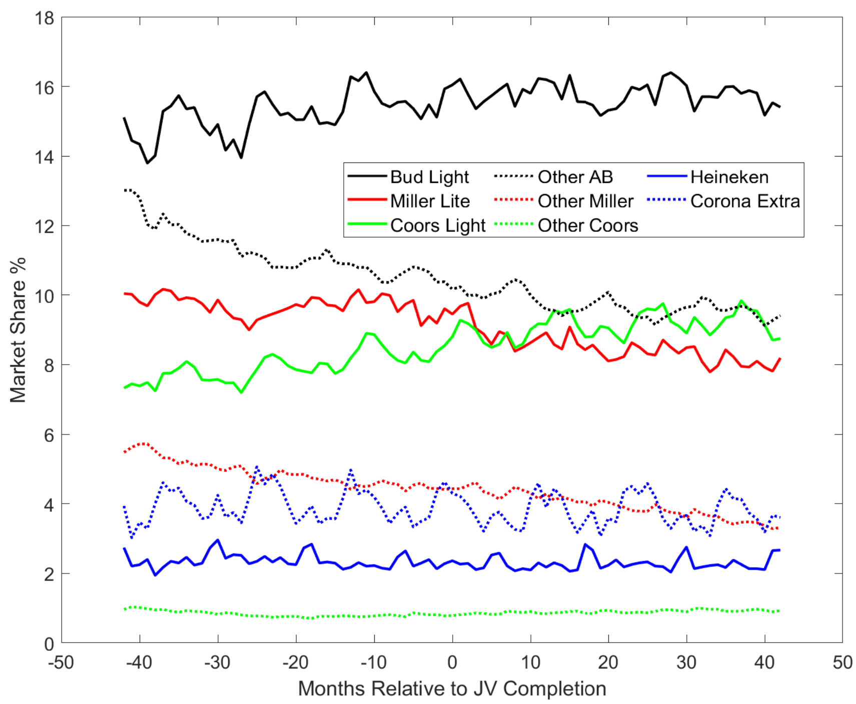
Table C.3: Estimates of Demand Parameters for Alternative Models

	(1) NL One Nest Month	(2) NL Two Nest Month	(3) RCNL One Nest Month	(4) RCNL Two Nest Month	(5) NL One Nest Week	(6) NL Two Nest Week	(7) RCNL One Nest Week	(8) RCNL Two Nest Week
Real Price	-0.058	-0.074	-0.084	-0.157	-0.023	-0.033	-0.054	-0.129
	(0.017)	(0.018)	(0.021)	(0.022)	(0.013)	(0.015)	(0.016)	(0.019)
Single All Brand Nest	0.735		0.836		0.706		0.851	
	(0.050)		(0.039)		(0.062)		(0.038)	
Two Nests								
Domestic Flagship		0.829		0.735		0.905		0.802
		(0.049)		(0.070)		(0.028)		(0.052)
Other Brands		0.633		0.760		0.627		0.731
		(0.046)		(0.030)		(0.053)		(0.032)
Income Coefficients (RCNL)								
* constant			0.015	0.008			0.016	0.007
			(0.005)	(0.006)			(0.005)	(0.006)
* price			0.001	0.002			0.000	0.002
			(0.000)	(0.000)			(0.000)	(0.000)
* calories			0.004	0.015			0.005	0.017
			(0.002)	(0.003)			(0.002)	(0.003)
Median product (brand-size) elasticity	-2.34	-2.53	-4.72	-4.81	-0.82	-1.13	-3.33	-3.74
Mean ML brand elasticity	-1.67	-2.92	-3.63	-3.34	-0.60	-2.29	-2.58	-3.41
Mean flagship diversion	0.39	0.82	0.47	0.72	0.38	0.90	0.48	0.78
Observations	94,656	94,656	94,656	94,656	405,003	405,003	405,003	405,003

Notes: all specifications include time period and product (brand*size) fixed effects, and use data from Jan 2005 to Dec 2011, excluding June 2008 to May 2009. All estimates use two-step optimal GMM. NL=Nested Logit. Instruments are the same as in MW for the relevant specification, apart from the two nest models where we define instruments for the number and distance measures for other products based on products in the same nest, and interact instruments with a flagship brand dummy. Market size is defined as 50% more than the highest sales observed in the geographic market for monthly specifications. For the weekly specifications, this market size is divided by 4.5. ML Brand Elasticity reflects the change in ML sales when the prices of all ML products are increased. Mean Flagship Diversion is the average proportion of lost sales that go to other flagship products (i.e., BL, ML and CL products) when the price of a flagship product is increased. Standard errors, clustered on the geographic market, in parentheses.

CL prices increased by very similar amounts), the shares of other brands appear to be stable or to follow paths that are consistent with pre-JV trends. In particular, there is no upwards jump in the shares of the imported brands, nor a clear decrease in the shares of the large domestic brewers (as might happen, for example, if there was limited substitution to imports but significant substitution to domestic craft beers).

Figure C.1: Brand Market Shares Around the Consummation of the Joint Venture



Notes: Budweiser, Michelob Ultra and Michelob Light aggregated into "Other AB"; Miller Genuine Draft and Miller High Life aggregated to "Other Miller"; Coors is "Other Coors"; Heineken and Heineken Premium Light are "Heineken" and Corona Extra and Corona Light are "Corona". Shares based on volume sold in packages equivalent to 6, 12, 18, 24, 30 and 36 12oz glass bottles and aluminum cans, and calculated at the monthly level.

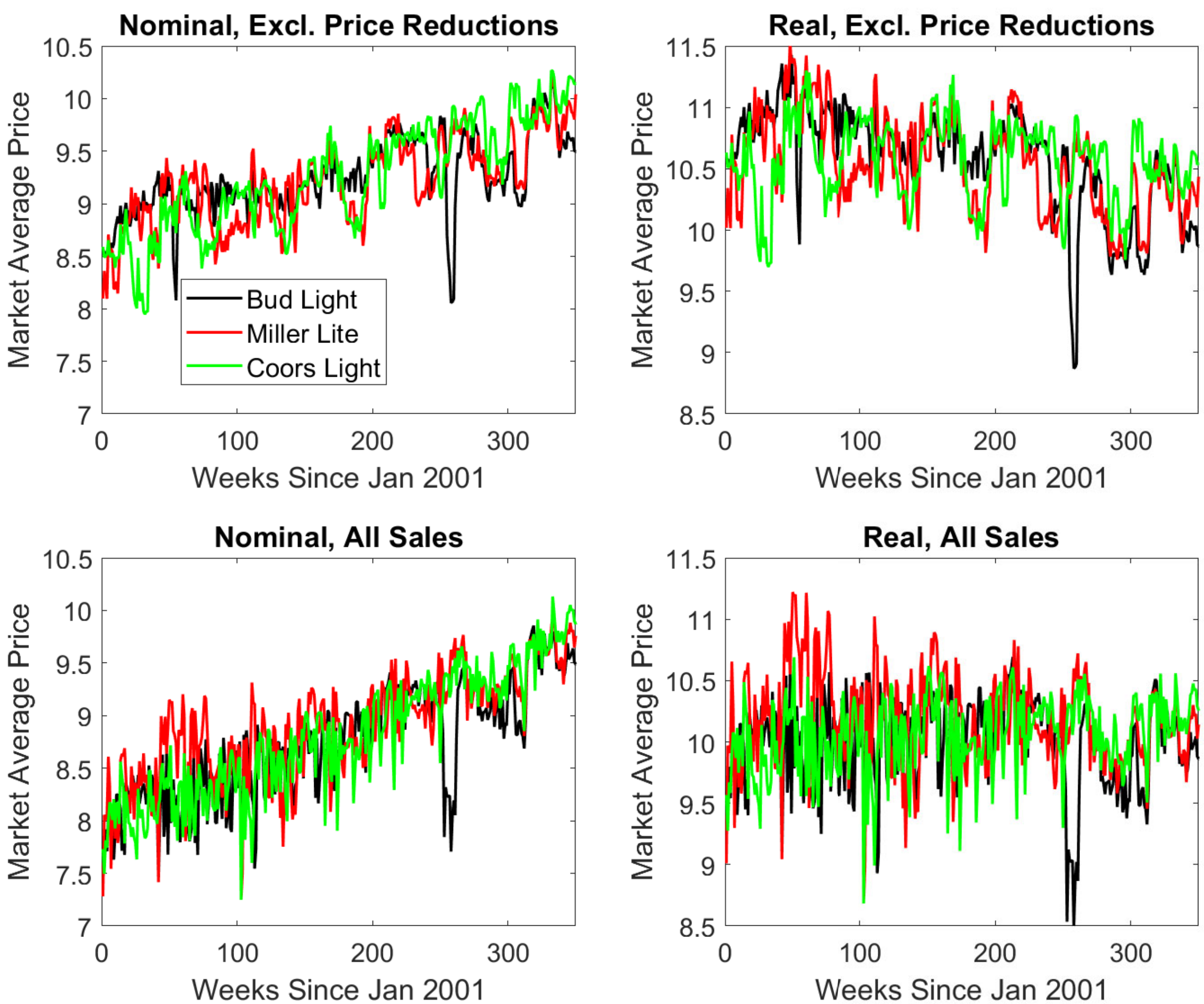
C.5 Pre-Merger Price Dynamics.

In this appendix we illustrate the price dynamics in the pre-JV data which we try to match when calibrating the parameters.

Figure C.2 shows weekly price series for the three flagship brands in the Seattle market for 12-packs. The panels show average prices (total revenues divided by units sold) calculated in four slightly different ways.

- top-left: nominal prices, excluding sales made at what the IRI data indicate are temporary store price reductions;
- top-right: real prices, excluding sales made at what the IRI data indicate are temporary store price reductions;
- bottom-left: nominal prices, including all sales;
- bottom-right: real prices, including all sales.

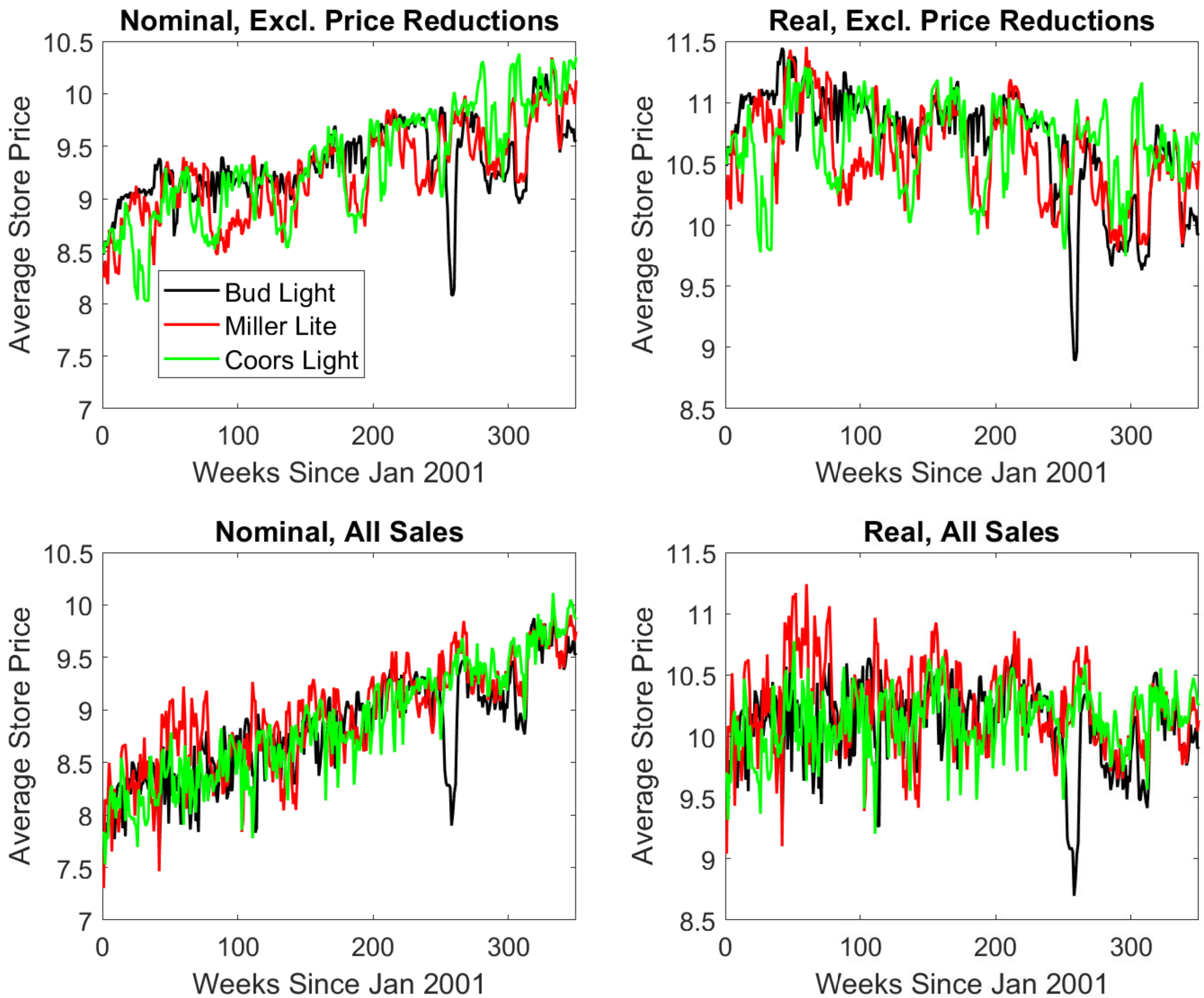
Figure C.2: Average Weekly Prices in the Seattle Market for 12-Packs of Flagship Brands in Seattle from January 2001 to the Annoucement of the JV. Prices are calculated based on total revenues in Seattle, divided by the total number of 12-packs.



Week-to-week variation in prices, with positive serial correlation, are clear features of each series. The series that include temporary price reductions are, unsurprisingly, more volatile and have more frequent large reductions lasting only one or two weeks.

One might wonder how much of the volatility reflects demand shifts across stores that tend to have different prices, rather than changes in actual prices. Figure C.3 shows the four series with market-week average prices computed as the unweighted (rather than volume-weighted) average of prices across the stores in the sample. These price series are similar to those in Figure C.2, indicating that the volatility is not driven primarily by demand shifts.

Figure C.3: Average Weekly Prices in the Seattle Market for 12-Packs of Flagship Brands in Seattle Prior to the Announcement of the JV. Prices are calculated as the unweighted average of the price in each store in the IRI sample in that week.



Our calibration matches coefficients from regressions where market-week-pack size average prices prior to the JV are regressed on the lagged prices of all three flagship brands. These regressions are done market-by-market, with the averages of the market coefficients being matched. Here we report the results of regressions that pool the data from different markets together to give a sense of the coefficients that are being matched and how they vary with exactly how the price series are defined. The fixed effects that are included will control for cross-market differences in prices.

The top-left panel of Table C.4 shows the coefficients where the weekly-brand-pack size-market price series are calculated excluding temporary price reductions. This specification is closest to matching the regressions used in our baseline calibration. The regressions include fixed effects for the set of stores observed in the market-week (interacted with pack size) and week-pack size dummies (which control for national promotions). The own lagged price coefficients are between 0.4 and 0.5, and are highly significant, whereas the cross-brand lagged coefficients are positive but smaller. The panel also reports mean prices and the standard deviation of the residuals from the price regression. The standard deviations are one of the non-targeted moments that we use to assess model fit.

The other panels use alternative price series that correspond to alternatives that we use to assess the robustness of our baseline calibrations. The top-right panel uses data only on 12-packs, a widely sold pack size for the flagship brands. The own brand lagged price coefficients become slightly larger. The bottom-left panel uses series that are calculated to include temporary price reductions. As promotions involve large price drops that may only last one week, the own price serial correlation coefficients fall significantly. The standard deviation of the residuals also doubles.

One might be concerned that using weekly data introduces noise into the price series that is inappropriate if brewers set prices less frequently. MW's preferred specifications assume monthly price-setting. If we estimate the specification from the top-left panel but using monthly average prices, the estimated own-brand lagged price coefficients fall (for example, for BL prices, the p^{BL}_{t-1} coefficients falls to 0.318). However, because we are using set-of-store fixed effects and the set of observed stores varies within a month for many markets, the sample size in the pooled regressions drops dramatically (to 2,806 observations), and some markets are dropped entirely. To preserve the sample size, it therefore makes sense to use market fixed effects (interacted with pack size). In this case, the estimated lagged price coefficients increase (for example, the own-brand lagged price coefficient for BL increases to 0.646), as cross-store differences in prices and the entry and exit of stores from the IRI sample over time makes prices appear to be more correlated than when we control for the set of stores. In our calibration, matching these higher values of the serial correlation coefficients requires that our parameters imply very strong signaling incentives, and in our counterfactuals, the conditions required for a separating equilibrium to exist are not satisfied. Our monthly specification therefore includes temporary price reductions in the price series (MW's monthly data also includes price reductions) which has an offsetting effect, and the associated coefficients are reported in the bottom-right panel.

Cross-Market Variation in the Estimated Pre-JV Price Dynamics and Market Concentration. As noted, the coefficients matched in the calibration are cross-market averages. It is, however, interesting to consider whether cross-market differences in the lagged price coefficients are potentially consistent with a signaling story. When we simulate data from our example specifications, we see that, holding the cost parameters fixed, there is greater serial correlation, both within and across brands, when we choose parameters that imply stronger signaling incentives. This includes increasing the extent of demand diversion between the firms that are playing the game.

If it is the three largest domestic brewers that play the game, we would expect higher diversion between the players when the combined market share of these firms is larger, all else equal. If our model is correct, we might therefore expect a positive relationship between this combined share and market-level serial correlation coefficients.

Figure C.4 presents these relationships as scatter plots, with univariate regression coefficients in the notes beneath the figure. For both own-brand and cross-brand coefficients there are positive relationships that are at least marginally statistically significant, consistent with our expectation.

C.6 Pass-Through of Distribution Costs to Price Before and After the JV.

In the text, we suggest that a pattern where prices increased more after the JV in markets where trucking distances are larger is consistent with our model. In this Appendix, we explain in more detail our rationale for this prediction and present our evidence that the predicted pattern holds in the data.

Our logic is the following: suppose that the underlying state variable that is private information is the "per mile per unit" efficiency of the firm's distribution network, and that this efficiency lies on the interval $[\underline{e}, \overline{e}]$. The privately observed marginal cost of a unit of beer sold in market m by brewer i will be the value of this state variable multiplied by the distance from the brewery to the

Table C.4: Pre-JV AR(1) Price Regressions Using Flagship Market-Week-Pack Size or Market-Month-Pack Size Data.

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

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

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

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

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

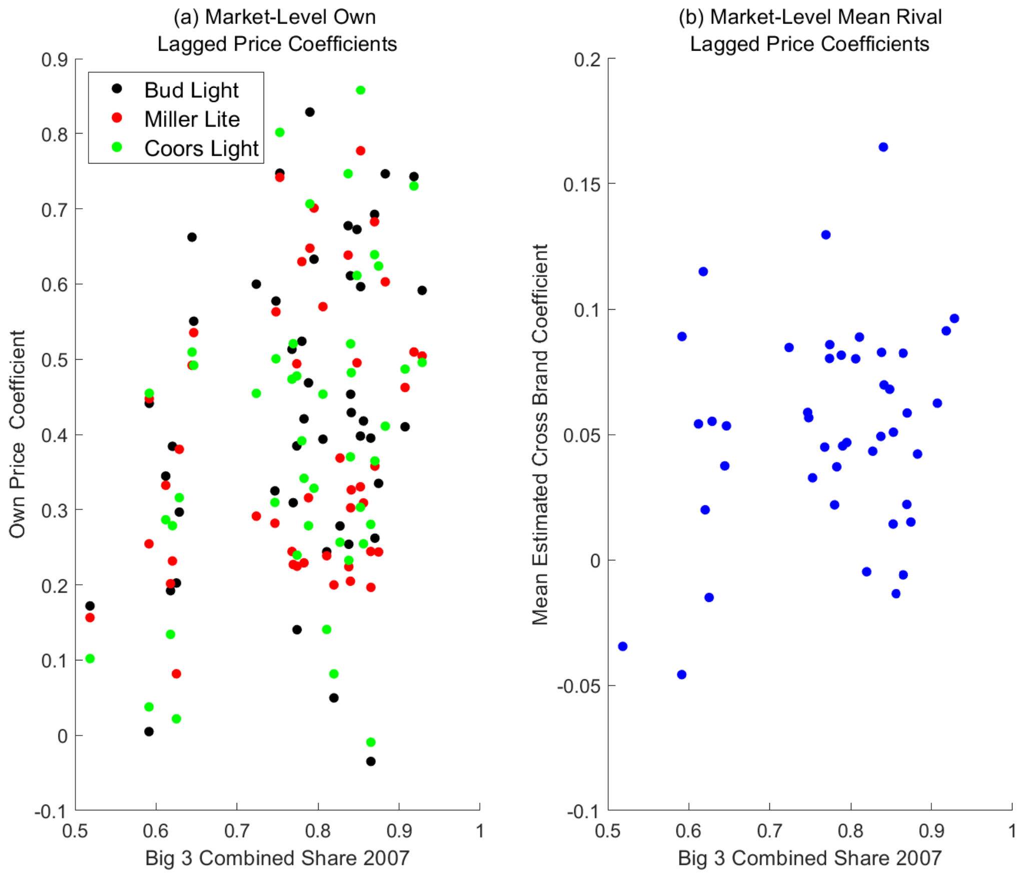
	(1)	(2)	(3)
	$p_{BL,t}$	$p_{ML,t}$	$p_{CL,t}$
$p_{BL,t-1}$	0.287	0.036	0.020
	(0.027)	(0.013)	(0.013)
$p_{ML,t-1}$	0.045	0.322	0.010
	(0.009)	(0.027)	(0.012)
$p_{CL,t-1}$	-0.023	-0.049	0.267
	(0.013)	(0.020)	(0.039)
Observations	37,449	37,431	37,442
R-squared	0.939	0.941	0.942
Mean Price ($)	9.79	9.67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

(d) Month, Price Reductions Included, Five Pack Sizes, Fixed Effects for Markets

	(1)	(2)	(3)
	$p_{BL,t}$	$p_{ML,t}$	$p_{CL,t}$
$p_{BL,t-1}$	0.386	0.106	0.114
	(0.037)	(0.019)	(0.018)
$p_{ML,t-1}$	0.093	0.385	0.061
	(0.019)	(0.031)	(0.020)
$p_{CL,t-1}$	0.147	0.115	0.498
	(0.017)	(0.017)	(0.028)
Observations	14,022	14,029	14,034
R-squared	0.944	0.945	0.948
Mean Price ($)	9.73	9.61	9.63
SD residuals ($)	0.329	0.337	0.321

Notes: regressions also include time period-pack size interactions. The five pack sizes are 6, 12, 18, 24 and 30 packs. Data from January 2001 to the announcement of the JV. Market or store fixed effects described in the label to each panel. Standard errors, clustered on the market, are in parentheses. The SD residuals statistic is the standard deviation of the residuals from the regression.

Figure C.4: Cross-Market Variation in the Estimated Pre-JV Price Dynamics and Market Concentration.



Notes: The estimated univariate regression coefficients (i.e., best fit linear relationship), with standard errors in parentheses, for panel (a) are BL: -0.027 (0.203) + 0.588C_3 (0.296), R^2 = 0.084; ML : 0.039 (0.227) + 0.450C_3 (0.290), R^2=0.066; CL : -0.172 (0.232) + 0.731C_3 (0.290), R^2=0.129; and for panel (b): -0.035 (0.054) + 0.117C_3 (0.069), R^2=0.064. 45 observations in each regression. The independent variable is the combined (volume) market share of AB, Miller and Coors brands in the market in 2007 (mean 0.771, std. dev. 0.102). The coefficients themselves come from a regression of the market-week-brand-pack size price on the lagged prices of the three flagship brands, with market-brand-pack size fixed effects and a linear time trend.

market, which should be observable to all firms. The per-unit marginal cost would then lie on the interval $[d_{i,m} \times \underline{e}, d_{i,m} \times \overline{e}]$, where $d_{i,m}$ is the trucking distance. Some examples suggest that the effects of dynamic signaling on equilibrium average prices tends to be larger when widths are wider, and that these effects will be more pronounced after a merger. We would therefore expect larger post-JV increases in average prices in those markets that are further from breweries.

Table C.5 reports the results of regressions where the dependent variable is the price measured in real dollars (equivalent regressions using log prices are reported in Table C.6). The sample includes the 39 markets and 13 domestic and imported brands that MW include in their demand analysis (for which we can use MW's distance variables) and all six common pack sizes. We exclude the period for one year following the JV in all regressions. Observations are at the market-week-brand-pack size level and the dependent variable is the real price (in dollars) per 12-pack equivalent. For comparison purposes, it is useful to keep in mind that the JV increased average real domestic prices by 40 cents to one dollar per 12-pack. All regressions include date fixed effects, and the reported specifications vary in how distances are measured and the combination of product and market fixed effects that are included.

The first column estimates how a brewer's own brewery-to-market trucking distance, measured in 1,000 miles, affects the brewer's prices before and after the JV, controlling for pre-/post-JV product (i.e., pre/post-brand-size) fixed effects, market fixed effects and week fixed effects. We estimate pre- and post-JV distance coefficients for imported products and each domestic brewer (based on pre-JV ownership). The coefficients show that after the JV there was a stronger, positive relationship between domestic brand prices and trucking distances, with no change in the relationship for imported brands, which, like MW, we would interpret as providing controls. The changes are also quite large: the average post-JV distance for Miller is 316 miles with standard deviation 269 miles, and a range of over 1,000 miles, so that, comparing markets with one standard deviation difference in distance, the coefficients imply that Miller-branded prices would go up 21 cents more in the more distant market, which is a large effect.

In specification (2), the distance variable is the trucking distance multiplied by the real price of diesel (which should also be observable). The coefficients change in scale, but the pattern remains the same.[10] The remaining columns use distance-only based measures, but the results are qualitatively similar using distance-diesel price interactions. Column (3) includes product-market

[10]Real diesel prices were quite flat two years before and after the JV.

Table C.5: Distance Pass-Through Regressions: Real Price Per 12-Pack for all Pack Sizes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Distance Measure	Own	Own × Diesel	Own	Min. Domestic	-	Own	Min. Domestic
Distance Measure ×							
AB	-0.276	-0.040					
	(0.185)	(0.067)					
Imports	0.451	0.088					
	(0.181)	(0.040)					
Coors	0.245	0.131					
	(0.069)	(0.020)					
Miller	-0.083	0.032					
	(0.093)	(0.035)					
Post-JV × Own Distance ×							
AB	0.619	0.189	0.537	0.819		0.274	0.565
	(0.283)	(0.091)	(0.302)	(0.227)		(0.292)	(0.271)
Imports	-0.166	-0.040	-0.107	0.294		-0.174	0.071
	(0.183)	(0.041)	(0.154)	(0.245)		(0.159)	(0.289)
Coors	0.451	0.116	0.692	0.813		0.564	0.552
	(0.145)	(0.046)	(0.133)	(0.186)		(0.117)	(0.233)
Miller	0.694	0.194	0.704	0.886		0.579	0.625
	(0.137)	(0.038)	(0.145)	(0.207)		(0.122)	(0.240)
Post-JV ×							
Reduction Coors Distance					-0.320	-0.372	-0.305
					(0.190)	(0.189)	(0.205)
Mkt HHI Increase Due to JV					3.735	3.556	3.505
					(3.215)	(3.145)	(3.448)
Post-JV × Domestic ×							
Reduction Coors Distance					-0.135	0.036	-0.023
					(0.128)	(0.135)	(0.142)
Mkt HHI Increase Due to JV					3.656	3.231	2.093
					(3.845)	(3.564)	(3.349)
Fixed Effects	Pre/Post× Product Market	Pre/Post× Product Market	Product× Market Pre/Post× Product	Product× Market Pre/Post× Product	Product× Market Pre/Post× Product	Product× Market Pre/Post × Product	Product× Market Pre/Post× Product
Observations	869,018	869,018	869,018	869,018	869,018	869,018	869,018
R-squared	0.923	0.922	0.947	0.947	0.947	0.947	0.947

Notes: standard errors, clustered on the market, in parentheses. See text for discussion of the sample. Date fixed effects in all specifications. Distances are measured in thousands of miles, and HHIs are measured between 0 and 1.

fixed effects, with the pattern of the post-JV coefficients unchanged.[11]

As noted in Section 2.3, price increases in our examples suggest that signaling price increases tend to be most sensitive to the smallest $\overline{c_i} - \underline{c_i}$ among the players. Column (4) therefore uses the smallest trucking distance for the domestic brewers as the distance measure for all brewers. The magnitude of the post-JV AB coefficient increases in size and statistical significance, with the coefficients for Miller and Coors products also increasing.

MW (Table III) also estimate regressions to understand cross-market differences in post-JV price increases. Using log price as the dependent variable, they find a positive relationship between price increases and the increase in concentration caused by the JV, and a negative relationship with the reduction in trucking distances for Miller and Coors brands that results from MC producing beers in the closest brewery, although they also find that the latter has some effect on imported brand prices. Column (5) presents a similar analysis for our sample.[12] The concentration levels are higher in our analysis as we include pack sizes that are sold in significant volumes only by the domestic brewers. When prices are in logs, we see qualitatively similar patterns to MW, although in levels the relationships are not significant. The remaining columns add these variables to the specifications in columns (3) and (4). We see that the change in distance pass-through pattern identifed in columns (3) and (4) is robust to including these variables, whereas in the changes in HHI and MC trucking distance have no additional effect for domestic brands, even in the log price regressions.

C.7 Brand Price Correlations Before and After the JV

In our calibration, we assume that each firm sets only a single price, which we would like to interpret as reflecting the prices that it is setting for its product portfolio. Our simplification is more plausible if the prices of the products within a portfolio are highly correlated. Similarly, as we assume that MC sets exactly the same price for both the representative Miller product, Miller Lite, and the representative Coors product, Coors Light, after the JV, we would like to see that these products have highly correlated prices in the post-JV period.

Table C.7 reports the pairwise correlations of market-week average prices of 12-packs of the flagship brands, plus Budweiser, Miller Genuine Draft and Coors, before and after the JV. The data

[11] In the MW data there are handful of small distance changes for markets before the JV. The post-JV distance coefficients are unchanged if additional pre-JV distance coefficients are estimated.

[12] MW estimate a separate regression for each brewer, and they only include time trends rather than date fixed effects.

Table C.6: Distance Pass-Through Regressions: Log(Real Price Per 12-Pack) for all Pack Sizes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Distance Measure	Own	Own × Diesel	Own	Min. Domestic	-	Own	Min. Domestic
Distance Measure ×							
AB	-0.022	-0.003					
	(0.016)	(0.006)					
Imports	0.039	0.011					
	(0.015)	(0.003)					
Coors	0.026	0.013					
	(0.006)	(0.002)					
Miller	0.003	0.006					
	(0.010)	(0.004)					
Post-JV × Own Distance ×							
AB	0.052	0.015	0.044	0.076		0.018	0.050
	(0.029)	(0.009)	(0.030)	(0.022)		(0.029)	(0.027)
Imports	-0.010	-0.005	-0.007	0.023		-0.013	0.006
	(0.013)	(0.003)	(0.011)	(0.017)		(0.011)	(0.020)
Coors	0.054	0.014	0.069	0.078		0.056	0.051
	(0.015)	(0.005)	(0.014)	(0.020)		(0.012)	(0.024)
Miller	0.076	0.021	0.076	0.094		0.063	0.067
	(0.015)	(0.004)	(0.016)	(0.023)		(0.013)	(0.026)
Post-JV ×							
Reduction Coors Distance					-0.028	-0.032	-0.027
					(0.014)	(0.014)	(0.014)
Mkt HHI Increase Due to JV					0.234	0.221	0.216
					(0.227)	(0.221)	(0.246)
Post-JV × Domestic ×							
Reduction Coors Distance					-0.018	-0.002	-0.007
					(0.010)	(0.011)	(0.012)
Mkt HHI Increase Due to JV					0.500	0.468	0.345
					(0.326)	(0.302)	(0.274)
Fixed Effects	Pre/Post× Product Market	Pre/Post× Product Market	Product× Market Pre/Post× Product	Product× Market Pre/Post× Product	Product× Market Pre/Post× Product	Product× Market Pre/Post × Product	Product× Market Pre/Post× Product
Observations	869,018	869,018	869,018	869,018	869,018	869,018	869,018
R-squared	0.915	0.915	0.942	0.942	0.943	0.943	0.943

Notes: standard errors, clustered on the market, in parentheses. Date fixed effects in all specifications.

Table C.7: Cross-Brand Correlations in Prices for 12-Packs

		Pre-JV					
		(1)	(2)	(3)	(4)	(5)	(6)
(1)	Bud Light	1					
(2)	Miller Lite	0.806	1				
(3)	Coors Light	0.806	0.831	1			
(4)	Budweiser	0.990	0.802	0.806	1		
(5)	Miller Genuine Draft	0.791	0.963	0.819	0.787	1	
(6)	Coors	0.772	0.795	0.959	0.774	0.789	1
		Post-JV					
		(1)	(2)	(3)	(4)	(5)	(6)
(1)	Bud Light	1					
(2)	Miller Lite	0.865	1				
(3)	Coors Light	0.887	0.940	1			
(4)	Budweiser	0.995	0.863	0.882	1		
(5)	Miller Genuine Draft	0.842	0.960	0.914	0.841	1	
(6)	Coors	0.846	0.908	0.956	0.841	0.885	1

Notes: the reported values are the pairwise correlations between market-week-brand average nominal prices of 12-packs, before the announcement of the JV and after its consummation. Average prices are calculated including price reductions. Correlations for brands with the same owner are slightly higher if price reductions are excluded.

come from 45 markets, and market-weeks where at least 5 stores carrying 12-packs are observed.

As we would like, the prices of products with the same owner (e.g., BL and Budweiser, or ML and Miller Genuine Draft) are particularly highly correlated, and the prices of Miller and Coors products become more highly correlated after the JV. Some of this correlation reflects beer being sold at different prices in different markets. To control for cross-market effects, we can also calculate correlations by regressing the nominal price of one brand on the price of another brand, and market and week fixed effects. These results also show significant increases in correlations of Miller and Coors products after the JV: for example, the coefficient on the CL price when the ML price is the dependent variable increases from 0.61 before the JV to 0.81 after the JV. Patterns in the table and the regressions are similar if we use prices defined to exclude temporary price reductions, or if we use real, rather than nominal, prices.

References

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

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

Cho, In-Koo, and David M. Kreps. 1987. "Signaling Games and Stable Equilibria." *Quarterly Journal of Economics*, 102(2): 179–221.

Conlon, Christopher, and Jeff Gortmaker. 2020. "Best Practices for Differentiated Products Demand Estimation with pyBLP." *RAND Journal of Economics*, 51(4): 1108–1161.

Hubbard, Timothy P, and Harry J Paarsch. 2013. "On the Numerical Solution of Equilibria in Auction Models with Asymmetries within the Private-Values Paradigm." In *Handbook of Computational Economics*. Vol. 3, , ed. Karl Schmedders and Kenneth L. Judd, 37–115. Elsevier: North-Holland.

Judd, Kenneth L. 1998. *Numerical Methods in Economics.* MIT Press.

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